

PAPER • OPEN ACCESS

## Object occlusion recognition to increase an accuracy of ToF-tracking in augmented reality systems

To cite this article: M Klyueva *et al* 2018 *J. Phys.: Conf. Ser.* **1096** 012090

View the [article online](#) for updates and enhancements.

You may also like

- [Transmission characteristics of investor sentiment for energy stocks from the perspective of a complex network](#)  
Yajie Qi, Huajiao Li, Nairong Liu et al.
- [Ultra small angle x-ray scattering in complex mixtures of triacylglycerols](#)  
Fernanda Peyronel, Bonnie Quinn, Alejandro G Marangoni et al.
- [Nanomodification of Refractories with Finely-Dispersed Additives with the Use of a Vortex Electromagnetic Homogenizer](#)  
D S Suvorov, B B Khaidarov, D V Lysov et al.



**ECS**  
The  
Electrochemical  
Society  
Advancing solid state &  
electrochemical science & technology

**DISCOVER**  
how sustainability  
intersects with  
electrochemistry & solid  
state science research

# Object occlusion recognition to increase an accuracy of ToF-tracking in augmented reality systems

M Klyueva<sup>1</sup>, M Boyarkin<sup>2</sup>, K Bychenkov<sup>3</sup> and A Nikonorov<sup>1,4</sup>

<sup>1</sup>Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

<sup>2</sup>OOO Detector, Samara, Russia

<sup>3</sup>OOO Medx, Samara, Russia

<sup>4</sup>Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Molodogvardeyskaya str. 151, Samara, Russia, 443001

e-mail: golova.m@bk.ru

**Abstract.** Today, the accuracy of modern tracking systems should meet the highest expectations. Tracking process is essential for obtaining position and orientation in augmented reality systems. In this paper we overviewed the task of consistent tracking for a partially overlapped objects during the Time of Flight (ToF) camera operating time to obtain objects positions with extra accuracy. Also we considered basic concepts and existing solutions of the problem. We proposed new algorithm of overlapped objects tracking.

## 1. Introduction

Augmented reality (AR) alters one's ongoing perception of a real-world environment. The aim of the AR systems is overlaying of superimposing virtual information onto three-dimensional field of human perception [1]. Usually the AR systems operate in real time.

In AR systems, the real environment is annotated or enhanced with computer-generated graphics. These graphics must be exactly registered to real objects in the scene and this requires AR systems to track a user's viewpoint. The AR system uses the tracking for solving problems of combining and display of real and virtual objects. Tracking is the procedure of capturing the location of a moving object. The object's location estimation based on the position and orientation coordinates of the object [2]. Incorrect objects tracking entails errors in the operation of the augmented reality system. Therefore, this process must provide high accuracy.

An augmented reality system for surgical navigation imposes virtual models of internal organs and pathologies over the video stream based on their real three-dimensional position and lets a surgeon to observe the situation before surgical incisions. Later, a surgical staff may slightly change position of patient during the operation and a ToF (time of flight) camera is one of the robust solution to tracks these changes in real time [10].



Occlusion between the ToF-camera and the patient lead to the tracking precision decrease. In this paper we propose the solution based on timely occlusion detection and appropriate adjustments of the object registration algorithm. Here we consider the problem of object occlusion recognition using the ToF-camera to increase the tracking precision and stability.

## 2. Object tracking using ToF-cameras

Tracking systems use one or more cameras. More accurate results are obtained when using multiple cameras [3]. This makes it much easier to get a depth map of the scene. Depth map is a two-dimensional single-channel image that contains information about a distance from the sensor to the scene objects [4]. Recently, many novel approaches to scene sensing appeared like an RGB-D cameras [13] or more innovative diffractive optic [14], here we consider using of the ToF camera.

Depth maps can be also obtained using the ToF-camera. This camera measures the time of flight of the light signal emitted and reflected by each point on the surface of the image being captured.

There are three main methods of measuring the distance of the point position:

- pulsed modulation;
- continuous modulation;
- range gated imagers.

*Pulsed modulation.*

This method is easier than other. The camera emits a pulse, and at each point of the matrix the return time. To efficiently apply this method, we have to use equipment for response time detection with high accuracy.

*Continuous modulation.*

In this method, the emitter emits a modulated wave, and the receiver finds the maximum correlation from the received wave. According to the formula (1) is the distance to the object:

$$d = \frac{c}{4\pi w} \phi, \quad (1)$$

where  $c$  - speed of light;  $\phi$  - phase angle between radiation and reflection;  $w$  - modulating frequency.

*Range gated imagers.*

In this method, the gate faces the matrix. At a time 0 the shutter opens, and the scene lights up, at a time  $t$  the shutter closes. Objects that are farther away  $d$  will not be visible:

$$d = \frac{ct}{2},$$

where  $c$  - speed of light;  $t$  - scene lighting time.

There are many ToF-cameras on the market, for example: Panasonic D-imager, PrimeSense, Asus Xtion, Kinect.

Kinect captures depth and color scenes with a frame rate of 30 frames per second. The sensor consists of a depth sensor, a color camera and a microphone array. The depth sensor is an infrared projector and a monochrome CMOS matrix. Kinect for Windows v2 was chosen to work in the developed system.

## 3. Occlusion problem in tracking systems

The tracked object could be overlapped by another object during the tracking process. This object overlapping is called occlusion. Overlaps of the objects has a great impact on the tracking accuracy. To persistently control patient's position, tracking system should recognize and prevent the occlusion. The analysis of existing algorithms for recognition the object occlusion was done. We consider the most typical and known solutions of this problem in the next section.

### 3.1. Adaptation to the changes of object

This algorithm adapts the model to slow changes of the object during operation [5]. It uses three component-based distributions:

- S - stable;
- W - transitional;
- L - emissions;

The stable component is trained during the whole period of work. The transition is trained over the two frames. The emission component has a uniform distribution and characterizes overlaps. The stable and transient components have a Gauss distribution. The model is trained using the EM-algorithm [5].

### 3.2. Block contour tracking

The occlusion is analyzed by comparing the motion parameters between the object and the image blocks [6]. The object is represented by a mask, which is obtained in several steps: the implementation of multi-valued segmentation in 4 directions and search for the area of equally moving segments.

The algorithm determines the blocks of the frame that belong to the object, its border and background. And work is being done with these blocks. By applying affine transformations to the initial blocks, you can determine the movement between frames.

Occlusion and disocclusion are considered as dual events and this relationship is explained in detail. The algorithm for detecting occlusion is developed by modifying the disocclusion detection algorithm in accordance with the duality principle.

### 3.3. Tracking a multivalued template

This algorithm uses templates to represent an object [7]. The template adjusts to the change of the object by smoothing the selected features over time, using a Kalman filter for each pixel in the frame. The algorithm tracks objects in real time with partial overlaps and changes in illumination.

The pattern is a feature vector for each point in the image. All vectors are processed independently, and their probability is calculated using the Bayes formula.

The distribution function contains the Huber function to handle the measurement error. If the error exceeds the specified threshold, the current state of the object is not subject to the Gauss distribution, and in the distribution formula the quadratic norm is replaced by a linear one.

The number of the overlapped pixels depends on the mode overlap and stopping adaptation of the template. A long overlap can change the pattern to such an extent that it cannot be mapped to an object that is no longer overlapped.

## 4. Proposed algorithm of occlusion detection

Now, there is no generic algorithm for processing the object occlusion. All solutions are based on the tracking algorithm that is used in the system. We propose a solution to the occlusion recognition based on the A-ICPtracking algorithm.

### 4.1. Anisotropic Iterative Closest Point

Iterative closest point (ICP) is the classical solution to the task minimize the difference between two clouds of points [8]. The algorithm works with point clouds that have zero expectation and isotropic Gaussian noise. The point cloud obtained by the ToF camera has an anisotropic noise of point localization. Therefore, we use the algorithm A-ICP (anisotropic iterative closest point).

A-ICP is a modification of the ICP algorithm for the possibility of taking into account anisotropic noise of point localization [9]. The error localization at each point  $\vec{p} \in X \cup Y$  is normally distributed with zero mean and covariance matrix  $K_{\vec{p}} = V_{\vec{p}} S_{\vec{p}}^2 V_{\vec{p}}'$ . The algorithm searches for the rotation matrix

$R$  and vectors of movement  $\vec{t}$  models so that the weighted error metric  $FRE_{weighted}^2(R, \vec{t})$  minimized:

$$FRE_{weighted}^2(R, \vec{t}) = \sum_{i=1}^{N_x} \|W_i(R\vec{x}_i + \vec{t} - \vec{y}_{idx(i)})\|_2^2,$$

where  $R$  - rotation matrix;  $\vec{t}$  - movement vector;  $W_i$  - rotation matrix function  $R$ ;

$idx(i)$  - the index of the point  $\bar{y}_{idx(i)} = C_{new}(R\bar{x}_i + \bar{t}, Y)$ ;

$$W_i = w(RK_{\bar{x}_i}R' + K_{\bar{z}_i})^{-\frac{1}{2}},$$

where  $w$  - variable to normalize;  $(\bar{x}_i, \bar{z}_i)$  - the point of compliance in the combine models.

The algorithm consists of the following steps:

1. Variable initialization:

- a.  $k = 1$  - the number of the iteration;
- b.  $FRE_{weighted}^0 = \infty$  - zero error metric;
- c.  $R^0 = I$  - zero rotation matrix;
- d.  $\bar{t}^0 = \bar{0}$  - zero movement matrix;
- a.  $K_{\bar{x}_i}^0 = K_{\bar{x}_i}$  - zero covariance matrix;

2. The calculation of the current corresponding points  $Z^k = \{\bar{z}_i^k\}$ ,  $i = 1, \dots, N_x$ .  $\bar{z}_i^k \in Y$ , as the nearest point  $\bar{x}_i \in X^k$  to according to the measure of distance:

$$idx(i)^k = \arg \min_j d_{new}(\bar{x}_i^k, \bar{y}_j); \quad (5)$$

$$\bar{z}_i^k = \bar{y}_{idx(i)^k}. \quad (6)$$

Calculate the current covariance matrix  $K_{\bar{x}_i}^k$  by formula, the current cross covariance matrix  $K_{ij}^k$  by formula, the weight matrix  $W_{ij}^{k-1}$  by formula, the current distance  $d_{new}(\bar{x}_i^k, \bar{y}_j)$  for each point transformation  $\bar{x}_i^k \in X^k$  and  $\bar{y}_j \in Y$  by formula:

$$K_{\bar{x}_i}^k = R^{k-1}K_{\bar{x}_i}^{k-1}(R^{k-1})'; \quad (4)$$

$$K_{ij}^k = K_{\bar{x}_i}^k + K_{\bar{y}_j}; \quad (5)$$

$$W_{ij}^{k-1} = w(K_{ij}^k)^{-\frac{1}{2}}; \quad (6)$$

$$d_{new}(\bar{x}_i^k, \bar{y}_j) = W_{ij}^{k-1}(\bar{x}_i^k - \bar{y}_j). \quad (7)$$

3. Calculate the rotation matrix  $R^k$ , the movement vector  $\bar{t}^k$  by the formula (11). And weighted error metric  $FRE_{weighted}^k$  by the formula (12):

$$(R^k, \bar{t}^k) = \arg \min_{R, \bar{t}} \sum_{i=1}^{N_x} \|W_i^k(R\bar{x}_i^k + \bar{t} - \bar{z}_i)\|_2^2; \quad (8)$$

$$FRE_{weighted}^k = \sqrt{\sum_{i=1}^{N_x} \|W_i^k(R\bar{x}_i^k + \bar{t} - \bar{z}_i)\|_2^2}. \quad (9)$$

4. Apply the transformations calculated in the previous step to  $X^k$  and obtain the transformed set of points  $X^{k+1}$  for the next iteration  $k+1$ :  $x_i^{k+1} = R^k x_i^k \bar{t}^k$ .

5. If the formula (13) is correct or the specified maximum number of iterations is reached, the algorithm stops working. Otherwise, it  $k$  is added the unit and move to step 2;

$$|FRE_{weighted,k} - FRE_{weighted,k-1}| < \varepsilon.$$

#### 4.2. Solution of the occlusion detection problem

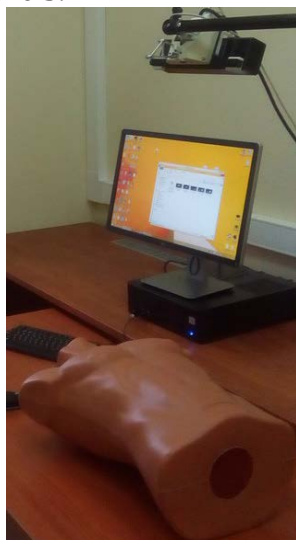
Finally, we solve the occlusion problem on the base of the proposed algorithm. Fiducial Registration Error (FRE) is the output value of the algorithm. We analyze this value to determine when the object overlaps. This value will be high if the object is moving quickly or something blocked it.

### 5. Experimental results

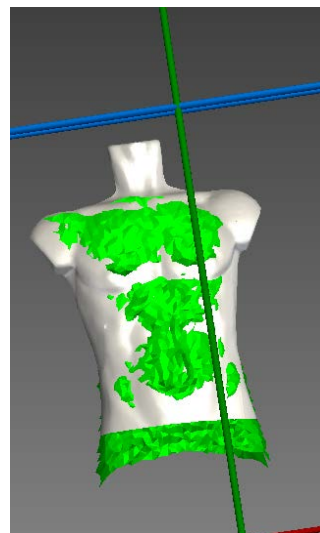
To evaluate the proposed algorithm, we analyzed the FRE values for three cases:

- 1) the object does not move and does not overlap;
- 2) the object moves but does not overlap;
- 3) the object does not move, but is overlapped.

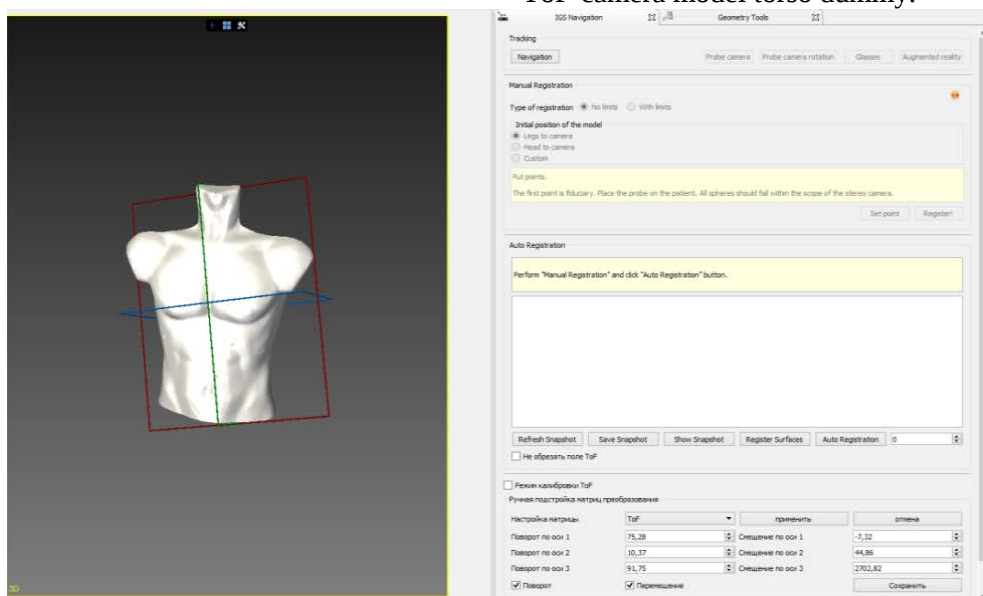
To obtain data we used a dummy torso, which is the patient's prototype. We used Autoplan surgery assistant and navigational system [11], [12] as a host platform to deploy our algorithm. The experimental setup is following. The dummy was located on the table and Kinect sensor on a tripod is placed directly above it. These conditions are shown in figure 1. The operation of this system is shown in figures 2 and 3.



**Figure 1.** Experimental setup.

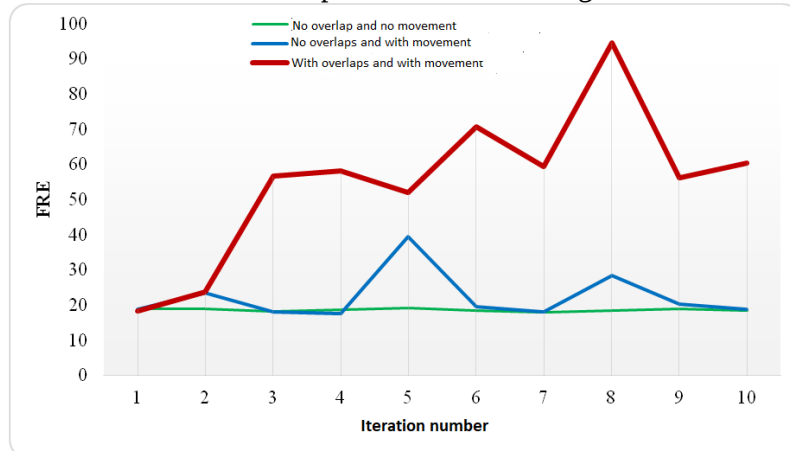


**Figure 2.** Alignment of the point cloud to the ToF-camera model torso dummy.



**Figure 3.** Interface of the developed system.

The average values of FRE before the implementation of the algorithm are shown in figure 4.



**Figure 4.** The average values of the FRE prior to the implementation of the recognition algorithm overlap.

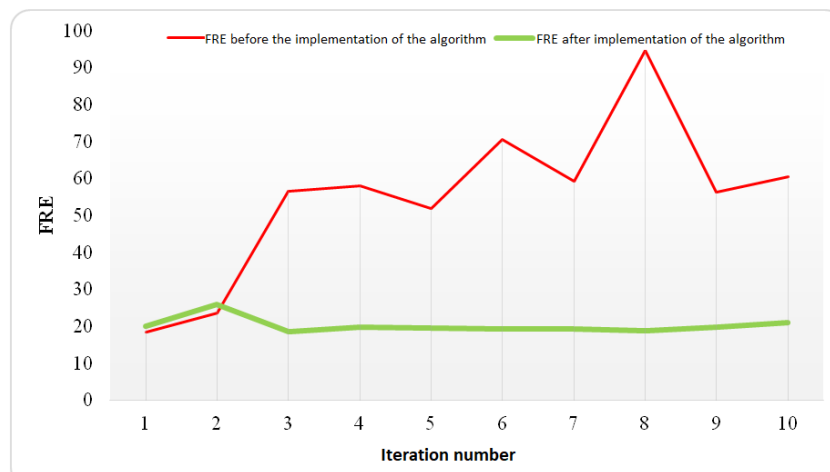
On the basis of the graph draw conclusions:

1. If the object is stationary and it does not overlap anything, the FRE value is low.
2. If the object is moved, the FRE increases and you can observe small jumps during the operation of the algorithm.
3. If the object is blocked, the FRE is much higher than the case of moving the object. If after a while to remove the obstacle, the AICP also gives the wrong result.

Designed algorithm:

1. Filling the array  $M$  dimension of 20 values  $FRE$ , which are obtained from the fixed and not overlapped model.
2. Calculate the average value  $s$  of the array  $M$ .
3. Compare current  $FRE$   $c$  with  $s$ :
  - a) if  $|FRE - s| > t$ , where  $t$  is a threshold, so there was an overlap and you do not need to track the object;
  - b) if  $|FRE - s| < t$ , so track the object and update the array  $M$  with the current  $FRE$ ;
4. Repeat the steps for each  $FRE$ .

The value of the threshold  $t$  was chosen by experimental studies, which found the optimal value of 7.



**Figure 5.** The decrease in FRE upon detection of the beams by the proposed algorithm.

We implement this algorithm within our system. Figure 5 demonstrates a graph when the recognition algorithm overlaps.

Figure 5 shows that the developed and implemented algorithm works well. Detection of overlapping object is carried out in real-time. The first 20 frames of the algorithm are accumulated and then works in real time, which depends on the operation of the entire system in real time. We have reduced the average TRE value from 55,017 to 20,186.

## 6. Conclusion

In this article we considered ToF tracking in the AR-based surgery assistant system. The problem of the object occlusion during tracking was addressed. Proposed solution may dramatically decrease tracking precession. We have analyzed the existing solutions of this problem and concluded that the existing solutions to handling occlusion are not suitable in the case of the surgery assistant system. We proposed the algorithm of occlusion detection and prevention. The developed algorithm was tested on the test object (dummy torso). The experimental results showed that the proposed algorithm recognizes the occlusions of the object in real time and reduces the average TRE (target registration error) value from 55,017 to 20,186. This allows to increase the accuracy and stability of ToF-tracking in the real-time AR-based surgery assistant system. The proposed algorithm was applied in the “Autoplan” surgery-assistant system [11], [12].

## 7. References

- [1] Azuma R T 1997 A Survey of Augmented Reality *Teleoperators and Virtual Environments* **6** 355-385
- [2] Yilmaz A, Javed O and Shah M 2006 Object tracking: A survey *ACM Journal of Computing Surveys* **38** 45
- [3] Suau X, Ruiz-Hidalgo J and Casas J R 2012 Real-time head and hand tracking based on 2.5D data *IEEE Transactions on Multimedia* **14**(3) 575-585
- [4] Ershov E, Karnaukhov V and Mozerov M 2016 Probabilistic choice between symmetric disparities in motion stereo for lateral navigation system *Optical Engineering* **55**(2) 1-20
- [5] Jepson D, Fleet D J and El-Maraghi T F 2003 Robust Online Appearance Models for Visual Tracking *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1296-1311
- [6] Hariharakrishnan K and Schonfeld D 2005 Fast object tracking using adaptive block matching *IEEE Trans. on Multimedia* 853-859
- [7] Nguyen H T and Smeulders A W M 2004 Fast occluded object tracking by a robust appearance filter *IEEE Trans. on Pattern Analysis and Machine Intelligence* 1099-1104
- [8] Besl P J and McKay Neil D 1992 A Method for Registration of 3-D Shapes *IEEE Transactions on pattern analysis and machine intelligence* **14**(2) 239-256
- [9] Maier-Hein L, Franz A M, dos Santos T R, Schmidt M, Fangerau M, Meinzer H P and Fitzpatrick J M 2012 Convergent Iterative Closest-Point Algorithm to Accomodate Anisotropic and Inhomogenous Localization Error *IEEE Trans. on Pattern Analysis and Machine Intelligence* p 34
- [10] May S, Droschel D, Holz D, Fuchs S, Malis E, Nchter A and Hertzberg J 2009 Three-dimensional mapping with time-of-flight cameras *Journal of Field Robotic* **26**(11-12) 934-965
- [11] Smelkina N, Kosarev R, Nikonorov A, Bairikov I, Ryabov K, Avdeev A and Kazanskiy N 2017 Reconstruction of anatomical structures using statistical shape modeling *Computer Optics* **41**(6) 897-904 DOI: 10.18287/2412-6179-2017-41-6-897-904
- [12] Nikonorov A 2016 Vessel Segmentation for Noisy CT Data with Quality Measure Based on Single-Point Contrast-to-Noise Ratio *Communications in Computer and Information Science* **585** 490-507
- [13] Sidorchuk D, Gusamutdinova N, Konovalenko I and Ershov E 2015 Problem-oriented stereo vision quality evaluation complex *Proceedings SPIE. Eighth International Conference on Machine Vision (ICMV)* **9875** 1-5

- [14] Nikonorov A, Petrov M, Bibikov S, Yuzifovich Y, Yakimov P, Kazanskiy N, Skidanov R and Fursov V 2016 Comparative Evaluation of Deblurring Techniques for Fresnel Lens Computational Imaging *Proc. Int. Conf. Pattern Recognition* 775-780

### **Acknowledgement**

This research was supported by the Federal Agency of Scientific Organizations (Agreement No 007-GZ/CH3363/26), a Presidential Grant of Russian Federation MD-2531.2017.9, and grants of RFBR 16-47-630721 r\_a, 18-07-01390, 16-29-09528 ofi\_m, 17-29-03112 ofi\_m.