



**HAL**  
open science

## A Camcorder for 3D Underwater Reconstruction of Archeological Objects

Arnaud Meline, Jean Triboulet, Bruno Jouvencel

► **To cite this version:**

Arnaud Meline, Jean Triboulet, Bruno Jouvencel. A Camcorder for 3D Underwater Reconstruction of Archeological Objects. OCEANS, Sep 2010, Seattle, WA, United States. 10.1109/OCEANS.2010.5664572 . lirmm-00523974

**HAL Id: lirmm-00523974**

**<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00523974v1>**

Submitted on 18 Nov 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A camcorder for 3D underwater reconstruction of archeological objects

A. Meline<sup>1</sup>, J. Triboulet<sup>1,2</sup>, and B. Jouvencel<sup>1</sup>

<sup>1</sup>LIRMM, Université de Montpellier 2 - CNRS UMR 5506

161 rue Ada, 34095 Montpellier Cedex 5, France

<sup>2</sup>Université de Nîmes, Place Gabriele Péri, 30021 Nîmes, France

**Abstract-** The underwater cartography has made great progress in the last decade. In this paper, we discuss of the 3D underwater cartography problem and propose a multimodal fusion approach. The work presented in this paper is about the analyze and 3D reconstruction of archeological objects. Using an uncalibrated single camera, we propose to describe a method with enables first, to calibrate this camera integrating variation of the water optical refractive index to the distortion model. And secondly, robust and stable features have been extracted. A comparison of two methods of detecting and matching points is presented. Finally, we estimate the movement between different images and compare two methods.

## I. INTRODUCTION

In this paper<sup>1</sup>, our motive is to analyze natural underwater scenes and especially the 3D cartography of submarine environments. Many techniques are widely used today to handle this problem. Using remotely operated vehicles in teleoperation or autonomous mode is the way underwater interventions tasks are solved. The originality of this work stand in the fusion of two kinds of maps obtained with sensors of different resolutions. An Autonomous Underwater Vehicle analyze the seabed with a lateral sonar (or a camera). From this operation results a first global map of the zone (coastal oceanographic or fluvial archaeology applications). This map is then decomposed into smaller cells composing a mosaic of the seabed. A second scanning is performed on particular cells using a second sensor with a higher resolution leading to a detailed 3D partial map. An AUV or a diver carrying a camera can perform this detailed acquisition.

This project includes two parts: the first part correspond to the 3D reconstruction of constrained submarine environment with a video camera. The Second part is the multimodal aspect of the problem and also the way the vehicle or the diver is driven inside the global map to the region of interest where a more precise acquisition is needed. Practical applications are the cartography of the seabed near pipelines or underwater sediment bed's profile analysis.

In our case, we aim at applying this method to submerged archaeological sites. The 3D precise reconstruction would be necessary where objects of interest (statues, plates, amphora...) are detected in the global map.

The work presented in this paper focuses on the first part of the project and deals with the 3D reconstruction of the submarine environment. Nowadays even if vision can solve several problems, the analysis of images in case of submarine environments is quite complicated. The methodology of 3D reconstruction presented in these lines is well known but these results were obtained with important turbidity and low luminance. In this experiment, a diver carrying a camcorder and camera acquires images. Pictures of the object were taken during a 360 degrees rotation around it.

This paper is organized as follows. Section II presents a brief survey on underwater calibration and 3D reconstruction. Section III describes the camera model and the method used to calibrate our camera. In section IV, a method to reconstruct underwater scene is presented. Section V presents the obtained experimentation results of the proposed reconstruction. Finally, conclusion and future work are presented in section VI.

## II. RELATED WORKS

3D reconstruction from image sequences is a widely studied topic in recent years in the community of computer vision. Many methods exist, but these methods cannot be applied to all underwater images. Indeed, they depend on the knowledge of the system and the environment. One example is the Microsoft Photosynth project [1] using Internet images to create a 3D urban database. This method uses a priori knowledge of the scene but has no information about the devices that took the pictures. In a different purpose, Sauvée [2] and Noce [3] have modeled the 3D movement of the heart during surgery. In this application, a priori knowledge of the observed scene cannot be used because the heart is a deformable object. However camera parameters are completely known from a camera calibration step. By focusing on underwater reconstructions, we find the work of Espiau [4] in 2002, which proposes a method to reconstruct textured scenes from an uncalibrated camera. To achieve this problem, he proposes to estimate the camera parameters from the scene rigidity. Brandou [5] works on the same images but he proposes to calibrate the camera in situ and uses a robot arm to move the camera. This enables him to know the exact camera movement and to simplify

---

<sup>1</sup> This work is supported thanks to ANR project CFLAM and ERIN team from UNIMES

the reconstruction. Hogue [6] proposes a different approach, he developed a stereo vision / inertial sensor. The system combines 3D information, resulted from using stereo vision with a 3DOF inertial sensor.

The difficulty of these methods varies depending on the constraints imposed by the application: type of scene, knowledge or not of the camera movement, calibration parameters, shape of the object. Some authors, like Hartley [7], Pratt [8] and Szeliski [9] propose general methods to reconstruct a 3D scene. Hartley and Zisserman [7] proposed a projective reconstruction method, which is divided into several phases. Initially, they extract primitives from the image that are usually points of interest. Then, the obtained points are matched. This step consists in correlating all the primitives in different images. After the matching, the projective reconstruction of the scene is carried out. Thereby they obtain a 3D scene that keeps only the projective properties of the structure.

To rebuild Euclidean structure of scenes, knowing the calibration parameters of the camera is essential. The calibration enable to establish the relationship between image dimensions and real dimensions. It can be done through knowledge of the scene geometry [10] [11] [12] [13] (parallel lines, shape and dimension of a known object, etc. ...), or knowledge of the camera movement [14] (pure rotation like a turning table) or simply through a calibration target [15].

All these methods have been developed for surface imaging, when the camera is submerged many problems linked to submarine environment are introduced. The first difficulty appears when the calibration target is used. Furthermore, another important constraint of this environment is the signal propagation variability which decreases image quality. This phenomenon is due to the presence of suspended particles (turbidity) which introduce reflection and refraction effects on the light beam. The refractive index represents the deviation of the light beam passing from one medium to another. This parameter is not fixed, it varies according to the environment characteristics such as water temperature, water salinity, light beam wavelength and camera depth. All these constraints affect the quality of the acquired images. A water refractive index variation changes the value of some internal camera model parameters and therefore all the 3D reconstruction is changed.

To avoid the two problems mentioned above, Pessel [16] propose self-calibration methods that adapts to the images used. They are based on the scene rigidity to extract data and use them in order to define the epipolar geometry between two images.

Other authors like Lavest and Rives [17] propose to calibrate the camera from the surface and adjust the settings to the underwater environment. To that end, they identified a relationship between surface parameters and the coefficient of refraction of water. The disadvantage of this method lies in the fact that the variation of the refractive index in a scuba-diving is not taken into account. It should continuously analyze the water and determine the refraction index of the water according to the depth, temperature and water salinity.

### III. CAMERA MODEL AND CALIBRATION

#### A. Modeling of the camera

Before determining the camera parameters, we must define the model that we will use, that is the pinhole model. It consists of four major frames (Fig. 1): the world frame ( $R_w$ ), the camera frame ( $R_c$ ) positioned at the optical center, the retinal frame ( $R_r$ ) placed at the plane of the image and the image frame ( $R_i$ ) in pixel. Consider the point  $\mathbf{M}$  in  $R_c$ , and the point  $\mathbf{m}$  its projection in the image. The transformations for  $\mathbf{M}$  can be described in three steps. The first transformation from  $R_w$  to  $R_c$  can be characterized by a homogeneous transformation with a rotation ( $R$ ) and a translation ( $t$ ). Those parameters are called *extrinsic parameters*.

$$M_c = T_1 \times M = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \times M \quad (1)$$

The two other steps are to move from  $R_c$  reference to  $R_r$  ( $T_2$ ) and from  $R_r$  to  $R_i$  ( $T_3$ ). Generally those two steps are combined into one transformation " $K$ " representing the matrix of *intrinsic parameters*.

$$K = T_3.T_2 = \begin{pmatrix} k_u & r & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} k_u.f & r & u_0 & 0 \\ 0 & k_v.f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \alpha_u & r & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (2)$$

In this equation,  $k_u$  and  $k_v$  represent the pixel size along x and y axes.  $u_0$  and  $v_0$  give the image frame position in pixel relative to the retina frame. The parameter  $r$  also called "skew coefficient" is the default angle that can occur between the axes u and v. Usually this term is equal to zero.

The parameters  $\alpha_u$ ,  $\alpha_v$ ,  $u_0$ ,  $v_0$  and  $r$  represent the *intrinsic parameters*. Thus the projective transformation of the camera can be determined without knowing the value of the focal length.

We can also determine the complete transformation  $P$  using intrinsic and extrinsic parameters.

$$P = \begin{pmatrix} \alpha_u & r & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \quad (3)$$

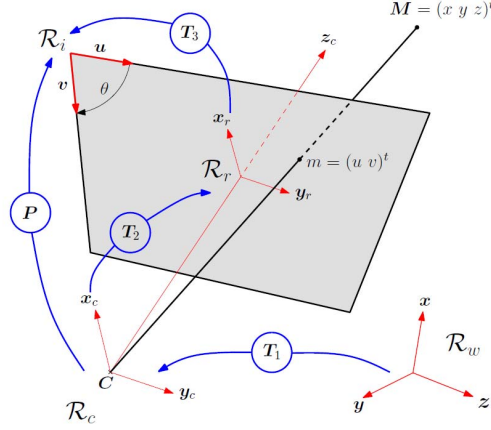


Fig. 1: Camera model and frame change

We also introduce a radial distortion term due to the submarine optical system:

$$\begin{cases} \hat{x}_n = x_0 + x_n (k_1 r^2 + k_2 r^4) \\ \hat{y}_n = y_0 + y_n (k_1 r^2 + k_2 r^4) \end{cases} \quad (4)$$

Where  $x_n$  and  $y_n$  are the coordinate of the point expressed from the image center  $(x_0, y_0)$  and  $r$  the distance between image center and current point with  $r^2 = x_n^2 + y_n^2$

#### B. Camera calibration method

We used Zhang's calibration method [15]. He proposes a technique based on the observation of a calibration plane (known dimensions) viewed from different angles, regardless of the calibration target's position. The algorithm which includes a non-linear model shows good robustness with respect to error detection in the image and errors on the geometric characteristics of the target. This non-linear method allows the integration in the linear camera model of the distortion parameters introduced by the change in environment air / water.

To calculate the camera intrinsic parameters, Zhang expressed the homography between the plane and its image. If we consider a plan  $\mathbf{M}$  and its image  $\mathbf{m}$ , the homography  $H$  is expressed as:

$$s\tilde{\mathbf{m}} = H\tilde{\mathbf{M}} \quad (5)$$

with  $H = K [r_1 \ r_2 \ t]$ ,  $K$  represents the intrinsic parameters matrix,  $r_1$  et  $r_2$  are the first and the second column of the rotation matrix ( $R$ ) and  $t$  the translation vector.

$\tilde{\mathbf{M}} = [X \ Y \ 1]$  is a 3D point. The  $Z$  coordinate equal to zero because we consider a plan.

Two constraints are expressed from the definition of the matrix  $H$  :

$$\begin{aligned} h_1^T K^{-T} K^{-1} h_2 &= 0 \\ h_1^T K^{-T} K^{-1} h_1 &= h_2^T K^{-T} K^{-1} h_2 \end{aligned} \quad (6)$$

Those two constraints can be rewritten as follows:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0 \quad (7)$$

with  $b = [B_{11} \ B_{12} \ B_{22} \ B_{13} \ B_{23} \ B_{33}]^T$  and  $B = K^{-T} K^{-1}$

$$v_{ij} = [h_{i1} h_{j1} \ h_{i1} h_{j2} + h_{i2} h_{j1} \ h_{i2} h_{j2} \ h_{i3} h_{j1} + h_{i1} h_{j3} \ h_{i3} h_{j2} + h_{i2} h_{j3} \ h_{i3} h_{j3}]^T$$

and  $h_i$   $i^{\text{ème}}$  column vector of matrix  $H$   $h_i = [h_{i1} \ h_{i2} \ h_{i3}]^T$

If "n" images of the plan are observed, "n" equations lead to the following system:

$$Vb = 0 \tag{8}$$

where V is a  $(2n \times 6)$  matrix

Vector b is estimated by solving (8) then the intrinsic camera parameters are extracted.

#### IV. 3D RECONSTRUCTION METHOD

As the geometry of the scene and the camera motion are unknown, we used the method proposed by Hartley and Zisserman [7] to achieve 3D reconstruction.

##### A. Points of interest detection

The characteristic points detection is an essential step for 3D reconstruction. The choice of these elements is crucial for the matching and for the quality of the 3D model.

There are many detection points methods such as: Beaudet operator, Moravec corner detector, Kitchen and Rosenfeld detector, Forstner operator, SURF, Harris and Stephens detector, SIFT detector... Only the last two methods are detailed below.

##### 1) Harris and Stephens detector

Harris and Stephens [18] have pointed out some limits to Moravec detector. Thus, they improved the approach by calculating a matrix related to the auto-correlation function, which takes into account the values of first derivatives of the signal on a window. This correspond to an improvement over Moravec, because it no longer requires the discretization used to compute the auto-correlation, due to displacement and to the chosen directions. As a result, Harris and Stephen propose the following operator to detect corners in an image:

$$k(x, y) = \det(C) - \lambda \cdot \text{trace}(C) \tag{9}$$

where  $\lambda$  is determined by experience, by maximizing the response on an ideal corner model:  $\lambda \approx 0, 04$ .

Matrix C is defined by:

$$C = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{10}$$

where  $I_x$  represent the image derivative along x and  $I_y$  along y.

The term  $\text{trace}(C)$  provides to edge information, while  $\det(C)$  contains "corner" information. The values of  $k(x, y)$  are positive near a corner, negative near a edge and low in a region of constant intensity.

##### 2) Scale Invariant Feature Transform (SIFT)

The SIFT method [19] is a method used to describe local image invariant to changes of scale, perspective, lighting, and particularly robust to noise in the image. This approach has been evaluated as one of the most successful in a comparative study of local descriptors. The algorithm extracts interest points in the image after a serie of treatments (pyramid of images, Gaussian differences, finding extrema). Then each point becomes a Keypoint by assigning the direction of the gradient at that point. For each Keypoint a descriptor is defined, which is a vector with 128 dimensions representing the gradient orientation near the point. A 4x4 window is defined around the point. In each zone, the gradient norm is calculated for eight different orientations. The descriptors enable to correlate SIFT points extracted in different images. The major drawback is the descriptors size, making it difficult to use in large image sequences or real-time image processing.

##### B. Matching feature points

To recover the 3D information it is essential to match informations contained in at least two images. The matching consist in finding the same point or group of points in the images. It can takes place either between primitives extracted from images or between all points of the images. The matching between primitives will mainly be used to find the transformation between two images or to facilitate a dense matching that will serve to dense 3D rendering of an object or scene. We combine a matching by correlation for the Harris points and a matching by invariant for SIFT points.

To achieve the Harris points matching, we used a correlation method. Indeed this method enables to find a point in the image even in the case of large motion between images and it doesn't takes into account the rigidity of the scene. Matching points searching is a global method. As shown in Fig. 2 only pairs of corresponding points in both directions are matched.

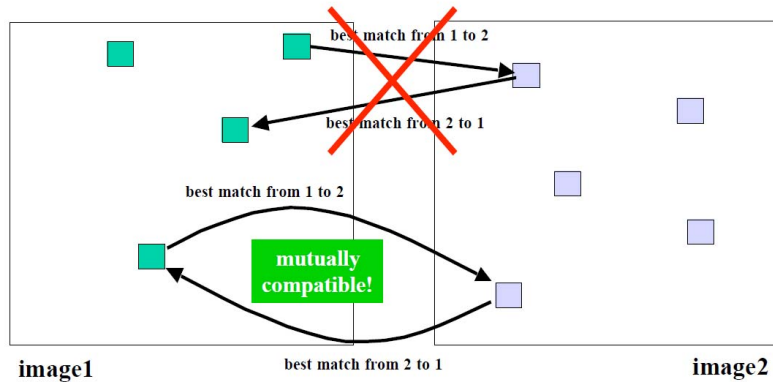


Fig. 2: Selection of matching points for the correlation method.

To improve performance and reduce the computation time, a guiding approach is done for the matching. To that end, a method of followed characteristic will be used to find an object or area in the second image. We deduce the global motion between two images and use it to limit the search points in a region of interest of the second image. Thus we also limit the selection of false matches.

The matching of SIFT points is directly issued from the invariant vectors defined by the method.

### C. Suppression of false matching and estimation of the fundamental matrix

Previous methods present some limits and can therefore provide false matches ("outliers"). To avoid using these points, algorithms such as Least Median of Squares (LMEDS) or Random Sample Consensus (RANSAC) can classify the matched data into good and false matches (inliers and outliers). These methods give results fairly close. However the RANSAC method can also estimate the fundamental matrix. Therefore, we have implemented the RANSAC method.

Fischler and Bolles [20] proposed the RANSAC algorithm designed to check the matches. The purpose of this algorithm is to define a model and identify its parameters with  $N$  sub-sets of  $n$  random data. From these  $N$  estimates of the model, a classification of matching points in good and bad matches can be done. The validation of a pairing of points is based on the calculation of the distance  $d$ . This distance is the measured error between the projection of the point from image 1 on the image 2 and the point matched in the image 2.

The algorithm used is as follow:

- Select randomly eight points and estimate the fundamental matrix
- Calculate for each point the distance between the projected point in the second image and the epipolar line. If the distance exceeds a threshold, reject the point. The remaining items are grouped in a set.
- The process is iterated  $N$  times then the set with the maximum of elements is selected. Finally, the fundamental matrix is estimated with these points.

The algorithm selects eight points because the method used to determine the fundamental matrix is the one presented by R. Hartley [13]. He improves the method of the eight points proposed by Longuet-Higgins [21] by proposing a normalization before estimating the fundamental matrix. The major advantage of this method is the linear estimation of the matrix.

Once the inliers obtained and the fundamental matrix estimated, the movement between two successive images can be deduced.

## V. EXPERIMENTATIONS AND RESULTS

### A. Underwater experimentations

A camera / camcorder was used for these experiments. The resolution used for pictures is 3264x2448 pixels and 1920x1080 for videos. A waterproof case fitted the device to take underwater pictures obtained at sea in shallow water (15m). We immersed various objects (a statue of a female bust, a statue of a fish, bowls, plates and a calibration grid) which we are able to compare with 3D scanned models. The results presented in this document are made with images of the female bust.

### B. Calibration results

As it has been introduced in section III-B, the method used for calibration is Zhang's method. Camera parameters are calculated with at least two pictures and without knowing the camera motion. Moreover, the radial distortion phenomenon is modelised. This is interesting because this distortion appears as soon as the system is submerged and may vary with the waterproof box. To achieve the different calibrations, we used the toolbox "camera calibration" developed by JY Bouquet [22] for Matlab.

The variation of the refractive index of water changes the focal length. To overcome this problem, we hypothesized that all these parameters could be incorporated into the second order radial distortion model of the camera. To justify this choice, we

made various experimentations with under-water and air photos. We calibrated the device with different sequences of images. To overcome the influences of the waterproof-box and compare only the parameters variations between water and air, we calibrated the device in air with and without the waterproof-box (see TABLE 1). We noticed that the waterproof-box had no influence on the results. Thus, we could verify that only environment change influenced the distortion parameters. As we can see in TABLE 1, the principal points  $u_0$  and  $v_0$  do not change whatever the conditions (respectively around 1600 and 1300), unlike the radial distortion parameter  $k_1$  which sign changes between calibrations in air (-0.15) and water (+0.12). In other word, the radial distortion changes significantly and moves from a pincushion distortion to a barrel distortion. As it has been demonstrated in Lavest [17], focal lengths ( $\alpha_u, \alpha_v$ ) vary from about 1.3 between air and freshwater. Regarding salt water, the coefficient is generally higher (around 1.4 in our case). TABLE 1 presents calibration results.

TABLE 1  
Results of camera calibration in diverse environment

	model parameters	Underwater pictures			Outside pictures					
					Without waterproof-box			With waterproof-box		
		Values	Errors	Errors (%)	Values	Errors	Errors (%)	Values	Errors	Errors (%)
Focal length	$\alpha_u$	4572.85	60.23	1.32 %	3200.45	37.38	1.17 %	3248.21	25.52	0.78%
	$\alpha_v$	4511.64	46.68	1.03 %	3225.97	56.66	1.66 %	3248.79	37	1.14%
Principal point	$u_0$	1666.98	26.29	1.58 %	1655.1	31.55	1.90 %	1643.3	24	1.46%
	$v_0$	1284.4	125.1	9.74 %	1343.82	29.01	2.16 %	1297.9	21.5	1.65%
Distortion	$k_1$	0.124	0.0156	12.58%	-0.142	0.0148	10.42%	-0.154	0.011	7.14%

From these calibration parameters, we can achieve a reprojection of points obtained on the calibration target images. With the points we use for calibration and those we get by reprojection, we can deduce the reprojection errors. TABLE 2 presents the reprojection errors in pixels along the axes x and y. These errors are low despite the fact that our pictures are disturbed by the underwater environment.

Thanks to the knowledge of the size and shape of the target calibration and the calibration parameters, we can deduce the extrinsic parameters of different pictures.

TABLE 2  
Calibration reprojection error along x and y axes

Error	Minimum	Maximum	Average	Standard deviation
Along x	$26.10^{-6}$	0.28	0	0.10
Along y	$22.10^{-6}$	0.36	0	0.11

### C. Detection and matching results

As we have seen in Section IV, the 3D reconstruction is divided into several steps. The first step consist in extracting characteristic points of the image. Harris points (blue) and SIFT points (green) are displayed from two different images on Fig. 3.

The advantage of using the Harris detector is that it determines the points near the structure corners. This enables us to obtain the characteristic points of the object geometry that we observe. The SIFT algorithm permits us to get points distributed more evenly over the image. These points are different from those found by the Harris method and they enable to increase the model resolution.

Fig. 4 and Fig. 5 represent the matched points between consecutive images and the estimated motion for each point between the two images with the Harris/correlation method and SIFT method.

We compared the numbers of detected and matched points from the two methods. TABLE 3 summarizes the results obtained from four pairs of images. The first three pairs of images are selected to have a small object displacement. The last image pairs show a greater movement.

TABLE 3  
Comparison of number points obtained with the two methods (Harris and SIFT)

	Method 1: Harris detector with correlation matching				Method 2: SIFT detector and matching			
	Detected points on first image	Detected points on second image	Matched points	Inliers points estimate by RANSAC	Detected points on first image	Detected points on second image	Matched points	Inliers points estimate by RANSAC
Image pair 1	805	1164	326 (40%)	289 (36%)	808	728	182 (25%)	156 (21%)
Image pair 2	913	1157	345 (38%)	300 (33%)	727	907	143 (20%)	131 (18%)
Image pair 3	380	414	143 (38%)	132 (35%)	236	261	37 (16%)	28 (12%)
Image pair 4	805	1578	213 (26%)	164 (20%)	802	1159	99 (12%)	72 (9%)

The detection threshold on both methods are adjusted so as to get roughly the same number of points with each of them. These experiments show that the method of Harris / correlation gives better results than the SIFT algorithm. Indeed, with Harris, we get about 40% of matched points for a small displacement, while we only have 20% with SIFT. With a large displacement between the images, the results are worse than in previous experiments but it can be explained by the movement which entails the loss of a larger image portion.

RANSAC algorithm enables to eliminate outliers, which represent about a 4% decrease of the useful points. With Harris, the inliers represent approximately 35% of the total points detected.

According to these results, we selected the Harris points to estimate the fundamental matrix, thus we were able to deduce the movement between the two images.

We estimated the global 2D motion in the image. TABLE 4 summarizes the different displacements estimated from all inliers.

TABLE 4  
Estimate movement from Harris points (in pixel)

		Estimate movement without fundamental matrix and RANSAC				Estimate movement with fundamental matrix and RANSAC			
		Minimum	Maximum	Mean	Std deviation	Minimum	Maximum	Mean	Std deviation
Image pair 1	Along x	-584	420	100	120	101	149	129	14
	Along y	-464	482	42	78	46	58	51	2
Image pair 2	Along x	-584	535	98	120	106	132	119	7
	Along y	-471	425	34	77	30	44	38	3
Image pair 3	Along x	-129	282	-58	42	-76	-50	-64	6
	Along y	-171	155	3	24	-4	9	3	4
Image pair 4	Along x	-613	554	265	236	321	415	368	33
	Along y	-472	424	93	134	129	149	139	7

As we can see in TABLE 4, when the estimation motion is based on all points, worse are the results. Indeed, the standard deviation of the estimated displacement for the four sequences is important, so we have a large number of errors in the movements. After using RANSAC and the fundamental matrix, we get better results with a smaller standard deviation. Fig. 4 and Fig. 5 illustrate this phenomenon, we have important errors on the image movement before using RANSAC and the fundamental matrix. All errors have almost disappeared after using of the fundamental matrix and RANSAC.

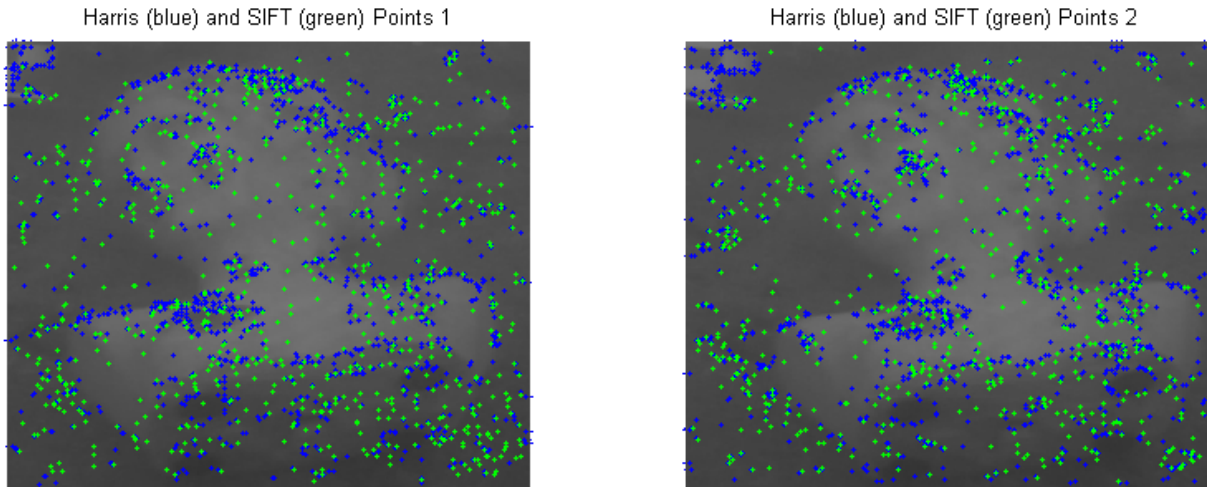
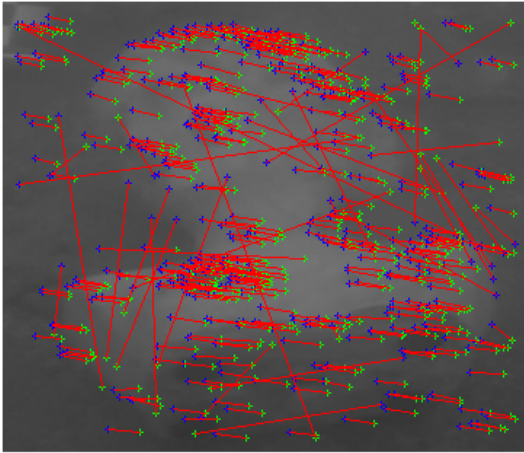


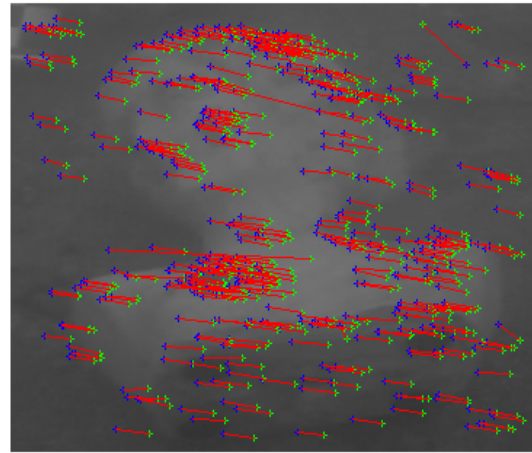
Fig. 3: Feature points detected with Harris method (blue) and SIFT method (green)



Harris Points movements (after correlation matching)

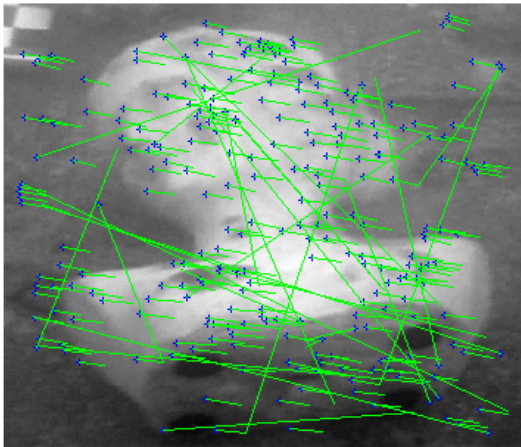


Harris Inliers points movements (after RANSAC fundamental matrix)

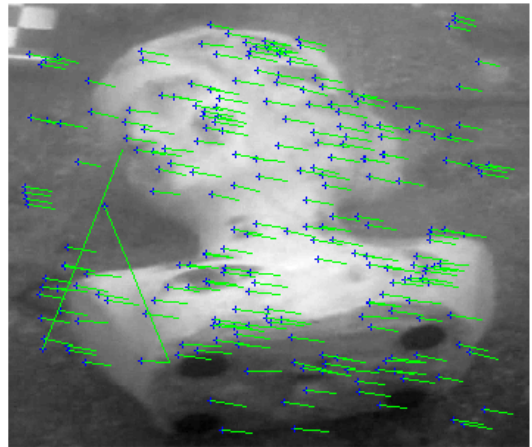


**Fig. 4: Estimate movement from Harris points between picture 1 (blue) and picture 2 (green).**

SIFT Points movements



Inliers SIFT points movements (after RANSAC fundamental matrix)



**Fig. 5: Estimate movement from SIFT points between 2 pictures.**

## VI. CONCLUSIONS AND FUTURE WORKS

The objective of this work is to obtain a good estimation of the camera calibration parameters and robust matching points to reconstruct 3D underwater archaeological objects with a single camcorder. To achieve this goal, we have integrated the underwater constraints in the camera distortion model. Then before realize the 3D point projection we have to make sure that the feature points and matching are robust enough to turbidity and low luminance. We have drawn a parallel between two methods through various experiments which enable us to single out the most robust method of the two. Then, we plan to used this robust points and camera parameters to create the 3D model with an euclidian reconstruction method. The multimodal aspect of the problem will be done thanks to a lateral sonar, GPS localization and inertial information.

## REFERENCES

- [1] N. Snavely, S. M. Seitz and R. Szeliski, "Photo tourism: Exploring photo collections in 3D", *ACM Transactions on Graphics*, 25(3), August 2006.
- [2] M. Sauvée et al., "Mitral valve leaflet motion tracking in ultrasound images", *SURGETICA'05: Computer-Aided Medical Interventions - tools and applications*, p. 237-244, 2005
- [3] A. Noce, J. Triboulet & P. Poignet, "Composite visual tracking of the moving heart using texture characterization", *Medical Robotics Workshop (MICCAI'06)*, p. 6, October 2006
- [4] F.X. Espiau and P. Rives, "Extracting robust features and 3D reconstruction in underwater images", *In Proc. OCEAN 2001*, vol. 4, p. 2564-2569, 2001
- [5] V. Brandou et al., "3D reconstruction of natural underwater scenes using the stereovision system iris", *In Proc. OCEAN 2007 - Europe*, p. 1-6, 2007
- [6] A. Hogue, A. German, J. Zacher and M. Jenkin, "Underwater 3D mapping: Experiences and lessons learned", *Computer and Robot Vision*, The 3<sup>rd</sup> Canadian Conference, p24-24, 2006.
- [7] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision", *Cambridge University Press*, 2003

- [8] W. K. Pratt, "Digital Image Processing: PIKS Inside", *John Wiley & Sons, Inc.*, 2001
- [9] R. Szeliski, "Computer Vision: Algorithms and Applications", unpublished, 2010
- [10] C. Zeller and O. Faugeras, "Camera Self-Calibration from Video Sequences: the Kruppa Equations Revisited", technical report n°2793, INRIA, February 1996.
- [11] M. Lourakis and R. Deriche, "Camera Self-Calibration Using the Singular Value Decomposition of the Fundamental Matrix: from Point Correspondences to 3D measurements", Technical report n°3748, INRIA, August 1999.
- [12] P. Mendonça and R. Cipolla, "A Simple Technique for Self-Calibration", *In Proceedings of Conference on Computer Vision and Pattern Recognition*, Vol. I, pp. 500-505, Fort Collins, Colorado, June 1999.
- [13] R. Hartley, "In Defence of the 8-points Algorithm", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 19(6), pp. 580-593, June 1997.
- [14] L. Dron, "Dynamic Camera Self-Calibration from Controlled Motion Sequences", *In Proc. of the Conference on Computer Vision and Pattern Recognition*, New-York, USA, pp. 501-506, IEEE Computer Society Press, 1993.
- [15] Z. Zhang, "A Flexible New Technique for Camera Calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, 22, 1330-1334, 2000.
- [16] N. Pessel, J. Opderbecke and M. J. Aldon, "Camera self-calibration in underwater environment", *In International Conferences in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, 2003
- [17] J.-M. Lavest, G. Rives and J.-T. Lapresté, "Underwater Camera Calibration", *In Proc. of European Conference on Computer Vision*, Vol. 2, pp. 654-668, January 2000.
- [18] C. Harris and M. Stephens, "A combined corner and edge detection", *In Proc. of The Fourth Alvey Vision Conference*, p. 147-151, 1988.
- [19] D.G. Lowe, "Distinctive image features from scale-invariant keypoints", *In International Journal of Computer Vision*, volume 60, 2, p. 91-110, 2004.
- [20] M. Fischler and R. Bolles, "Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography", *Communications of the ACM*, 24 :381-385, 1981
- [21] H. C. Longuet-Higgins: "A computer algorithm for reconstructing a scene from two projection", *Nature*, 293(10): p. 133-135, September 1981
- [22] J. Y. Bouguet and P. Perona, "Camera calibration from points and lines in dual-space geometry", technical report, California Institute of Technology, 1998