

# Histogram of Oriented Gradients for Image Mosaicing

Ridhi Arora and Parvinder Singh

**Abstract** Image Stitching can be defined as a process of combining multiple input images or some of their features into a single image without introducing distortion or loss of information. The aim of stitching is to integrate all the imperative information from multiple source images to create a fused mosaic. Therefore, the newly generated image should contain a full description of the scenery than any of its individual source images and is more suitable for human visual and machine perception. With the recent rapid developments in the domain of imaging technologies, mosaicing has become a reality in wide fields such as remote sensing, medical imaging, machine vision, and the military applications. Image Stitching therefore provides an effective way by reducing huge volume of information by extracting only the useful information from the source images and blending them into a high-resolution image called mosaic.

**Keywords** Image mosaicing · Image reconstruction from gradient measurements  
RANSAC estimation · Fundamental matrix · Geometric transformation  
HoG

## 1 Introduction

There are many applications where images with larger field of view are of great importance. In areas ranging from medical imaging to computer graphics and satellite imagery, a computationally effective and easy to implement a method to produce high-resolution wide angle images will continue to draw research interest. This production of wide angle image from source images is called Image Stitching. Research into algorithmic image stitching requires image registration, alignment, calibration, and blending which is one of the oldest topics in computer vision. Image stitching for the purposes of mosaic creation can be used by amateur pho-

---

R. Arora (✉) · P. Singh  
Department of CSE, DCRUST, Sonapat (Murthal), Haryana, India  
e-mail: ridhiarora584@gmail.com

tographers and professionals without requiring detailed knowledge, using digital photography software such as Photoshop that provides easy-to-use instructions and interfaces. Depending on the scene content, the future panorama, and luminance, two major steps are being followed. The first is the registration, that is to be done on the source images, and second is choosing an appropriate blending algorithm.

Registration is the process of spatially aligning images. This is done by choosing one of the images as the reference image and then finding geometric transformations which map other images on to the reference frame. Upon completion of this step, an initial mosaic is created which is made by simply overlapping the source images by the common region in them. This helps in reducing visual and ghosting artifacts. Various registration techniques are described in [1].

The initial mosaic may contain visual and ghosting artifacts because of varying intensities in each of the source images. They can be significantly reduced by choosing an appropriate blending algorithm through which transition from one image to other becomes imperceptible. Good blending algorithms should produce seamless and plausible mosaics in spite of ghosting artifacts, noise, or lightning disturbances in the input images. Previously, proposed blending algorithms work on image intensities or image gradient, at full resolution of the image or at multiple-resolution scales.

Image Stitching is performed by adopting techniques such as feature-based and direct techniques. Direct techniques also called as area-based technique [2] minimize the sum of absolute differences between the overlapping pixels. This technique is computationally ineffective as it takes every pixel window into consideration. The use of rectangular window to find the similar area limits the application of these methods. This method is generally used with applications that have small translation and rotation because they are not invariant to image scale and rotation [3]. The main *advantage* of direct technique is that they incorporate every pixel value which optimally utilizes the information in image alignment but fail to correctly match the overlapping region due to limited range of convergence. On the other hand, feature-based technique works by extracting a sparse set of essential features from all the input images and then matching these to each other [4]. Feature-based methods are used by establishing correspondences between points, lines, edges, corners, or any other shapes. The main characteristics of robust detectors include invariance to scaling, translation, image noise, and rotation transformations. There are many feature detector techniques existing some of which are Harris corner detection [5], Scale-Invariant Feature Transform (SIFT) [6, 7], Speeded Up Robust Feature (SURF) [3, 8], Features from Accelerated Segment Test (FAST) [9], HoG [10], and ORB [11].

The use of image stitching in real-time applications has proved to be a challenging field for image processing. Image stitching has found a variety of applications in microscopy, video conferencing, video matting, fluorography, 3D image reconstruction, texture synthesis, video indexing, super resolution, scene completion, video compression, a satellite imaging, and several medical applications. Stitched images (mosaics) are also used in topographic mapping. For videos, additional challenges are imposed on image stitching. As videos require motion of

pictures with varying intensities so, camera feature-based techniques which aim to determine a relationship zoom, and to visualize dynamic events impose between the images is used which poses additional challenges to image stitching.

## 2 Related Work

Image Stitching is a process of combining source images to form one big wide image called mosaic irrespective of visual and ghosting artifacts, noise addition, blurring difficulties, and intensity differences. For stitching, many techniques have been proposed in many directions and fields. They include gradient field, stitching images in mobiles, SIFT algorithm, SURF algorithm, Haar wavelet method, corner method, and many more.

A well-known intensity-based technique is *feathering* [12], where the mosaic is generated by computing a weighted average of the source images. In the composite mosaic image, pixels are assigned weights proportional to their distance from the seam, resulting in a smoother transition from one image to the other in the final mosaic. But this can introduce blurring, noise, or ghosting effect in the mosaic when the images are not registered properly. Therefore, the intensity-based multiscale method is presented that relies on *pyramid decomposition*. The source images are decomposed into band-pass components, and the blending is done at each scale, in a transition zone inversely proportional to the spatial frequency content in the band.

Gradient field is also exclusively utilized for mosaicing. Sevcenco et al. [13] presented a gradient-based image stitching technique. In this algorithm, the gradients of the two input images are combined to generate a set of gradients for the mosaic image and then to reconstruct the mosaic image from these gradients using the Haar wavelet integration technique.

Further, advancements are made in the field of gradients using Poisson image editing. Perez et al. [14] proposed Poisson image editing to do seamless object insertion. The mathematical tool used in the approach is the Poisson partial differential equation with Dirichlet boundary conditions with the Laplacian of an unknown function over the domain of interest, along with the unknown function values over the boundary of the domain. The actual pixel values for the copied region are computed by solving Poisson equations that locally match the gradients while obeying the fixed Dirichlet condition at the seam boundary. To make this idea more practical and easy to use, Jia et al. [15] proposed a cost function to compute an optimized boundary condition, which can reduce blurring. Zomet et al. [16] proposed an image stitching algorithm by optimizing gradient strength in the overlapping area. Here, two methods of gradient domain are discussed. One optimizes the cost functions, while others find derivative of the stitched image. Both the methods produce good results in the presence of local or global color difference between the two input images.

Brown and Lowe [17] use SIFT algorithm to extract and match features in the source images that are located at scale-space maxima/minima of a difference of

Gaussian function. After successful features are mapped, RANSAC algorithm is applied to remove all the unnecessary outliers but include the necessary inliers that are compatible for Homography between the images. Afterward, bundle adjustment [18] is used to solve for all the camera parameters jointly.

Zhu and Wang [19] proposed an effective method of mosaic creation that includes sufficient iterations of RANSAC to produce faster results. For this, multiple constraints on their midpoint, distances, and slopes are applied on every two candidate pairs to remove incorrectly matched pairs of corners. This helps in making effective panoramas with least no RANSAC iteration. Stitched images (mosaics) are also used in topographic mapping and stenography [20]. Stenography is a technique that is used to hide information in images. Various other operations can also be applied on images being stitched [21–24].

Various other techniques are also introduced that either used area-based or feature-based techniques, but an effective technique is introduced in this paper, viz. HoG (Histogram of Oriented Gradients).

### 3 Proposed Methodology

The main feature for selection and detection used here is HoG (Histogram of Oriented Gradients) [25]. The method is based on evaluating well-normalized local histograms of image gradient orientations in a dense grid. They not only exaggerate the essential feature points, but in turn make small grids around imperative features and make gradients on features which appear densely in the overlapping region between the source images. This method uses linear as well as nonlinear images (images that differ in angle and camera position).

In this work, we have used HoG with RANSAC, with geometric transformation, with Blob method of feature selection and detection. The RANSAC algorithm uses *fundamental matrix*, which estimates the fundamental matrix from corresponding points in linear (stereo) images. This function can be configured to use all corresponding points or to exclude outliers by using a robust estimation technique such as random sample consensus (RANSAC). Other method makes use of *geometric transformation*. Geometric transform returns a 2D geometric transform object, *tform*. The *tform* object maps the inliers in *matchedPoints1* to the inliers in *matchedPoints2*. The *matchedPoints1* and *matchedPoints2* inputs can be corner point objects, SURF Point objects, MSER objects, or M-by-2 matrices of  $[x, y]$  coordinates. The next method is *Blob* method. BLOB stands for Binary Large Object and refers to a group of connected pixels in a binary image. The term *large* indicates that only objects of a certain size are of interest as compared to those *small* binary objects which are usually considered as noise which in turn is significantly reduced by applying a binary mask that makes the final mosaic independent of ghosting effect and visible seams.

### **3.1 Main Steps of Image Stitching**

All stitching can be divided into several steps. First, registration for the image pair is done. Registration is done on the overlapping region to find the translation which aligns them. It is done to make the images photometrically and geometrically similar. After successful registration, some feature-based methods are applied to extract fine and strongest points from the input images in order to match images to their maximum. Next, these matched points are used either to create Homography using RANSAC or are used in gradient medium to reconstruct the final mosaic.

#### **3.1.1 Feature Matching**

Feature matching requires similar features from input images; therefore, it is the integration of the direct method and feature-based method. Direct method matches every pixel of one image with every pixel of other images; i.e., it incorporates every pixel value which makes it less popular technique of feature matching. It is generally used with images having large overlapping region and small translation and rotation because it cannot effectively extract overlapping window region from the referenced images, whereas feature-based method is used over large overlapping region as they only cull requisite and important features. Many feature extraction algorithms have been used such as FAST and Harris corner detection method. FAST features return corner point object points. The object contains information about the feature points detected in a 2D grayscale input image. The detection of FAST feature function uses the features from accelerated segment test (FAST) algorithm to find feature points.

#### **3.1.2 Image Matching**

In the following step, the SURFs extracted from the input images are matched against each other to find nearest-neighbor for each feature. Connections between features are denoted by green lines. When using RANSAC, unnecessary lines of matching are removed as outliers and necessary lines remained as inliers. But when geometric transformation is used with HoG to match the images, RANSAC estimation is eliminated because geotransform is a complete package of inlier and outlier detection and removal of unwanted noise from the input set of images.

#### **3.1.3 Image Calibration**

When HoG is used with RANSAC for matching images, calibration is done after feature extraction and matching because it returns projective transformations for

rectifying stereo images. Calibration rectification function does not require either intrinsic or extrinsic camera parameters. The input points can be M-by-2 matrices of M number of  $[x, y]$  coordinates, or SURF Point, MSERs, or corner point object.

## 4 Experimental Setup

### 4.1 Performance Matrices

The *Peak Signal-to-Noise Ratio (PSNR)* and *mean squared error (MSE)* are the two error metrics used to compare mosaic quality. The MSE represents the mean squared error between the final (de-noised image) and the original image, whereas PSNR represents a measure of the peak error.

#### 4.1.1 PSNR

As the name suggests, peak signal-to-noise ratio is the measure of peak error which is used to depict the ratio of maximum possible power of signal (image) to the power of the corrupting noise that affects the fidelity of its representation. It is represented in terms of mean square error (MSE) as follows:

$$\text{PSNR} = 10 * \log_{10}(256^2/\text{mse})$$

PSNR being popular and accurate in terms of prediction is commonly used in image processing. The higher the PSNR value, the better the quality of the reconstructed final mosaic.

#### 4.1.2 MSE

Mean square error is the term used to present the average between the estimator and what is estimated, i.e., between the final (de-noised image) and the original image before introduction of noise.

$$\text{MSE} = \text{sum1}/\text{numel}(\text{image})$$

This enables us to compare mathematically as to which method provides better results under same conditions like image size noise.

## 5 Experimental Results

In this section, the performance of the proposed method is illustrated with experiments. Given two partially overlapping, previously registered images with photometric inconsistencies (i.e., with differences in light intensity within the overlap region), the objective is to stitch them and produce a larger mosaic image which looks smooth and continuous without any noticeable artifacts such as seams or blurring.

The two images have an overlapping region indicated by a vertical black patch. The intensity levels in the two images have been modified and are clearly different. This has the effect in a “direct paste” mosaic image. The result of pasting the two images directly is shown in Fig. 1c. Later, when the HoG method (in Fig. 2a, b) is applied, the resultant seam is seen to have been absorbed and a final segmented panorama is produced in Fig. 3.

Later, the mosaic is improved upon by removing noise and ghosting artifacts from the generated panorama which can be visualized in Fig. 4.



Fig. 1 From left to right, a and b are original images, and c is the result of direct pasting

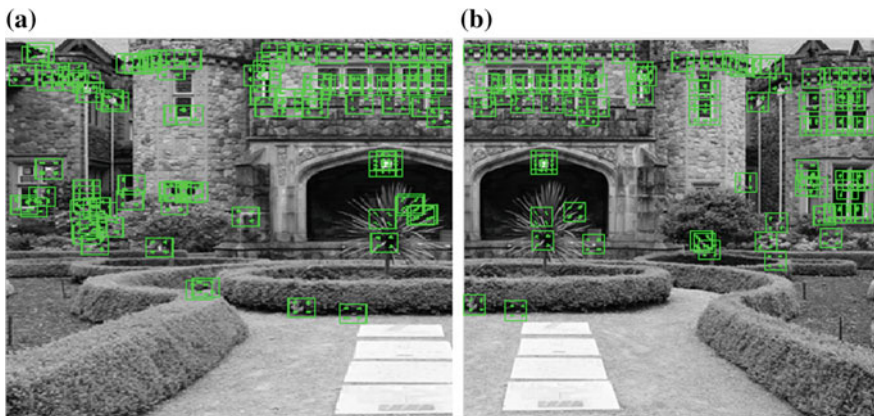


Fig. 2 From left to right, a and b shows the HoG descriptor in the input images



**Fig. 3** Final segmented panorama



**Fig. 4** Output mosaic without ghosting

The above experiments illustrate that the proposed method can lead to mosaic images without any visual artifacts despite differences in the intensities of the input images in the overlapping region. The algorithm was implemented using MATLAB and was running on a  $\times 86$  64 bit PC Architecture, Intel Core Duo T2080, 1.73 GHz (Figs. 5, 6 and 7; Tables 1 and 2).



Fig. 5 Temporal comparison

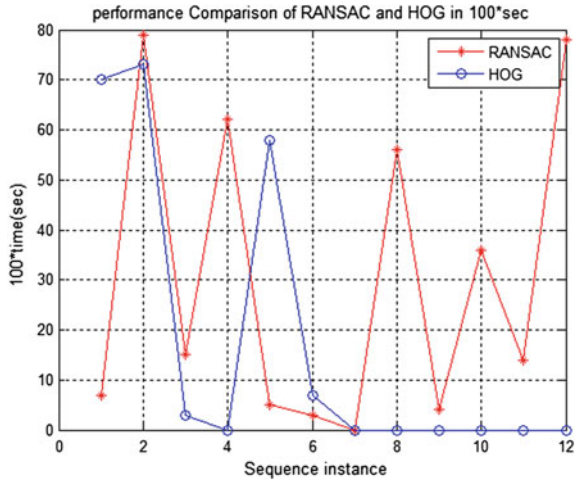
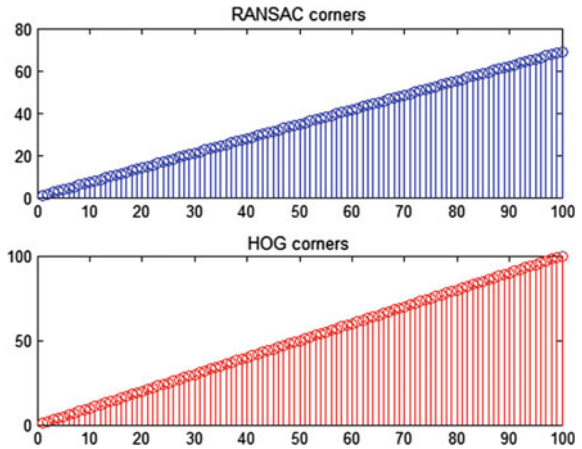


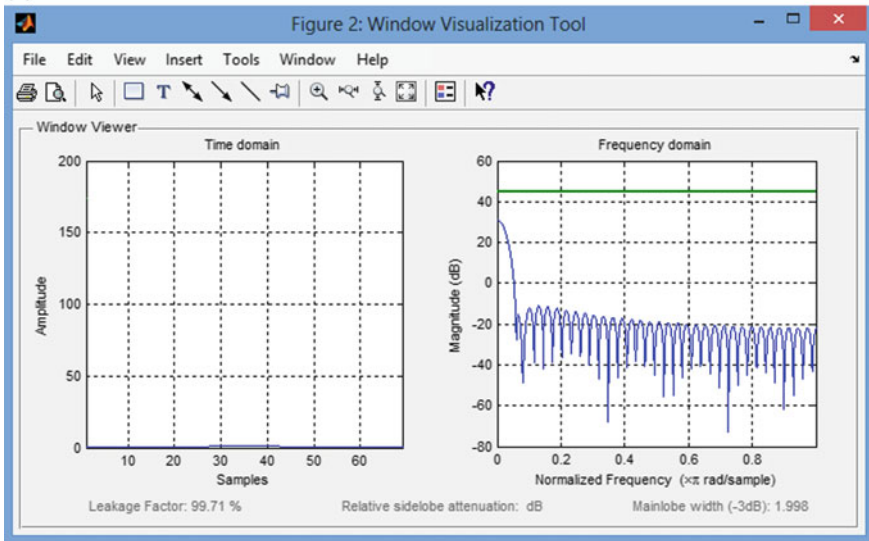
Fig. 6 In search for strongest points in both input images, the matched points are compared against each other



```

p1,p2 = image inputs; Im1 = image output;
X = imread ('cameraman.tif');
X1 = X;
X1(X <=100) = 1;
[psnr,mse,maxerr,L2rat] = measerr (X, X1)
To find energy distribution in image:
Entropy (p1)
Entropy (p2)
Entropy (Im)
PSNR—Peak (pixel) SNR (psnr = 10*log10 (256^2/mse));
MSE—Mean Square Error (mse = sum1/numel (image));
    
```

(a)



(b)

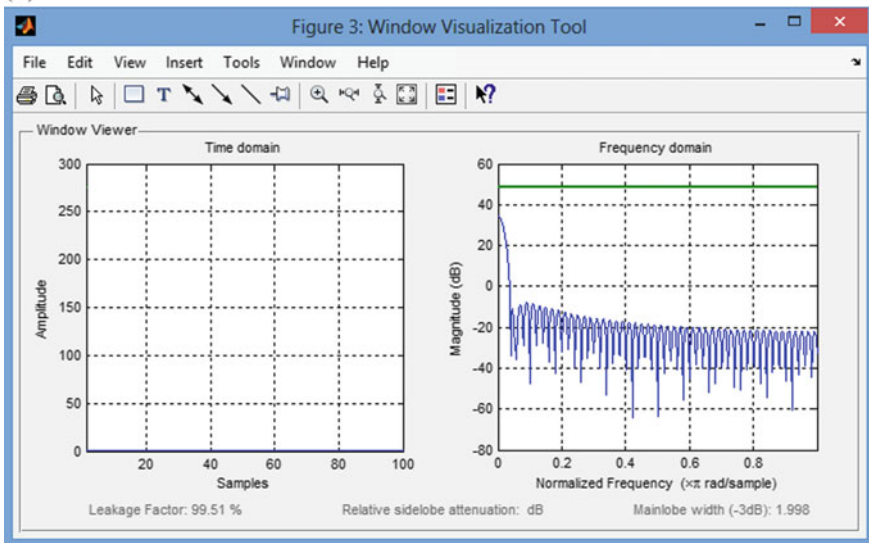


Fig. 7 a and b depicts the inliers comparison in time and frequency domain of HoG and RANSAC

**Ghosting:** If objects are not static in the overlap image, the image will appear blurry and ghosted. So far, scholars both domestic and foreign have done a lot of researches and made remarkable achievements in de-ghosting. Shum and Sezliski [26] also proposed a method for eliminating small ghosting based on computing optic flow and then doing a multiway morph.

**Table 1** Image properties comparison

	PSNR (p1)	PSNR (p2)	PSNR (Im1)	MSE (p1)	MSE (p2)	MSE absolute (Im1)	Maxerr (p1 ~ p2 ~ Im1)	L2RAT (p1)	L2RAT (p2)	L2RAT (Im1)
Blob_Ransac 200a500.m	18.6938	19.8816	20.3460	878.4119	668.2186	600.4480	99	0.9602	0.9724	0.9687
Blob_Ransac 1000a10000.m	18.6938	19.8816	20.3569	878.4119	668.2186	598.9536	99	0.9602	0.9724	0.9688
Blob_Ransac 2k10k.m	18.6938	19.8816	20.3246	878.4119	668.2186	603.4152	99	0.9602	0.9724	0.9689
Hog_geo 1000a10000.m	18.6938	19.8816	20.3102	878.4119	668.2186	605.4297	99	0.9602	0.9724	0.9686
Hog_geo 200a2k.m	18.6938	19.8816	20.2615	878.4119	668.2186	612.2477	99	0.9602	0.9724	0.9683
Hog_geo 100a1000.m	18.6938	19.8816	20.2649	878.4119	668.2186	611.7767	99	0.9602	0.9724	0.9681
Hog_Ransac 1000a20000.m	18.6938	19.8816	20.3772	878.4119	668.2186	596.1544	99	0.9602	0.9724	0.9688
Hog_Ransac 600a5000.m	18.6938	19.8816	20.3772	878.4119	668.2186	596.1544	99	0.9602	0.9724	0.9688
Hog_Ransac 300a10000.m	18.6938	19.8816	20.4004	878.4119	668.2186	592.4782	99	0.9602	0.9724	0.9688

Table 2 Performance comparison of methods used

Features	Unitary mode detection—Blob linear spatial RANSAC	Gradient HoG linear spatial RANSAC	Gradient HoG nonlinear spatial geometric transformation
Detection	Only one detection mode—SURF	Two detection methods are used: FAST and HoG	Two detection methods are used: FAST and HoG
Algorithms and no. of corner points	'Metric Threshold' = 2000 and compare strongest feature points	FAST detection for 300 level points, with strong feature point comparison	FAST detection for 300 level points, with strong feature point comparison
Feature detection iterations	First 2000 followed by 10,000 RANSAC trials	First 300 followed by 10,000 RANSAC trials	First 200 followed by 2000 trials
Spatial reference	Linear image spatial—no reference to light source/viewpoint/field-of-depth/background reference	With reference to light source/viewpoint/field-of-depth/background reference	With reference to light source/viewpoint/field-of-depth/background reference
Directional spatial reference	No directional reference	Gradient type detection	Gradient type detection
Blending	Via Alpha Blender	Via Alpha Blender	Via Alpha Blender

But here, an impression-based method HoG is used to get better PSNR and MSE values. The proposed algorithm shows better results than the standard median filter (MF) and decision-based algorithm (DBA). The method performs well in removing low-to-medium density impulse noise with detailed preservation up to a noise density of 70%, and it gives better peak signal-to-noise ratio (PSNR) and mean square error (MSE) values.

## 6 Conclusion

Image gradient blending techniques prove effective because the human visual system is known to be more sensitive to local contrast changes than to global intensity variations. Therefore, the gradient domain provides an excellent framework for image processing applications such as image editing, high dynamic range imaging, and compression and image mosaicing. In all these methods, the gradients of the source images are modified and the modified gradients are used to obtain the final image. In general, these gradient modifications lead to a gradient which typically is not a conservative vector field, and image reconstruction from this gradient no longer has an exact solution, but can be formulated as an optimization problem.

Therefore, a method of seamless stitching of images with photometric and geometric inconsistencies in the overlapping region has been discussed in this dissertation. This is done by generating a set of stitched gradients of the input images and then reconstructing the mosaic from the stitched gradients. This requires solving Poisson equation for the input images leading to mosaic formation without visual artifacts. Experimental results illustrate the method and show that it can lead to seamless mosaic images despite intensity differences in the overlap region of the input images.

## 7 Future Scope

The given work was based on the implementation and evaluation of the gradient-based HoG mosaicing technique. An approach and methodology have been proposed to enhance the performance producing the best quality panoramas by the fusion of the complimentary features specific to these algorithms. The test input images used for the present work were the planar images, but it can further be extended for MR images and cylindrical or spherical images.

Moreover, for computational enhancements, optimization of the algorithms used can be performed to dramatically increase timing results. Currently, the code developed for the proposed system runs mostly in MATLAB and is not designed especially for a high speed of computation. Multithreaded software or GPU-based

algorithm implementations are all ways to reduce the time of computation, making use of parallelism in different stages of the mosaicing process. These techniques can be implemented in addition to actual optimization of the mosaicing algorithms and procedures themselves.

## References

1. L.G. Brown, A survey of image registration techniques. *Comput. Surv.* **24**(4), 325–376 (1992)
2. B.Z.B. Zhang, Computer vision vs. human vision, *Cogn. Informatics (ICCI)*, in *2010 9th IEEE International Conference* (2004), p. 318108
3. Survey (ebsam, elmogy)
4. CV\_Overview
5. C. Guo, X. Li, L. Zhong, X. Luo, A fast and accurate corner detector based on Harris algorithm, in *3rd International Symposium Intelligence Information Technology Applied IITA 2009*, vol. 2, no. 4, pp. 49–52 (2009)
6. L. Zeng, S. Zhang, J. Zhang, Y. Zhang, Dynamic image mosaic via SIFT and dynamic programming. *Mach. Vis. Appl.* **25**(5), 1271–1282 (2014)
7. J. Melorose, R. Perroy, S. Careas, No Title No Title. *Statew. Agric. L. Use Baseline* **1**, 2015 (2015)
8. Z. Wang, F. Yan, Y. Zheng, An adaptive uniform distribution surf for image stitching, in *2013 6th International Congress Image Signal Processes*, vol. 2, (2013), no. Cisp, pp. 888–892
9. E. Rosten, T. Drummond, Machine learning for high speed corner detection, *Computer Vision—ECCV 2006*, (2004), pp. 430–443
10. N. Dalal, W. Triggs, Histograms of oriented gradients for human detection, in *2005 IEEE Computer. Social Conferences Computer Vision Pattern Recognition CVPR05*, vol. 1, (2004), no. 3, pp. 886–893
11. A. Kulkarni, J. Jagtap, V. Harpale, *Object recognition with ORB and its implementation on FPGA*, vol. 3 (The accents Organization, 2013)
12. R. Szeliski, Video mosaics for virtual environments. *IEEE Comput. Graph. Appl.* **16**(2), 22–30 (1996)
13. I.S. Sevcenco, P.J. Hampton, P. Agathoklis, Seamless stitching of images based on a HAAR wavelet 2D integration method. Department of Electrical and Computer Engineering University of Victoria (2011)
14. P. Pérez, M. Gangnet, A. Blake, Poisson image editing. *ACM Trans. Graph.* **22**(3), 313 (2003)
15. J. Jia, J. Sun, C.-K. Tang, H.-Y. Shum, Drag-and-drop pasting. *Siggraph* **25**(3), 631 (2006)
16. A. Levin, A. Zomet, S. Peleg, Y. Weiss, Seamless image stitching in the gradient domain. *Eccv* **3024**, 377–389 (2004)
17. M. Brown, D.G. Lowe, Recognising panoramas, in *Proceeding IEEE International Conference Computer Vision*, (2003) pp. 1218–1225
18. B. Triggs, P. Mclauchlan, R. Hartley, A. Fitzgibbon, Bundle adjustment—a modern synthesis 1 introduction. *Vis. Algorithms Theory Pract.* **34099**, 1–71 (2000)
19. M. Zhu, W. Wang, B. Liu, J. Huang, A fast image stitching algorithm via multiple-constraint corner matching. *Math. Prob. Eng.* **2013** (2013)
20. P. Singh, S. Batra, H.R. Sharma, Steganographic methods based on digital logic, in *Proceedings of the 6th WSEAS International Conference on Signal Processing, Dallas, Texas, USA* (2007), pp. 157–162

21. A.V. Krishna, Performance evaluation of new encryption algorithms with emphasis on probabilistic encryption and time stamp in network security, vol. 3, no. 5, pp. 39–46 (2009)
22. D. Hooda, P. Singh, A comprehensive survey of video encryption algorithms. *IJCA* **59**(1), 14–19 (2012)
23. R. Singh, P. Singh, M. Duhan, An effective implementation of security based algorithmic approach in mobile adhoc networks. *Hum—centric Comput. Inf. Sci.* **4**(1), 7:1–7:14 (2014)
24. D. Singh, D. Sethi, Reduction of noise from speech signal using HAAR and biorthogonal wavelet. **7109**, pp. 263–269 (2011)
25. G. Tsai, Histogram of oriented gradients. Lecture Series University of Michigan (2010), pp. 1–6
26. H.-Y. Shum, R. Szeliski, Construction of panoramic mosaics with global and local alignment. *Int. J. Comput. Vis.* **36**(2), 101–130 (2000)