Underwater Object Recognition in Photo Images

*Alexander Pavin
Laboratory of Underwater Vehicles
Far Eastern Federal University
Vladivostok, Russia
pavin@bk.ru

***Abstract*—The identification problem of artificial or natural objects in underwater images is considered in this paper. Any object (including a previously received image) with unequivocally determined contours can be used as a pattern. Recognition of the required object is made by an exhaustive search of key points or small pieces of borders with a subsequent increase of accumulator cells. Processing of real sea-bottom images has confirmed the algorithm's efficiency (including dead reckoning, marker recognition and line recognition tasks).**

Keywords—pattern recognition; identification; image processing; detection; underwater vehicle; photo system; photo reckoning; AUV; ROV; UUV.

I. Introduction

Seabed image processing used in many applications related to the control and navigation of remotely operated vehicles (ROV) or autonomous underwater vehicles (AUV). These tasks may include: maneuvering near the underwater structures, inspection of underwater communications [1-3], task of underwater docking using visual markers, dead reckoning of the trajectory using photos [4], positioning over a sea bottom fragment [5], image stitching, marine organisms counting and others. For all of these tasks the desired object can be of any geometric shape, e.g. different markers for docking. In many cases the object remains anonymous up to the mission start, as in the case of dead reckoning relative to the ground (where previously captured picture is the desired object). This work is part of integrated project "Modern technologies and technical means of monitoring the state of marine ecosystems and marine biological resources" [6-8].

In many cases, the desired object is uniquely determined by the shape of its boundaries, but the contours of objects in a picture can be noisy, intermittent or blurred. This appears when AUV shooting in muddy water, siltation, with the presence of foreign objects, algae, marine organisms, etc. The basic requirement to the recognition algorithm is the ability to identify the objects in a real time mode with frequency 1-2 Hz (depend on flash of underwater robot).

II. Preliminary Image Processing

Image preprocessing is necessary to reduce the algorithm operations, smooth the noise in the original image and obtain baseline data for the identification procedures. Thus, the preprocessing consists of three steps (fig. 1):

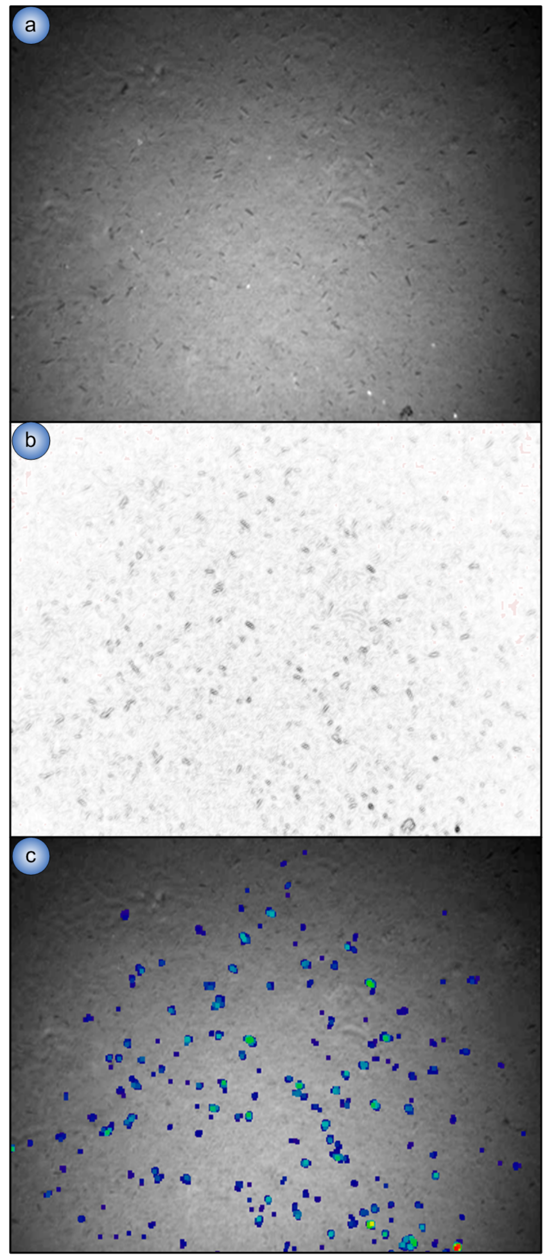

Fig. 1. Underwater photographs obtained by the underwater vehicle: a) scaled seabed image; b) gradient maps; c) mos contrasted key points on the photo.

*This work was supported by Russian Science Foundation, (project No. 14-50-00034)

- Scaling the original frame of the image (many image noises smoothed out in this case).
- Emphasize object boundaries using gradient maps.
- Search on the map gradients most contrasting points (points of interest, key points) or segments (segments of interest) belonging to the boundaries of objects.

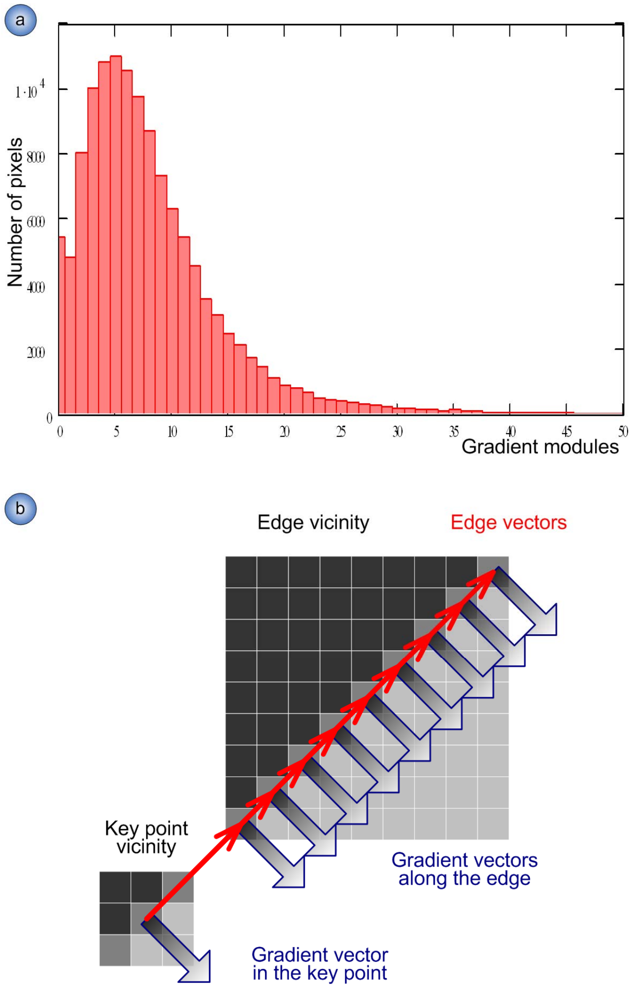

Fig. 2. Key points and segments detection in the image: a) histogram of the gradient modules; b) the detection of the boundaries of objects.

Object recognition in a photo image is made on the basis of borders, which are in turn defined by the size and direction of pixel gradients. To calculate a gradient, it is possible to use approximations in a sliding window of a size of 2x2, 3x3 pixels and more. For the estimation of a gradient in the central point of a sliding window of 3x3 pixels, it is possible to use Prewitt, Sobel or Scharr operators [9]. Because the object recognition method used a gradient direction, the Scharr operator (as possessing the best circular symmetry) was used. Each of these operators determines the brightness gradient at each point of the image and allows to roughly estimating the direction of the border (perpendicular to the gradient). Among the set of points belonging to the contours of the objects about 1000-2000 most contrasted points is enough to use in identification algorithm.

To define a fixed number of the most contrasting points we need to choose the minimum threshold contrast for each frame. Determination of the contrast threshold is produced by plotting the histogram of gradients. This histogram represents the number of pixels having respective contrast (fig. 2-a). Thus, to detect the threshold it is enough to summarize the points with the greatest contrasts until the sum does not exceed the required value.

Detection of the most contrasting lines (segments) is made after selection the required number of points. Algorithm investigates neighborhood points in a direction perpendicular to the gradient in order to detect the next point belonging to the boundary (edge) of the object (fig. 2-b). Points with maximum of gradients and approximately the same direction are selected as the second point (fig. 3-a, 3-d). These segments are considered as contours of the objects and used later to detect the object. The experiments showed that in most cases less than 10000 segments is necessary for the reliable identification of the objects like "line" or "box."

III. Object Recognition and Determination of Its Parameters

For object recognition in a photo image, the Hough transformation with a special modification has been applied. This transformation represents an object as a point in a four dimensional accumulator. The four dimensions correspond to the *X-Y* position of any two points in a pattern (e.g. the centre of the coordinates and the top of the object). These two points determine the position, angle and scale of the pattern relative to the source image. Pattern coordinates are connected with coordinates in the image using the following equation:

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} dx \\ dy \end{bmatrix} + \rho \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \tag{1}$$

where: $[x'; y']^T$ – the pattern coordinates vector, $[x''; y'']^T$ – the image coordinates vector, ρ – the scale coefficient, φ – the rotation angle and $[dx; dy]^T$ – the pattern offset relative to the image.

The basic idea consists of a point comparison of the image with a key point in a pattern. Only the most contrasted points in the image (with a maximum gradient) take part in this comparison. All associated points increase weight in a 4D-accumulator. The maximum weight in the accumulator corresponds to the recognized object, and value of this weight can be interpreted as a probability of the object's existence in the image. Unfortunately, there is no way to represent the condition of the 4D-accumulator in the paper, but we can use 2D-projections of this array. Fig. 4 represents source images (a, b) and accumulator states (c, d) for a recognized rectangle-marker and white line on a seabed. The maximum in the accumulators corresponds to the position of the pattern's centre in the image. We can see a similar result for an AUV navigation task solved on the basis of photo images (fig. 4).

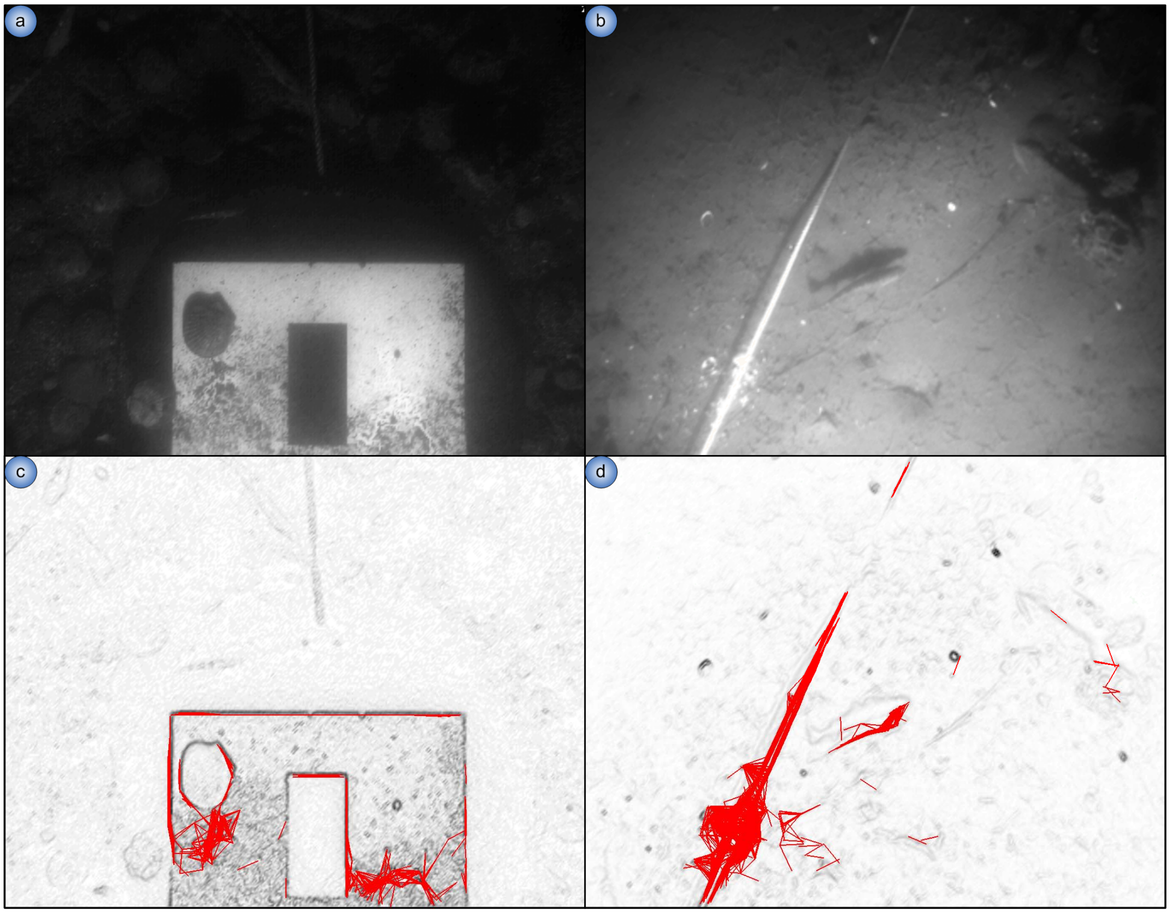

Fig. 3. Underwater photographs obtained by underwater vehicle: a) rectangular marker; b) white tether; c) a gradient map and tether edges; d) a gradient map and marker edges.

If the orientation of the image is unknown, then requires a system of four equations to identify all object parameters (*dx*, *dy*, ρ and φ). This system can be obtained from (1), knowing the coordinates of the segments belonging to the objects contours in the pattern $[x'_1; y'_1]$, $[x'_2; y'_2]$ and image $[x''_1; y''_1]$, $[x''_2; y''_2]$:

$$\begin{cases} \begin{bmatrix} x''_1 \\ y''_1 \end{bmatrix} = \begin{bmatrix} dx \\ dy \end{bmatrix} + \rho \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x'_1 \\ y'_1 \end{bmatrix} \\ \begin{bmatrix} x''_2 \\ y''_2 \end{bmatrix} = \begin{bmatrix} dx \\ dy \end{bmatrix} + \rho \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x'_2 \\ y'_2 \end{bmatrix} \end{cases} \tag{2}$$

The system (2) easy solved for $\rho*cos(\varphi)$ and $\rho*sin(\varphi)$ if equate the first equation to the second relatively $[dx; dy]^T$. Thus, we arrive at the following system of two equations with two unknowns:

$$\begin{bmatrix} x''_1 \\ y''_1 \end{bmatrix} - \rho \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x'_1 \\ y'_1 \end{bmatrix} = \\ = \begin{bmatrix} x''_2 \\ y''_2 \end{bmatrix} - \rho \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix} \begin{bmatrix} x'_2 \\ y'_2 \end{bmatrix} \tag{3}$$

$$\begin{cases} (x'_1 - x'_2)\rho\cos\varphi + (y'_1 - y'_2)\rho\sin\varphi = (x''_1 - x''_2) \\ (y'_1 - y'_2)\rho\cos\varphi - (x'_1 - x'_2)\rho\sin\varphi = (y''_1 - y''_2) \end{cases} \tag{4}$$

Denoting the coordinate difference as $(x_1'–x_2')=A$, $(y_1'–y_2')=B$, $(x_1''–x_2'')=C$, $(y_1''–y_2'')=D$ we obtain the following decision on $\rho*cos(\varphi)$ and $\rho*sin(\varphi)$:

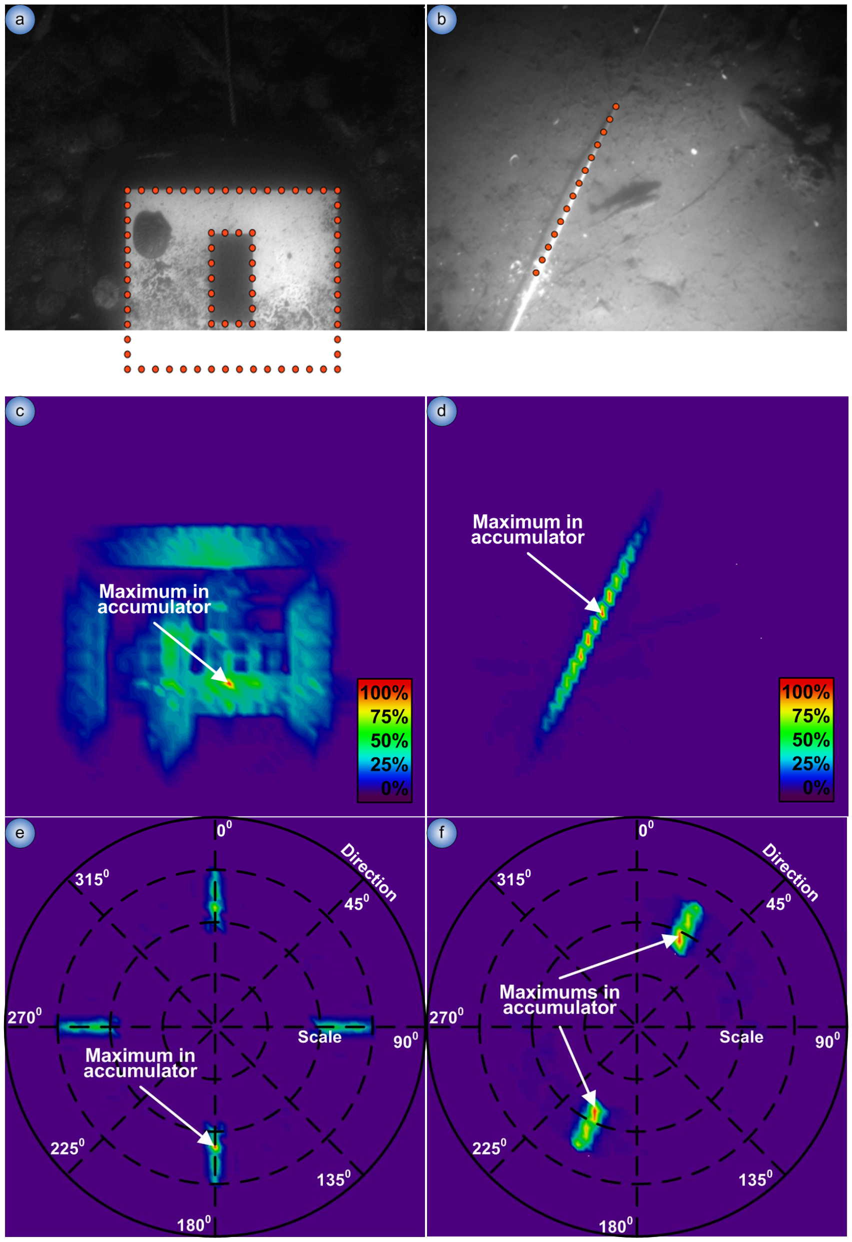

Fig. 4. Accumulator view in the plane-projection of displacement (c, d) and scale-rotation (e, f) when detected rectangular marker (a) and white rope (b).

$$\begin{bmatrix} \rho\cos\varphi \\ \rho\sin\varphi \end{bmatrix} = \frac{1}{-A^2 - B^2} \begin{bmatrix} A & B \\ B & -A \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} \tag{5}$$

Thus, to solve the problem of identification is sufficient to know the coordinates of any point of the pattern in the image (if the scale and rotation of the image is known) or the coordinates of any segment (if the scale and rotation is unknown).

Unfortunately there is no correspondence of segments on the pattern with the image segments. Because of this the algorithm has cycle through all the possible matches to determine (by the formula (1) and (5)) the necessary parameters. Obviously, if the desired object is actually present on the image, then parameters *dx*, *dy*, ρ and φ must be identical for different segments. Therefore, for the detection of the object is sufficient to create a Hough-like accumulator [10] and increase the accumulator cells with coordinates [*dx, dy, ρ, φ*] during busting segments. The cell with the maximum value corresponds to the coordinates of the object, and the higher the value in the cell, the greater the likelihood that the desired shape is present in the image.

IV. SEABED IMAGES PROCESSING AND MODELING RESULTS

Algorithm verification was performed on simulated images and on real photos obtained by underwater vehicles. Simulation experiments with models and real photos were carried out using special simulator developed in collaboration FEFU and IMTP FEB RAS. This simulation software [11] includes a model of navigation, control system and information-measuring system of the AUV and the environment. Fig. 4-a,b show real photos of white rope and rectangular marker. Seabed images used for dead reckoning are shown in fig. 5-a,b.

Experiments on the detection of the rectangular marker and white rope assumed that the location, direction and scale of the objects is unknown, ie, calculated by the algorithm. In both cases all objects was successfully recognized (fig. 4-a, fig. 4-b). Besides that, the marker was detected, even with the interruption of the object borders and absence of about 20% of the object in the picture.

To assess the reliability of the algorithm we can analyze the accumulator condition at the time of the object identification. If the accumulator has a distinct maximum and its value several times higher than a value in other cells, then it is possible to speak about the reliability of the recognized object.

Let consider in details the accumulator state during detection of white rope (fig. 4-d and 4-f). It clear to seen that maximum values in accumulators correspond to the object location (fig. 4-d) and its direction (fig. 4-f). One advantage of the algorithm is relatively small noise in accumulator cells, for example, in comparison with original Hough transform. It explained by using not all contrasting image points (and all possible line directions), but only segments corresponding to the objects boundaries.

Fig. 5-a. represents the result of the shift calculation of two consequent photos. In the center of the image the first (base) frame is located. The next (second) frame is shifted by a distance corresponding to the maximum value in accumulator (fig. 5-b). It can be seen that the maximum in accumulator is clearly defined and corresponds to the picture shift for the shooting period. Notice, the shift is correctly determined without any distinct objects with significant dimensions (there are only small objects, which differ little from each other).

V. CONCLUSION

The developed algorithm allows detecting the existence and location of the sought-for object on a seafloor. The advantages of the described method consist of the low computational complexity of the recognition algorithm. Realization of this algorithm consists of two main nested iterations (key points of the pattern, approx. 100, and key points of the image, approx. 1000~2000) and small cycles in scale and rotation. Offline image processing results and modeling experiments allow concluding that this method can be used in photo vision and navigation systems of underwater robots. The prospects of the method development include the use of color component of photographic images and applying this method to detect objects in the sonar images (like side scan and sector scan sonar echograms).

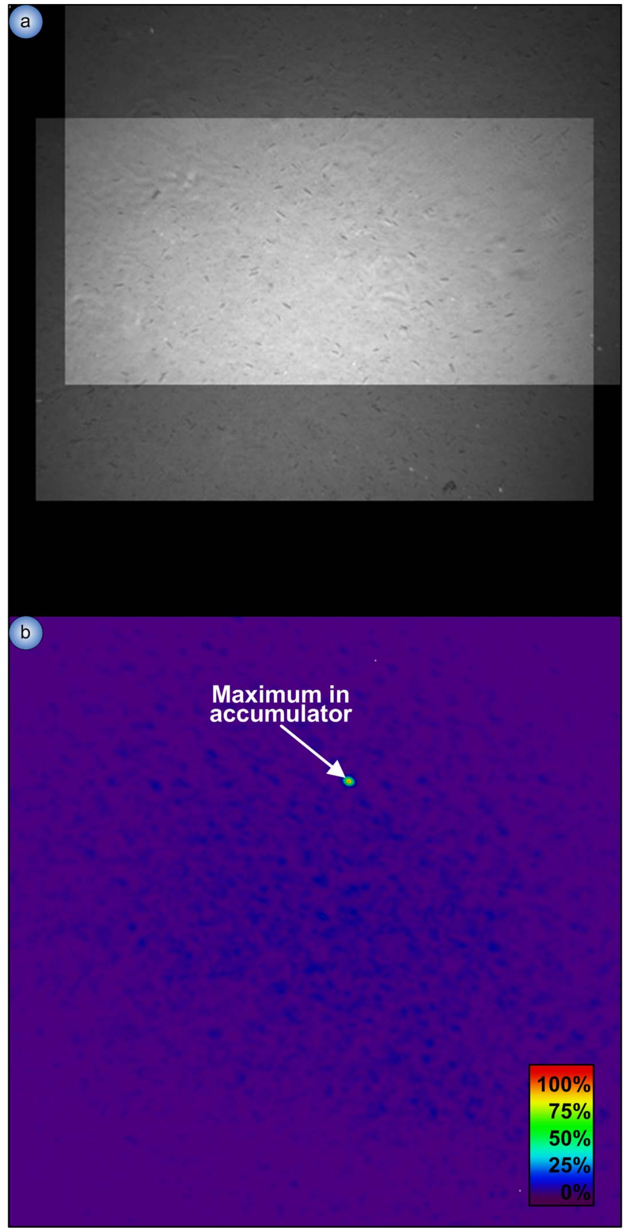

Fig. 5. AUV offset determination for two images: a) the combined sequence of two images; b) the accumulator state.

ACKNOWLEDGMENT

Author expresses his gratitude to colleagues from FEFU & IMTP FEB RAS for numerous discussions, the original data, valuable comments and assistance. Also author acknowledge Russian Science Foundation for the financial support of this study (Grant No. 14-50-00034).

REFERENCES

[1] Inzartsev A.V., Pavin A.M., Bagnitckii A.V. Actions planning and execution in the inspection underwater robot on the basis of behavioral methods // Underwater Investigation and Robotics, 2013,No 1(15), pp. 4-16.

[2] Andrey Bagnitsky, Alexander Inzartsev, Alexander Pavin, Sergey Melman, Mikhail Morozov. Side Scan Sonar using for Underwater Cables & Pipelines Tracking by Means of AUV // Proceedings of Symposium on Underwater Technology 2011, 5-8 April, 2011, Komaba Research Campus, The University of Tokyo, Tokyo, Japan, ISBN: 978-1-4577-0163-4.

[3] Alexander Inzartsev, Alexander Pavin "Experience of Underwater Cable Tracking by Means of AUV" AUV Sensors and Systems Workshop, 6-12 Nov 2010, Hawaii, USA.

[4] Scherbatyuk, A.; Boreyko, A. & Vaulin, Yu. (2000). “AUV Operation Based on Video Data Processing: Some IMTP Experience”, Workshop on Sensors and Sensing Technology for Autonomous Ocean Systems, Oct -Nov 2000, Hawaii.

[5] Alexander Inzartsev, Lev Kiselyov, Andrey Medvedev and Alexander Pavin «Autonomous Underwater Vehicle Motion Control during Investigation of Bottom Objects and Hard-to-Reach Areas» // “Motion Control” edited by Federico Casolo, In-Tech Publishers, Vienna, January, 2010, 590 pages, pp. 207-228, ISBN 978-953-7619-55-8, open access: http://sciyo.com/books.

[6] Yu N Kulchin, A Yu Mayor, D Yu Proschenko, A Yu Zhizhchenko, S S Golik, M Yu Babiy, A G Mirochnik, Modification of a new polymer photorecording material based on PMMA doped with 2,2-difluoro-4-(9-antracyl)-6-methyl-1,3,2-dioxaborine by ultrashort pulses Quantum Electronics, 2015, 45 (5), 477–481 DOI: 10.1070/QE2015v045n05ABEH015764.

[7] Kameneva P.A., Imbs A.B., Orlova T.Yu.. 2015. Distribution of DTX-3 in edible and non-edible parts of Crenomytilus grayanus from the Sea of Japan, Toxicon 98, pp. 1-3.

[8] Beleneva I.A., Shamshurina E.V., Eliseikina M.G. Assessment of the toxic effect exerted by fluorescent pseudomonads on embryos and larvae of the sea urchin strongylocentrotus nudus //Ecotoxicol Environ Saf. 2015;115:263-71.

[9] Bidyut Baran Chaudhuri, Bhabatosh Chanda The equivalence of best plane fit gradient with Robert's, Prewitt's and Sobel's gradient for edge detection and a 4-neighbour gradient with useful properties / Signal Processing, Volume 6, Issue 2, April 1984, Pages 143-151.

[10] Priyanka Mukhopadhyay, Bidyut B. Chaudhuri. A survey of Hough TransformOriginal Research Article / Pattern Recognition, Volume 48, Issue 3, March 2015, Pages 993-1010.

[11] Alexander Pavin, Alexander Inzartsev, Grigoriy Eliseenko, Oleg Lebedko, Mikhail Panin "A Reconfigurable Web-based Simulation Environment for AUV" // Proceedings of the OCEANS 2015 MTS/IEEE Conference, October 19-22, 2015, Washington, D.C., USA (in this issue).