International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Increasing Accuracy of Correspondences in Stereo Vision by Implementing Image Registering Technique

Indrit Enesi [1], Anduel Kuqi *[2]

Submitted: 14/11/2022 **Accepted:** 16/02/2023

Abstract*:* 3D reconstruction on stereo vision is widely used mainly for measuring distances of objects in a scene or sizes of the object. It is based on the triangulation of point correspondences in stereo images. The accuracy of point correspondences as well their number has great impact in the performance of 3D reconstruction of the object in a scene. In many cases stereo images are obtained by handheld cameras, with same parameters or not. The vertical camera axes are assumed to be parallel, but usually they are approximately parallel or not parallel, obtained stereo images are rotated related to each other. The paper considers the pixel correspondences in case where cameras have different parameters, and the vertical axes of stereo images create an angle between them. Images are first aligned according to each other by a transformation function defined by control points. Point correspondences between reference and registered image are defined, and their similarity is calculated based on respective normalized descriptive vector. Statistics about correspondences in registered images is compared with the ones in rotated images. Experimental results show that accuracy of correspondences is increased.

Keywords: *3D reconstruction, stereo vision, point correspondences, image registration, Euclidian distance, descriptive vector.*

1. Introduction

Stereo vision is an important technique in 3D reconstruction. It employees two cameras targeting the same object, they may have the same parameters or not. Mainly the triangulation approach of point correspondences is involved for the 3D reconstruction [1]. Finding the right correspondences is the main problem in stereo vision, false ones have a significant impact in the reconstruction process. Two approaches are used for finding correspondences, sparse stereo matching and dense stereo matching [2]. In sparse matching algorithms, distinctive points are extracted from the image pair and the corresponding points are established. Two sparse matching algorithms are developed, the first is based on KLT tracking algorithm and the second is based on correlations between regions of potential matches [2, 3]. The first works better for image pairs with narrow baselines, it is fast and accurate, the second is well used for image pairs with a larger baseline. Dense matching, in contrast to sparse on, finds matches for all points in the images. The matching search is constrained by the epipolar geometry. Two dense matching algorithms are developed. The first works better for narrow baseline, and assumes a known approximate solution for matching, the second is well suited for larger baselines and computes the matching score along the corresponding epipolar line [4]. The dense match algorithm returns a disparity map used for the 3D stereo reconstruction of the object. Finding exact correspondences is a challenge in 3D stereo reconstruction process.

Image registering process overlays two or more images of the same scene taken from different cameras and angles to spatially align the images. It calculates the spatial transformation to align an image according to another one. Usually, it is used for image analysis when images from different sources need to be combined. It is implemented in two ways, as image-to-image registration, where the images integrate matching pixels and as image-to-map registration where the image is distorted to match the map information of the other image keeping the original resolution [5,6,7,8]. The aim of the paper is to increase the matching correspondences found in the sparse stereo matching based on correlation of regions. The two cameras used in stereo vision usually have the same parameters, but in some cases they may be different. The images taken from them may be aligned, but in cases where handheld cameras are used, usually the left-hand and right-hand images are not aligned to each other. The paper analysis the match correspondences found from the image pair obtained where the vertical axes of the cameras are not parallel and the handheld cameras are not the same. The issue of the paper is to compared the number of correspondences points between left-hand and right-hand images with the one where the images are first

1 Department of Electronic and Telecommunication, Faculty of Information Technology, Polytechnic University of Tirana, Tirana, Albania
ORCID ID : 0000-0002-2695-2726
2 Department of Electronic and Telecommunication, Faculty of Information Technology, Polytechnic University of Tirana, Tirana, Albania
ORCID ID : 0000-0002-7141-0475
** Corresponding Author: Indrit Enesi*
Email: ienesi@fti.edu.al

registered regarding each other. Registering the images increases the similarity between them, which influences in the matching correspondences. The two stereo images are taken from different handheld cameras, with different resolutions and focus and their vertical axes are not parallel, the object in one image is inclined regarding the same object in the next image. The corresponding points are found, and the histogram of the number of corresponding points related to the distance is performed. The inclined image is registered regarding the reference one, corresponding points are found and the histogram is done. Experimental results show that the number of corresponding points of the original images and in the registered ones is increased.

The paper is organized in seven sections, related works are treated in section two, an analysis of stereo vision with passive and active categories, point and feature correspondences and their challenges are discussed in section three. Section four treats image registration process and different methods. The proposed algorithm is treated in section five, discussion of results is part of section VI, conclusions and further work close the paper.

2. Related Works

In [9] a high precision measurement schema based on stereo vision is proposed. Existing methods for the optimization of image registration procedure are covered. Based on the global polynomial optimization a hybrid method combines the SVD method, interleaving parameterization algorithm and the convex matrix inequalities relaxation achieving a global solution for the fundamental matrix.

In [10] the point cloud registration and generation from stereo images is discussed. Point clouds are extracted from stereo images of the same object taken from different angles are merged to create e 3D model of the object.

In [11] a highly accurate joint photometric and geometric image alignment is performed based on a new depth driven approach. The state of minimal energy is solved based on a real time capable alternating iterative optimization schema. Performance of disparity estimation is improved based on energy formulation on pixel wise color similarity.

In [12] a hybrid stereo vision system is proposed by a dynamic predictor. A low overhead accurate hardware predictor that uses the percentage of dark regions and the perceived brightness construct a real times system. The predictor provides necessary hints to activate a higher-level control system in case that disparity estimation is low.

In [13] a software tool is proposed for generation of stereo-vision disparity for Xilinx FPGAs.

In [14] the robust registration method for ultra-field infrared and visible images is proposed. The control points are extracted utilizing phase congruency and are optimized based on the guidance map. The ROI pair matching is performed based on epipolar curve. A multiple phase congruency directional pattern is established. A weighted histogram of phase congruency amplitudes encoded as binary patterns is presented to distribute information.

3. Stereo Vision

Stereo vision (SV) is the process of extracting 3D information from two 2D images taken from two cameras on the same scene by examining the relative positions of objects in these images. Depth information is obtained by the comparison of these images, which yields the disparity map. Usually some preprocessing steps are performed, barrel and tangential distortion are removed, image rectification and some measure information from the object. 3D reconstruction process concludes in a 3D point cloud, providing measurement at a known scale [15].

SV is used in obtaining 3D images. Two or more cameras, with known distance between them are employed in 3D reconstruction by capturing multiple images. Triangulation of the distance between cameras, the image matches and specific procedures are used to form the 3D image [15,16]. A number of methods are carried out for 3D reconstruction based on the collected information from images, but all of them are based on the disparity map in correlation of the distance of the object from cameras system with the baseline according to (1), where B represents the baseline distance and F the focal length of the camera lens.

$$d = (BF)/z \quad (1)$$

Stereo vision consists in passive, active and multiple-baseline stereo vision.

3.1. Passive Stereo Vision

In passive stereo vision (PSV), distance of the object from the camera system must be known.

The dynamic movement of the cameras prevents accurate 3D reconstruction since it is difficult to determine multiple baselines without assistance. In PSV, the inclusion of more than two cameras significantly increases the accuracy of 3D modeling because the number of matching points based on correspondences increases [17]. Two of the limitations of PSV are the existence of false boundaries, as illustrated in Fig. 1 and the problems that come from reflectance, illustrated in Fig. 2. These two problems prevent obtaining true boundaries and surface topography due to the diversity of light intensity from the surface of the object due to high levels of reflectivity.

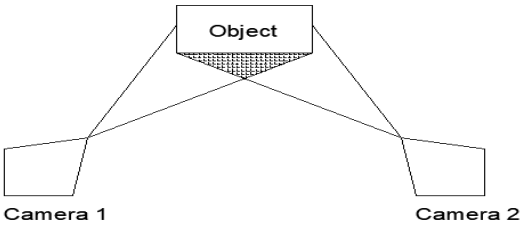

Fig. 1. Some parts of the object are not seen from other camera.

3.2. Active Stereo Vision

Unlike PSV, active stereo vision (ASV) allows dynamic movement of the cameras and the system. ASV processes the frames received from the cameras and merge them to form a more detailed 3D image. The light projected onto objects provides images captured by cameras to be matched together based on point correspondences [18].

Fig. 3 illustrates the acquisition of object images from different camera positions enabling automatic processing and 3D modeling through ASV.

ASV has its own limitations also. Simpler shapes of objects provide fewer matching points than more complex ones as pentagons or hexagons [16,18].

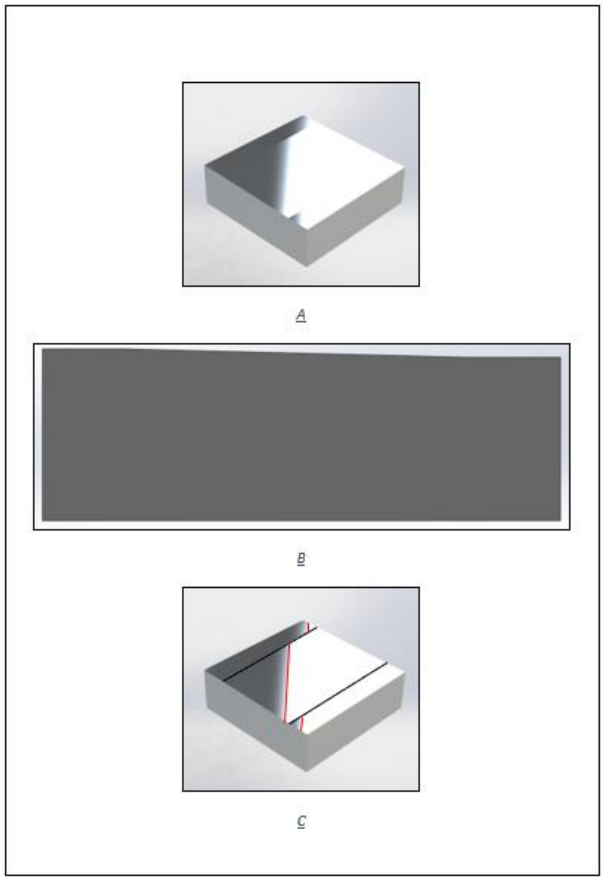

Fig. 2. Reflection of the object surface

.

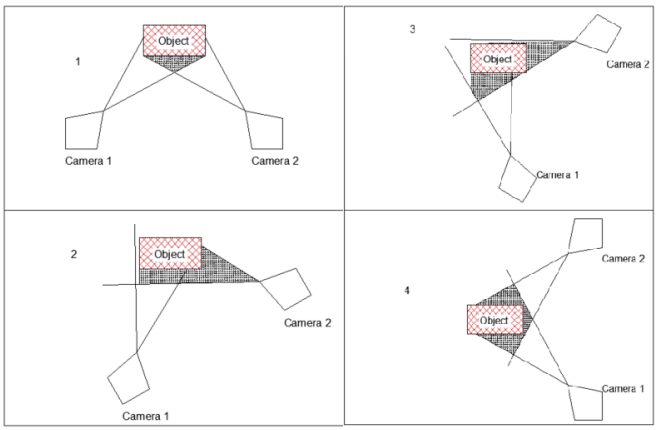

Fig. 3. Object seen from different camera position

3.3. Multiple-Baseline Stereo

In Multiple Baseline Stereo (MBS) utilizes many cameras for capturing images, processing and matching them to increase the accuracy of 3D reconstruction and at the same time allowing a wider view of the object. In MBS, a trade-off must be set between the accuracy of pixel correspondences and the accuracy of 3D reconstruction of the object [18,19]. From eq. 1 is observed that disparity and distance are proportional if the distance doesn't change. If the distance between cameras is extended, the disparity is increased yielding a more accurate distance from the object. But, increasing disparity increases the search region of pixel matches which increases the possibility for false matching.

The relation between disparity and baseline is illustrated in Fig. 4.

A proposed solution is to combine coarse and fine image captures. A short baseline is used for a coarse image, which results in an increased number of point correspondences. Repeating the process for a group of cameras make it possible for a high precision 3D reconstruction of the object.

3.4. Challenges Appeared in SV

Although the SV has advanced in the way to give solution to challenges, some new challenges has also come to place. Challenges consists in image matching, false boundary, reflection issues and identification to obstacles.

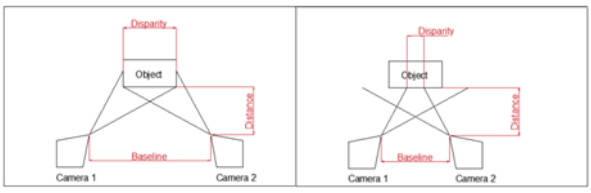

Fig. 1. Relation between disparity and baseline

3.4.1. Point Correspondences in Image Matching

Stereo Image matching is the basic procedure in 3D reconstruction in stereo vision. It consists in finding pixel correspondences in the left and right images based in the correlation of regions surrounding the corresponding pixels.

Some methods match all common features between images, some methods include line matching on boundaries or edges of the object, some methods find correlation between image patches using area based measurements, others analyze the correlation of windows template surrounding the current pixel in both stereo images, requiring a large amount of processing power. The last method enables the implementation of optimization techniques of the correlation of the surrounding area of the pixels.

3.4.2. False Boundaries

Based on different position of cameras, usually some edges are out of view from respective cameras, which may result in false matching. The solution can be moving the camera along the object, but the problem can be complicated when the background of the object is uniform.

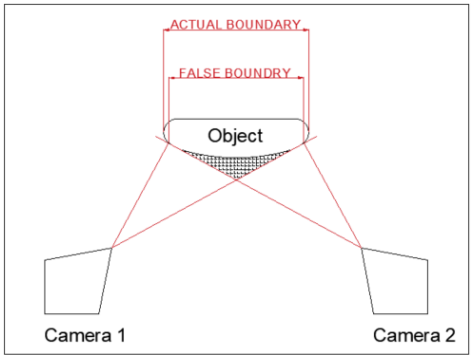

Fig. 2. Boundaries with no correspondence

3.4.3. Reflection of the Surface

Changing the light reflection from the object's surface has consequences in 3D reconstruction of the target object. Due to the uniformity of the surface, the structured light may lead to obtain a false surface which leads to false matching. Fig. 6 illustrates a false reflection of the structured light.

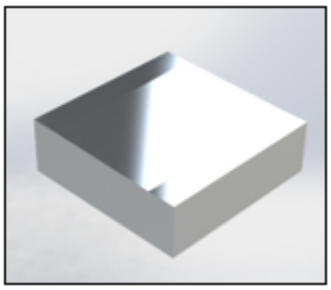

Fig. 3. Surface reflection

Usually this happens due to the textured geometry of the object as illustrated in Fig. 7.

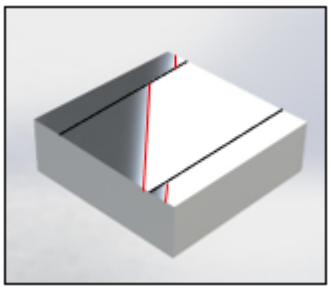

Fig. 4. False contour perceived from surface reflection

In Fig. 7 in black is shown the true geometry of the object, the red lines show the false geometry of it caused by the texture surface of the material.

4. Image Registration Technique

Image registration is the process of determining the correspondence between points in images of a scene, usually to fuse the complementary information. Relationship between two or more images usually is needed in the combination of the information content of them [7,20].

In general, the method involves a number of corresponding points to generate a transformation function of one image according to the other. Assuming that N corresponding points are known, as illustrated in equation (1), the function f( ) need to be determined.

$$\{(x_i, y_i) \leftarrow f(\,,) \rightarrow (X_i, Y_i) : i = 1, \dots, N\} \quad (1)$$

$$(X_i, Y_i) = f_{x,y}(x_i, y_i), i = 1, \dots, N \quad (2)$$

After the transformation function f(x,y) is determined, the for every pixel with coordinates (x,y) in the reference image, the respective transformed coordinates are calculated to obtain the registered image .

Transformation functions are dependent of geometric differences between images, the accuracy of the based corresponding points, the density and organization of them. These corresponding points are otherwise called control points, their accuracy, density and organization is very important, for more in the presence of noises. Due to the nonlinearity of image acquisition devices, usually the images to be registered have nonlinear geometric differences. In some cases it is possible to replace the nonlinear transformation functions with linear ones without significative difference. In case of noises, the approximated methods are preferable over interpolated ones to determine the transformation function. The mathematics of transformation functions is based on the similarity transformation which involves the rotation, translation and scaling between two images, as described in (3).

$$\begin{aligned} X &= x\cos\,\theta - y\sin\,\theta + h \\ Y &= x\sin\,\theta + y\cos\,\theta + k \end{aligned} \quad (3)$$

In (3) s means scaling, θ means rotational and (h,k) mean translational differences between two images. To solve (3) according to above parameters, two known point correspondences are needed. The scaling difference is determined from the ratio of distances of corresponding points, rotational difference is determined from the angle of lines connecting two corresponding points in each image. The two other parameters are determined using the coordinates of one corresponding point and solving (3).

In matrix language the (3) can be written as (4).

$$\boldsymbol{P} = s\boldsymbol{R}p + \boldsymbol{T} \quad (4)$$

In (4) **P** stands for coordinates of control points in the sensed image, **p** stands for the coordinates of the corresponding one in the reference image, **R** is the orthonormal matrix representing the rotation, s represents scaling and **T** represents translation. To determine the transformation function, the scaling s is calculated first, then rotational matrix **R** and translation vector **T**. Equation (5) illustrates the determination of matrix **R** and translation vector **T** based on coordinates of corresponding points

$$\mathbf{P}_i = \mathbf{R}\mathbf{p}_i + \mathbf{T}, i = 1, \dots, N \quad (5)$$

Considering the noise or other inaccuracy of the corresponding points, the matrix R and vector T are calculated by the minimization of results difference according to (6).

$$E^2 = \sum_{i=1}^{N} \|\mathbf{P}_i - (\mathbf{R}\mathbf{p}_i + \mathbf{T})\|^2 \quad (6)$$

According to […] the rotational matrix is determined using (7).

$$E_R^2 = \sum_{i=1}^{p} \|\mathbf{Q}_i - \mathbf{R}\mathbf{q}_i\|^2, \quad (7)$$

Where $\mathrm{Q}_i = \mathrm{P}_i - \mathrm{P}, \mathrm{q}_i = \mathrm{p}_i - \mathrm{p}$ and $\bar{P}$ and $\bar{p}$ are centers of gravity of the corresponding control points in sensed and reference images.

And the translation vector is calculated based on (8).

$$\mathrm{T} = \mathrm{P} - \mathrm{Rp} \quad (8)$$

Correspondence between images is mainly established through image registration process, which is basically classified as:

1. Image to image registration where the images are aligned according to a base image to integrate the pixels that represent the same object
2. Image to map registration, the input image is warped according to a base image to match the map information retaining its spatial resolution.

In image registration process, an image is spatially transformed according to another image referred as the reference one. Registration is a key step in the analysis of the image especially when data from different sources are combined. During the registration process two challenges appear, first, need to be determined the corresponding points between images, known as the matching process, and second, the information transformation process minimizing the distance between the aligned images [21].

The algorithms for image registration are classified in area and featured based methods. Area based methods are preferable when the prominent details of image are absent and the distinctive information is based colors or grey levels, feature based methods are used when there is information about local shapes and structures in the image. Features like points, edges, contours, corners, intersection points, surfaces are used for image matching. Four steps are followed to perform image registration: feature detection, feature matching, design of mapping function, image transformation and resampling [20, 21].

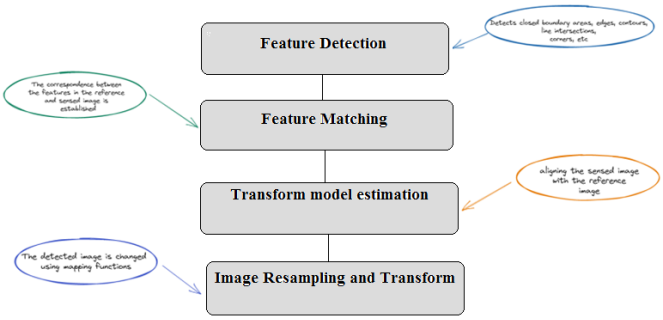

Fig. 5. Steps in image registration process

Image registration methods are classified in two general classes: area-based approaches and feature-based methods, but a more specific classification exists, as described following:

4.1. Pixel Based Method

The method is widely used for template matching or pattern recognition, it performs finding of location and orientation of a pattern or template in an image. As a matching metric it uses cross-correlation. The key disadvantages of the method are the high processing complexity and the maximum measured by flatness similarity.

4.2. Feature Based Methods

Feature based methods extract and abstracts information from images by reducing the redundant one. A sufficient number of features which are more tolerant to distortions are selected, then a matching process is performed. Usually, the feature based method is used when misalignment between images is not known. It consists in three stages, image features are extracted, correspondences are established between the features' image and the reference one, the spatial transformation is performed based on these correspondences.

A drawback is that features are preferred to be in distinctive parts of the image, but usually the selection is not based on their location.

4.3. Multimodal Image registration

In some cases, images are captured from different devices, yielding in images with different characteristics and general techniques based on spatial correlations may not be applied directly. In these cases, the registration is based on mutual information. Registration is performed using the entropy of the images, it doesn't change and in cases when the histogram changes. The method is accurate, but its performance lacks when images are of low resolution, contain little information or the overlapped region is smaller than mutual information.

4.4. Image Registration Using Frequency Information

The method is used when the images to be registered have only translation between them. The accuracy is higher than correlation method but less than other methods. Applying the Fourier transform only on the extracted features, results in an increased performance.

4.5. Image Registration based on Genetic Algorithms

Genetic algorithms use intrinsic parallelism and iterative procedures. The fitness function evaluates the quality of each candidate for each generation. The selected candidates then are combined and new candidates are produced. Evaluation and selection process terminates when a criteria is reached or a convergence is obtained. It is a reliable method usually in digital maps.

5. Algorithm

Two images of Tower Eiffel are obtained from two cameras with different parameters, at the same time. Images are shown in Fig. 9 and Fig. 10.

Fig. 6. Left image or the reference one

Fig. 7. Right image or the sensed one

Fig. 9 is captured from handheld camera 7M pixels with focal length f= 7.4mm, tower is 700 pixels tall. Fig. 10 is captured from handheld camera 10M pixels with focal length f= 5.2mm, tower is 600 pixels tall.

In each of the images distinctive points are defined and a window surrounding the distinctive point is defined, with size 11x11 pixels. It means that a descriptor of 121 elements corresponds to every distinctive point. The values of the pixels within window are offset by the mean value, rearranged into a vector and normalized to create a unit vector. ZNCC similarity measure is performed to compare the descriptor vectors as shown in (9).

$$s = \frac{\sum_{i=1}^{N}(I_1[i]-\bar{I}_1)(I_2[i]-\bar{I}_2)}{\sqrt{\sum_{i=1}^{N}(I_1[i]-\bar{I}_1)^2 \cdot \sum_{i=1}^{N}(I_2[i]-\bar{I}_2)^2}} = \underbrace{\frac{I_1[i]-\bar{I}_1}{\sqrt{\sum_{i=1}^{N}(I_1[i]-\bar{I}_1)^2}}}_{f_1} \cdot \underbrace{\frac{I_2[i]-\bar{I}_2}{\sqrt{\sum_{i=1}^{N}(I_2[i]-\bar{I}_2)^2}}}_{f_2} \quad (9)$$

The ZNCC of the two descriptor vectors is their dot product, resulting in the similarity measure between [-1, 1], where perfect match is 1, but values ≥ 0.8 are good matches [22]. These descriptors are distinctive and invariant to changes of intensity but not to scale or rotation. The aim of the paper is to increase the number of point correspondences based on ZNNC by registering images. Image register process improves the similarity between surrounding windows of the distinctive corners for rotated and scaled images. Similarity between normalized descriptive vectors is calculated using Euclidian distance. Harris operator is used to detect points in both images and HOG is the respective method to emerge the respective descriptor [23,24].

The matched corners serve as control points in the image registration process. Transformation function is defined based on the corresponding control points, the entire sensed image is transformed to be aligned according to the reference image. Pixel based method is performed for image registration.

The corresponding points are defined for the following three cases:

1. Case 1 – corresponding points are defined their similarity is measured for the two original images
2. Case 2 – corresponding points are defined their similarity is measured for the two images, where one is rotated
3. Case 3 – corresponding points are defined their similarity is measured for the two images, referenced and registered one

The results are compared to see the effect on the number of corresponding points when the image register process is implemented. The cumulative histogram of distribution of calculated distances between corresponding points and the histogram of distributed distances of corresponding points are created. Statistical results are analyzed and compared between three cases.

6. Implementation

Two images, as shown in Fig. 11, with different resolutions are taken, handheld cameras are considered not parallel to each other, and with not the same focal length. The two images have different sizes and orientations, determining the corresponding points is not easy.

Determine the number of corners found in each of the images. Both images are aligned, so their vertical axes are considered parallel to each other. Implementing the algorithm results that from the first image are detected 1288 corners and from the second one are detected 1426 corners. From the comparison of descriptors, 644 good correspondences are found, from a total of 1288 ones. 35% of all matches have descriptor less than 0.05.

The output of the implementation in Matlab is as following:

1288 corners found (0.1%), 1288 corner features saved

1426 corners found (0.1%), 1426 corner features saved

m = 644 corresponding points (listing suppressed)

m2 = 1288 corresponding points (listing suppressed)

Fig. 8. Left and right original images

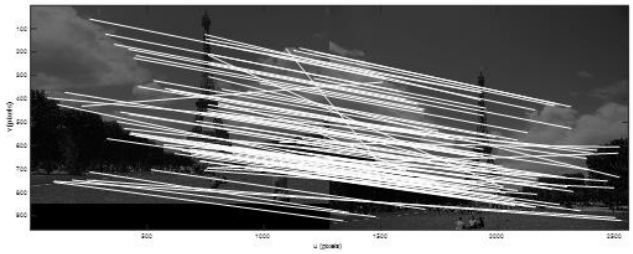

Fig. 9. Lines connecting corresponding points

The Fig. 12 shows lines connecting the corresponding points, most of the correspondences are correct, but there are also wrong correspondences. The lines connecting the incorrect correspondences do not have the same trend as the lines connecting the correct correspondences, whose trend is decreasing to the right.

The Euclidean distance is determined between the descriptors of the corresponding corners, the cumulative histogram of the distance distribution is done, without defining the threshold, as shown in the Fig. 13.

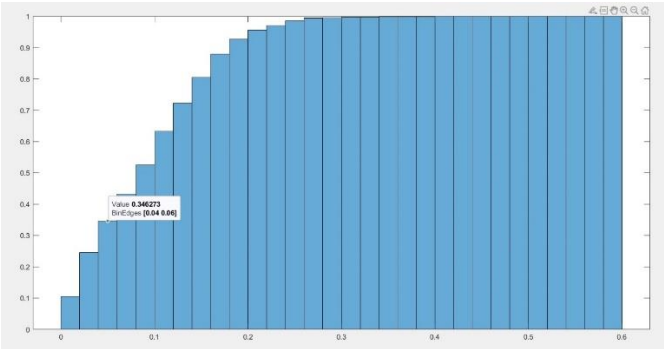

Fig. 10. Cumulative histogram of distribution of distances

From the cumulative histogram it is seen that 34.6% of all corresponding points have Euclidian distance less than 0.05. A large number of descriptors have Euclidian large distances, these correspondences usually are not valid. Analyzing fig. 13 yields that most of the correspondences are correct, but it obvious that some correspondences are false, the pattern of the line connecting false correspondences doesn't agree with the one of the true correspondences.

In the Fig. 14 the histogram of the distance distribution is presented, most of the correspondences have a Euclidean distance not greater than 0.15.

To analyze the effect of rotation in the number of point correspondences, one of the images is rotated, the correspondences are determined based on the Euclidean distance of the distinctive points in each image.

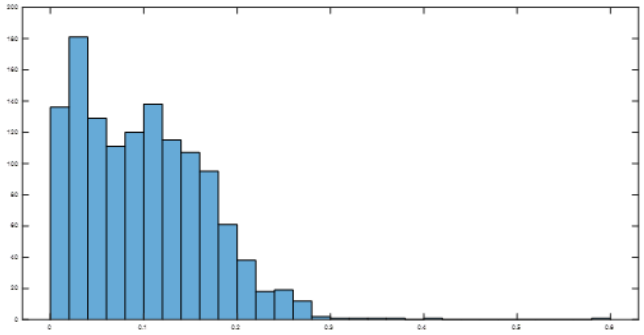

Fig 11 Histogram of distances

One of the images is rotated by 5 degrees clockwise, new correspondences are determined. Number of correspondences doesn't change, but the cumulative distribution of distances is reduced, 33% of pixels have descriptor distance less than 0.05.

Fig. 12. Right image rotated by 5 degrees clockwise

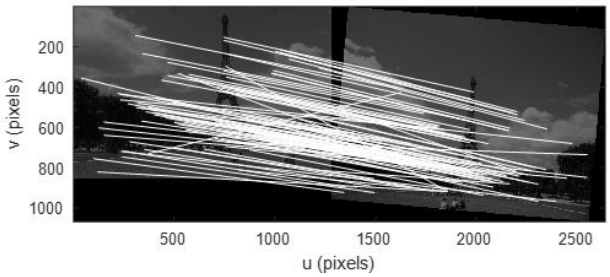

Fig. 13. Lines connecting corresponding points

The output of the algorithm is as following:

1288 corners found (0.1%), 1288 corner features saved

1877 corners found (0.1%), 1877 corner features saved

m = 644 corresponding points (listing suppressed)

m2 = 1288 corresponding points (listing suppressed)

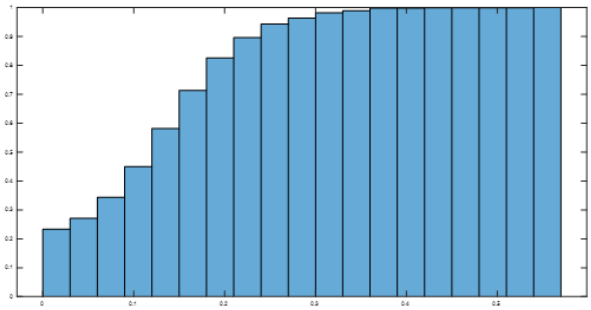

Fig. 14. Cumulative histogram for rotated image for rotated right image

From the cumulative histogram presented in the Fig. 17 it is noted that about 27% of correspondences have distances smaller than 0.05, the rest have distances about 10 times greater, increasing the probability of false correspondences.

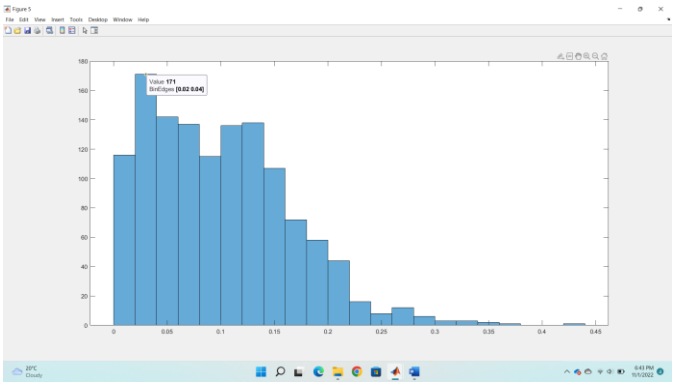

Fig. 15. Histogram of distances for rotated right image (labeled values)

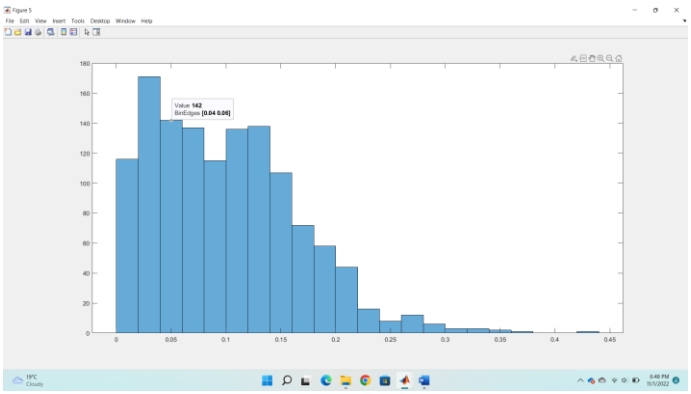

Fig. 16. Histogram of distances for rotated right image (labeled values)

From the distribution histogram of Euclidian distances between corresponding corners, it is noted that most of the correspondences are smaller than 0.15. The distribution of distances in the range 0-0.15 is almost uniform.

Implement image registration for these two images. The second image, rotated with 5 degrees clockwise, is oriented according to the first one. It is determined the number of correspondences between the two images, the first image is fixed (considered as reference image) while the second one (considered as sensed image) is registered regarding the first one.

The output of algorithm from Matlab is as following:

1288 corners found (0.1%), 1288 corner features saved

510 corners found (0.0%), 510 corner features saved

m = 644 corresponding points (listing suppressed)

m2 = 1288 corresponding points (listing suppressed)

Fig 17. Right image registered according to the left one

From the cumulative histogram presented in the Fig. 23 it is noted that the number of correspondences that have a distance of less than 0.05 is increased to about 33%.

From the distance distribution histogram, it is seen that the number of correspondences with distances very close to zero has increased significantly, 300 correspondences have distances in the interval [0 0.03].

Fig. 18. Rotated and registered images

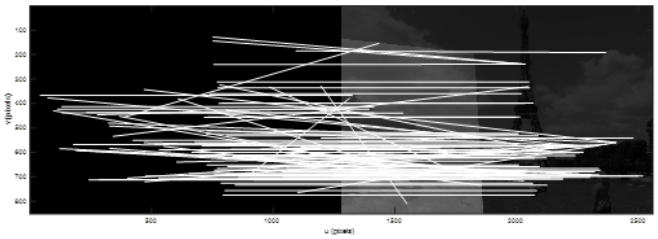

Fig. 19. Lines connecting corresponding points in registered images

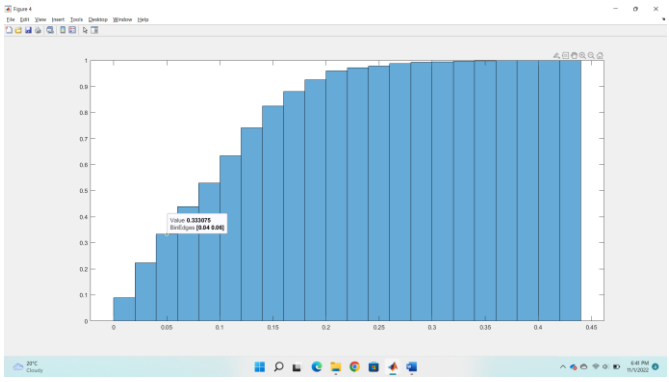

Fig. 20. Cumulative histogram for registered images

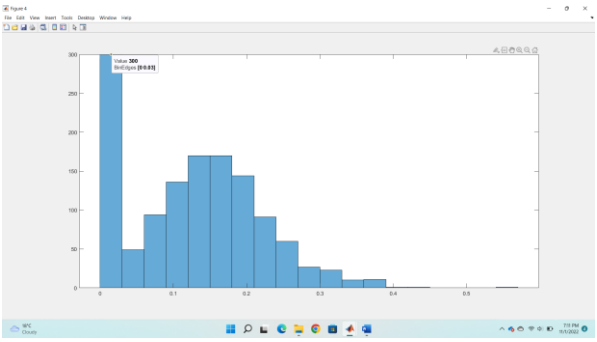

Fig. 21. Histogram of distances for registered images

7. Discussion

From experimental results it is concluded that the determination of true correspondences is dependent from the relative position between two cameras of stereo vision and cameras' parameters, such as focal length and resolution. A small number of true correspondences serve as control points which define the transformation function between two stereo images during the registration process. Based on the transformation function, the registration process is performed on the sensed image to be aligned as the reference one. The new correspondences are defined again on distinctive points detected from Harris operator on every image and their distances are calculated. The cumulative histogram and distribution histogram of distances are created as a statistic for measuring the performance of corresponding points. Analyzing histogram values results that the registration process increases number of correspondences, which are very important in 3D reconstruction process of the objects of the scene.

8. Conclusion

3D reconstruction process from stereo images is based on triangulation of corresponding points. The accuracy and number of point correspondences is very important in this process. Correspondences are determined by the correlation of regions surrounding distinctive points in stereo images. Not rarely the stereo images are obtained from cameras with different parameters and not aligned to each other. Statistical analysis of experimental results show that implementing the registration process to stereo images, increases the match of corresponding points by decreasing respective distance of description vectors comparing with the match correspondences of original stereo images.

Acknowledgements

This work was supported by the National Agency for Scientific Research and Innovation under the Contract no. 831.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Cheng, Yong-Qing & Wang, Xiaoguang & Collins, Robert & Riseman, Edward & Hanson, Allen. (2001). Three-Dimensional Reconstruction of Points and Lines with Unknown Correspondence across Images. International Journal of Computer Vision. 45. 129-156. 10.1023/A:1012424014764.

[2] Bhandari, Piyush & Wu, Meiqing & Aslam, Nazia & Lam, Siew Kei & Kolekar, Maheshkumar. (2020). Efficient Sparse to Dense Stereo Matching Technique. 10.1007/978-981-32-9088-4_14.

[3] Ahmed, Dr. Tasneem & Singh, Dharmendra & Raman, Balasubramanian. (2016). Potential application of Kanade-Lucas-Tomasi tracker on satellite images for automatic change detection. Journal of Applied Remote Sensing. 10. 026018. 10.1117/1.JRS.10.026018.

[4] Dall'Asta, E., & Roncella, R. (2014). A comparison of semiglobal and local dense matching algorithms for surface reconstruction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5*, 187-194. doi:10.5194/isprsarchives-xl-5-187-2014

[5] Pitiot, A. (2021). Multimodal Image registration. *Imaging Modalities for Biological and Preclinical Research: A Compendium, Volume 2*. doi:10.1088/978-0-7503-3747-2ch25

[6] Indrit Enesi, Miranda Harizaj, Betim Çiço, "Implementing Fusion Technique Using Biorthogonal Dwt to Increase the Number of Minutiae in Fingerprint Images", *Journal of Sensors*, vol. 2022, Article ID 3502463, 13 pages, 2022. https://doi.org/10.1155/2022/3502463

[7] Pataky TC, Yagi M, Ichihashi N, Cox PG. 2021. Landmark-free, parametric hypothesis tests regarding two-dimensional contour shapes using coherent point drift registration and statistical parametric mapping. *PeerJ Computer Science* 7:e542 https://doi.org/10.7717/peerj-cs.542

[8] Fathy GM, Hassan HA, Sheta W, Omara FA, Nabil E. 2021. A novel no-sensors 3D model reconstruction from monocular video frames for a dynamic environment. *PeerJ Computer Science* 7:e529 https://doi.org/10.7717/peerj-cs.529

[9] Ke, F., Liu, H., Zhao, D., Sun, G., Xu, W., & Feng, W. (2020). A high precision image registration method for measurement based on the Stereo Camera System. *Optik, 204*, 164186. doi:10.1016/j.ijleo.2020.164186

[10] Vladimir A. G. (2016). POINT CLOUDS REGISTRATION AND GENERATION FROM STEREO IMAGES. International Journal "Information Content and Processing", Volume 3, Number 2.

[11] Waizenegger, Wolfgang & Feldmann, Ingo & Eisert, Peter. (2011). Depth Driven Photometric and Geometric Image Registration for Real-Time Stereo Systems.. VMV 2011 - Vision, Modeling and Visualization. 25-32. 10.2312/PE/VMV/VMV11/025-032.

[12] D. Moolchandani, N. Shrivastava, A. Kumar and S. R. Sarangi, "PredStereo: An Accurate Real-time Stereo Vision System," 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022, pp. 4078-4087, doi: 10.1109/WACV51458.2022.00413.

[13] Jayasena, A. (2021). Register Transfer Level Disparity generator with Stereo Vision. *Journal of Open Research Software*, *9*(1), 18. DOI: http://doi.org/10.5334/jors.339

[14] S. Zhang, F. Huang, B. Liu, G. Li, Y. Chen, L. Sun, and Y. Zhang, "Robust registration for ultra-field infrared and visible binocular images," Opt. Express 28, 21766-21782 (2020).

[15] Achmad, M.S. & Bayuaji, Luhur & Daud, Mohd & Pebrianti, Dwi & Ann, Nurnajmin. (2019). Study on 3D Scene Reconstruction in Robot Navigation using Stereo Vision. 10.1109/I2CACIS.2016.7885292.

[16] Costineanu, D., Fosalau, C., Damian, C., & Plopa, O. (2008). Triangulation-based 3D image processing method and system with compensating shadowing errors. In *16th IMEKO TC4 Int. Symp.: Exploring New Frontiers of Instrum. and Methods for Electrical and Electronic Measurements; 13th TC21 Int. Workshop on ADC Modelling and Testing—Joint Session, Proc* (pp. 460-3).

[17] Gao, L. *et al.* (2022) "Research on multi-view 3D reconstruction technology based on SFM," *Sensors*, 22(12), p. 4366. Available at: https://doi.org/10.3390/s22124366.

[18] O' Riordan, Andrew & Newe, Thomas & Toal, Daniel & Dooly, Gerard. (2018). Stereo Vision Sensing: Review of existing systems. 10.1109/ICSensT.2018.8603605.

[19] Mingyou Chen, Yunchao Tang, Xiangjun Zou, Kuangyu Huang, Lijuan Li, Yuxin He. High-accuracy multi-camera reconstruction enhanced by adaptive point cloud correction algorithm. Optics and Lasers in Engineering. Volume 122, 2019, Pages 170-183, ISSN 0143-8166, https://doi.org/10.1016/j.optlaseng.2019.06.011.

[20] Ms. Ritu Singh Phogat, Mr. Hardik Dhamecha, Dr. Manoj Pandya, Mr. Bharat Chaudhary, Dr. Madhukar Potdar. Different Image Registration Methods – An Overview. International Journal of Scientific & Engineering Research, Volume 5, Issue 12, December-2014 44 ISSN 2229-5518.

[21] Balluff B, Heeren RMA, Race AM. An overview of image registration for aligning mass spectrometry imaging with clinically relevant imaging modalities. J Mass Spectrom Adv Clin Lab. 2021 Dec 18;23:26-38. doi: 10.1016/j.jmsacl.2021.12.006. PMID: 35156074;

PMCID: PMC8821033.

[22] Cournet, M. *et al.* (2016) "2D sub-pixel disparity measurement using QPEC / Medicis," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLI-B1, pp. 291–298. Available at: https://doi.org/10.5194/isprsarchives-xli-b1-291-2016.

[23] Gueguen, Lionel & Pesaresi, Martino. (2011). Multi scale Harris corner detector based on Differential Morphological Decomposition. Pattern Recognition Letters. 32. 1714-1719. 10.1016/j.patrec.2011.07.021.

[24] Danielsson, M. *et al.* (2016) "Feature detection and description using a harris-hessian/freak combination on an embedded GPU," *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods* [Preprint]. Available at: https://doi.org/10.5220/0005662005170525.