Performances Analysis of Underwater Image Preprocessing Techniques on the Repeatability of SIFT and SURF Descriptors

Amine Mahiddine, Julien Seinturier, Jean-Marc Boï, Pierre Drap, Djamal Merad
LSIS umr CNRS 7296
Centre National de la Recherche Scientifique
Marseille, France
surname@univ-amu.fr

ABSTRACT

ROV 3D project aims at developing innovative tools which link underwater photogrammetry and acoustic measurements from an active underwater sensor. The results will be 3D high resolution surveys of underwater sites. The new means and methods developed aim at reducing the investigation time *in situ*, and proposing comprehensive and non-intrusive measurement tools for the studied environment.

In this paper, we made an investigation to find at first a pre-processing method of underwater images that do not require a priori knowledge of the scene in order to increase the repeatability of SIFT and SURF descriptors and, in a second time, finding a method to compute distances which will be less costly in terms of execution time for finding corresponding points.

Keywords

Relative orientation, SIFT, SURF, K-nearest neighbour, Automatic Color Equalization, IACE, Correlation.

1. INTRODUCTION

ROV3D[1] project goal is to develop automated proceedings of 3D surveys, dedicated to underwater environment, using both acoustic and optic sensors. The acoustic sensor allows acquiring a great amount of low resolution data, whereas the optic sensor (close range photogrammetry) allows acquiring a low amount of high resolution data. In practice, a 3D acoustic scanner produces a range wide scan of the scene, and an optic system allows a high resolution restitution (larger scale) of different areas in the scene.

In underwater environment, the image quality is degraded by the significant changes undergone by the light. This is due to two factors: first water absorption by the suspended and dissolved materials, and diffusion, mainly due to particles that scatter the radiation [Pet08a] [Que04a].

When the light crosses the dioptre air / water, one part is reflected while the rest effectively penetrates into the water. However the amount of light that penetrates the water decreases with the height of the water column crossing because water molecules absorb a certain amount of light (which reduces its energy). As a first result, underwater images are becoming darker with increasing depth. Not only the amount of light is reduced with depth, but also the light undergoes a change color depending on the amount of water crossing. The wavelength corresponding to red disappears after a few meters (see Fig.1) and beyond 25m only blue remains [Sch10a].

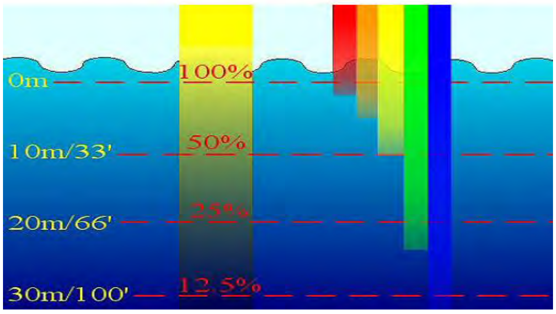


Figure 1. Colour appearance in underwater. [Iqb07a].

[1] http://www.rov3D.eu

Beyond the excessive amount of blue, underwater images therefore have a low brightness and contrast [Sch04a]. They are more affected by suspended particles called sometime « *Marine snow* ».

One of the most important issues in this work is to obtain an image quality for analysis, measurement, and extracting points of interest to optimize image processing and their orientation for the photogrammetric use.

2. UNDERWATER IMAGE PRE-PROCESSING

The underwater image pre-processing can be addressed from two different points of view: image restoration techniques or image enhancement methods.

Fan *et alii* proposed a restoration method based on blind deconvolution and the theory of Wells [Fan10a]. As a first step an arithmetic mean filter is used to perform image denoising, and then an iterative blind deconvolution using the filtered image is carried out. The calculation of the PSF of water is done using the following equations:

$$b = c\omega \quad (1)$$

$$H_{medium}(\psi, R) = \exp\left\{-cR + bR\left[\frac{1-\exp(-2\pi\theta_0\psi)}{2\pi\theta_0\psi}\right]\right\} \quad (2)$$

where θ_0 is referred to the median scattering angle, ψ is the spatial frequency in cycles per radian, R is distance between sensor and object, b scattering coefficient, c attenuation coefficient and albedo ω.

Image restoration techniques need some parameters such as attenuation coefficients, scattering coefficients and depth estimation of the object in a scene. For this reason in our works, the preprocessing of underwater image is devoted to image enhancement methods, which do not require *a priori* knowledge of the environment.

Bazeille *et alii* [Baz06a] proposed an algorithm to enhance underwater image, this algorithm is automatic and requires no parameter adjustment to correct defects such as non-uniform illumination, low contrast and muted colors.

In this algorithm which is based on the enhancement, each disturbance is corrected sequentially. The first step is to remove the *moiré* effect is not applied, because in our conditions this effect is not visible. Then, a homomorphic filter or frequency is applied to remove the defects of non-uniformity of illumination and to enhance the contrast in the image.

Regarding the acquisition noise, often present in images, they applied a wavelet denoising followed by anisotropic filtering to eliminate unwanted oscillations. To finalize the processing chain, a dynamic expansion is applied to increase contrast, and equalizing the average colors in the image is being implemented to mitigate the dominant color. Fig.2 shows the result of applying the algorithm Bazeille *et alii.*

To optimize the computation time, all treatments are applied on the component *Y* in *YCbCr* space. However the use of homomorphic filter changes the geometry, which will add errors on measures after the 3D reconstruction of the scene, so we decided not to use this algorithm.

(a) (b)

Figure 2. Images before (a) and after (b) the application of the algorithm proposed by Bazeille et alii. (Photo by Olivier Bianchimani on the Arle-Rhone 13 roman wreck in Arles, France)

Iqbal *et alii* have used slide stretching algorithm both on RGB and HIS color models to enhance underwater images [Iqb07a]. There are three steps in this algorithm (see Fig.3).

First of all, their method performs contrast stretching on *RGB* and then it converts the result from *RGB* to *HSI* color space. Finally, it deals with saturation and intensity stretching. The use of two stretching models helps to equalize the color contrast in the image and also addresses the problem of lighting.

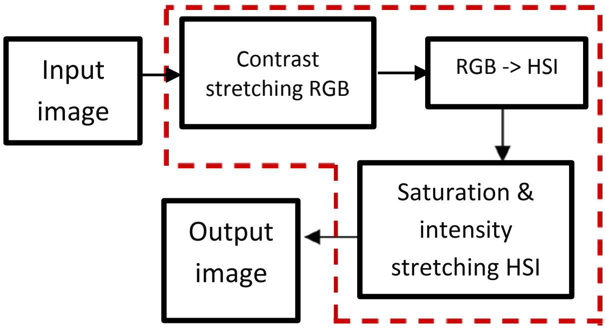


Figure 3. Algorithm proposed by Iqbal et alii .

Chambah *et alii* proposed a method of color correction based on the *ACE* model [Riz04a]. *ACE* "*Automatic Color Equalization"* is based on a new calculation approach, which combines the Gray World algorithm with the Patch white algorithm, taking into account the spatial distribution of information color. The ACE is inspired by human visual system, where is able to adapt to highly variable lighting conditions, and extract visual information from the environment [Cha04a].

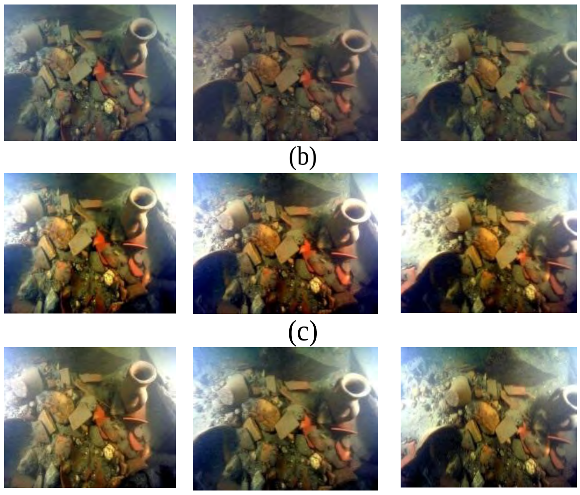
(a) *(b)*

Figure 4. Phtographs of the wreck Arles-Rhône 13, before (a) and after (b) the enhancement by ACE method.[Cha04a].

This algorithm consists of two parts. The first one consists in adjusting the chromatic data where the pixels are processed with respect to the content of the image. The second part deals with the restoration and enhancement of colors in the output image [Pet10a]. The aim of improving the color is not only for better quality images, but also to see the effects of these methods on the *SIFT* or *SURF* in terms of their feature points detection. Three examples of images before and after restoration with *ACE* are shown in Fig.4.

Kalia et alii [Kal11a] investigated the effects of different image pre-processing techniques which can affect or improve the performance of the SURF detector [Bay08a]. And they proposed new method named IACE 'Image Adaptive Contrast Enhancement'. They modify this technique of contrast enhancement by adapting it according to the statistics of the image intensity levels.

(a)



(b)

(c)

(d)

Figure 5. Phtographs of the wreck Arles-Rhône 13, (a) original images, (b) results by ACE method, (c) results by IACE method « Image Adaptative Contrast Enhancement », (d) results by the method proposed by Iqbal et alii.[Iqb07a].

If P_{in} is the intensity level of an image, it is possible to calculate the modified intensity level P_{out} with equation (3).

$$P_{out} = \frac{(P_{in} - c)}{(d - c)} \times (b - a) \tag{3}$$

where a is the lowest intensity level in the image and equal to 0, b is its corresponding counterpart and equal to 255 and c is the lower threshold intensity level in the original image for which the number of pixels in the image is lower than 4% and d is the upper threshold intensity level for which the number of pixels is cumulatively more than 96%. These thresholds are used to eliminate the effect of outliers, and improve the intrinsic details in the image while keeping the relative contrast.

The results of this algorithm are very interesting. One can observe that the relative performance of IACE method is better than the method proposed by Iqbal *et alii* in terms of time taken for the complete detection and matching process

3. FEATURE EXTRACTION AND MATCHING

The purpose of preprocessing is improving the quality of images to enhance the detection of interest points. Thereafter, these points of interest will be matched and used for 3D reconstruction of the scene.

There are several methods for extracting interest points such as Edge detector, Corner detector [Guo09a]. Juan *et alii* [Jua09a] made a comparison between *SIFT*, *PCA-SIFT* and *SURF*.

In our work, we decided to use two methods most robust in terms of invariance to the transformation and distortion of images: Scale Invariant Feature Transform "*SIFT*" and speeded-Up Robust Features "*SURF*".

Scale-invariant feature transform "*SIFT*" is a detector and descriptor at the same time proposed by Lowe [Low04a]. it is a method of extracting points of interest that are invariant to changes during image acquisition, these points of interest are local maxima or minima of the difference of Gaussians.

Each point has detected a descriptor vector which is the norm and direction of the gradient in the region around the point of interest. [Lin09a]

Speeded-Up Robust Features "*SURF*" proposed by [Bay08a] is a descriptor invariant to change of scale, rotation and image, this method is divided into two parts, the first part is devoted to the detection of points of interest, where in each scaling the local maxima are calculated using the Hessian matrix. From these local maxima, we choose the candidate points that are above a given threshold which will subsequently be invariant to scaling.

The purpose of the second part of this algorithm is to find a descriptor that will make the points detected invariant to rotation, the *SURF* descriptor is much faster but less robust than *SIFT* and can therefore be used in applications for real time processing.

After using *SIFT* and *SURF* to extract features from images, we implemented some methods for measuring distances, (see Table 1). These methods are often used to compute the similarity between points in a source image and points in a target image. Functions *SSD*, and *SAD*, *NSSD* calculate the level of dissimilarity between two points where the best result corresponds to the minimum value obtained after the computation.

We also added the method proposed by Lowe, this method is based on the K-Nearest Neighbour algorithm (*KNN)* with a modified search using the *kd-tree* to find corresponding points using Euclidean distance and optimize the calculation time.

Distance	Definition
Sum of Squared Differences **SSD**	$\sum_u (f_i(u) - f_j(u))^2$
Normalise Sum of Squared Differences **NSSD**	$\sum_u \left(\frac{(f_i(u)-\overline{f}_i)}{\sqrt{\sum_u ((f_i(u)-\overline{f}_i))^2}} - \frac{(f_j(u)-\overline{f}_j)}{\sqrt{\sum_u ((f_j(u)-\overline{f}_j))^2}} \right)^2$
Sum of Absolute Distances **SAD**	$\sum_u \left| f_i(u) - f_j(u) \right|$
Euclidean Distance	$\sqrt{\sum_{u=1}^{n} (f_i(u) - f_j(u))^2}$

Table 1 Distance measurements.

4. EXPERIMENTS

In our experiments, we took two sets of 14 images taken by photographer Olivier BIANCHIMANI in July 2012 on the *Arles-Rhone 13* roman wreck in the Rhodano river, south of France. We reduced the resolution of these photographs to 639 x 425 pixels in order to reduce computation time. These two sets are taken in two different situations. We choose for the first, a scene where we can see in Fig.2 the presence of wood pieces and in the second (see Fig.5), a scene with amphora and stones. The choice of these situations was to work on a real underwater scene and test the robustness of detectors and descriptors in the different conditions that we can cross in a marine environment.

The implementation was run on an *Intel Core i7* CPU *980* at *3.33* GHZ with *12*GB of *RAM* under Windows 7 operating system. We studied the effects of different methods that can affect or improve the performance of repeatability of a descriptor. Initially, we noticed improvements in color quality and we also see that the algorithm proposed by *Iqbal et alii* gives the best visual results.

Our approach is to detect points of interest on all images using *SIFT* or *SURF* descriptors. Subsequently, images are matched two by two with one of methods of distance measurement mentioned above in Table 1. For each matched pair of images, the relative orientation is computed using the *5* points algorithm proposed by [Ste06a].

From these orientations, an approximate value of orientations and coordinates of object points are calculated. Then a bundle adjustment is applied for optimal estimation of orientation parameters and 3D coordinates as illustrated in Fig.14.

We cannot give the results for all tests because of the space limitations. In the Table 2 and Table 3, we present some results obtained after several tests. These tables summarize the tests performed with SURF and SIFT descriptors on the original images and preprocessed images, the purpose of these tests as a first step is to find the best preprocessing algorithm in terms of color correction and preprocessing time and which mainly increases the repeatability of descriptors. In a second step, we seek to find the most appropriate method for calculating distances with the type of images that we used in our work which will give more points matched and remove outliers.

We judge the quality of these descriptors according to the number of image pairs oriented, the number of corresponding points and the reprojection error calculated both with the Root Mean Square (*RMS)* and the average error methods. We found that the *RMS* is always less than 0.5 then we have focused only on the average error (see Fig.7 and Fig.9).

The results obtained with images from the three preprocessing methods are better in comparison to results obtained with original images. However, the *ACE* took an hour and *35* seconds, for the same image *IACE* took *0.13* seconds and the method proposed by Iqbal *et alii* took *0.15* seconds almost the same time as the *IACE* method.

Before choosing the best method of preprocessing to be used in our future work, we started first by the choice of method of measuring distances where it was found that the method used by D. Lowe which is based on the algorithm *KNN* performed best in terms of points matches and computation time (see Fig.6, Fig.8, Fig.10, Fig.12), otherwise the *SSD* method and its normalized version *NSSD* also produce good results in terms of matched points and the number of pairs oriented but requires more time for the computation.

Finally after several tests, we found that the *IACE* and the method proposed by Iqbal *et alii* are quite efficient in terms of preprocessing time and number of matched points. However we cannot make a choice between these methods because the results depend on image quality and nature of objects which are located in the scene. In Table 2 and Table 3 we presented the results obtained in two different situations of a marine environment, where the *IACE* method with *SIFT* and *SUFT* descriptors gave the best result for the first situation. However, the second situation the method proposed by Iqbal *et alii* with *SURF* descriptor gave *671* matched points against *413* matched points with *IACE* and the same descriptor.

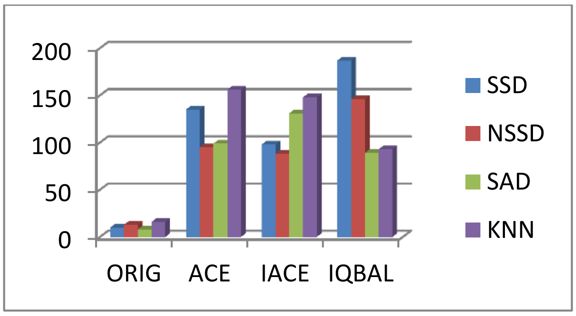


Figure 6 Matching points obtained with SURF descriptors in the first scene with presence of wood pieces.

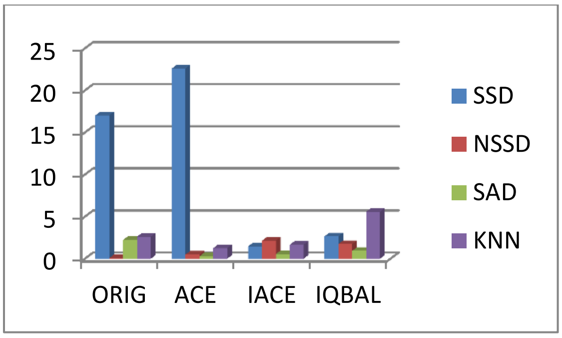


Figure 7 Average error.

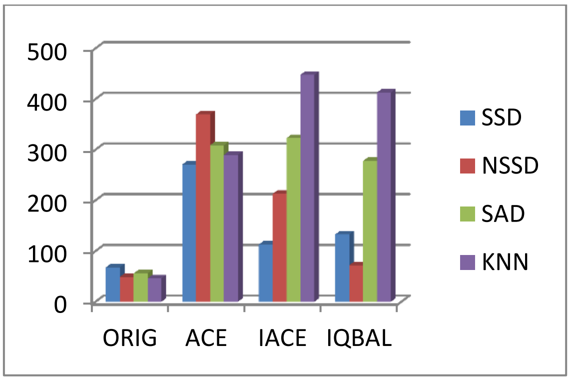


Figure 8 Matching points obtained with SIFT descriptors in the first scene with presence of wood pieces.

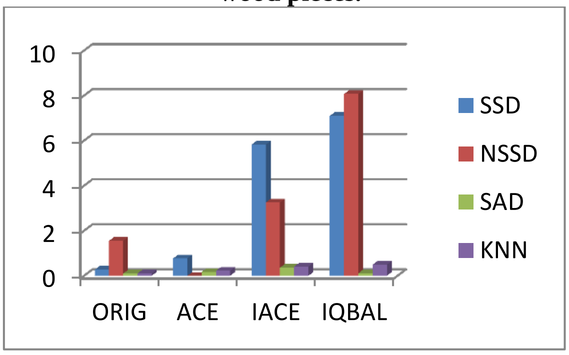


Figure 9 Average error.

SURF		SSD	NSSD	SAD	KNN
ORIG 630x425 pixels	Pairs /91	2	2	2	3
	Time (s)	22	22	20	21
	RMS	0.05	0.06	0.01	0.05
ACE 1h35	Pairs /91	14	10	7	9
	Time (s)	67	122	36	49
	RMS	0.05	0.04	0.35	0.036
IACE 0.13s	Pairs /91	12	13	9	8
	Time (s)	85	184	44	66
	RMS	0.07	0.07	0.03	0.03
IQBAL 0.15s	Pairs /91	14	18	7	8
	Time (s)	78	134	42	52
	RMS	0.04	0.05	0.03	0.05

(a)

SIFT		SSD	NSSD	SAD	KNN
ORIG 630x425 pixels	Pairs /91	4	5	4	4
	Time (s)	103	113	96	99
	RMS	0.03	0.05	0.01	0.011
ACE 1h35	Pairs /91	16	25	9	11
	Time (s)	253	253	178	238
	RMS	0.04	0.04	0.12	0.01
IACE 0.13s	Pairs /91	17	29	13	11
	Time (s)	268	1493	201	275
	RMS	0.11	0.07	0.01	0.009
IQBAL 0.15s	Pairs /91	15	11	12	12
	Time (s)	311	1413	195	264
	RMS	0.09	0.11	0.01	0.01

(b)

Table 2 Test on a set of photographs of a scene with wood pieces, (a) results with SIFT (b) results with SURF.

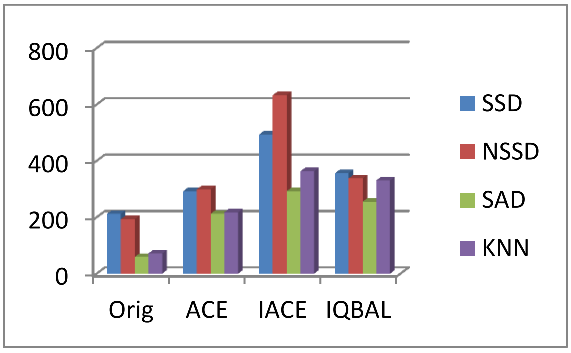


Figure 10 Matching points obtained with SURF descriptor in the second scene with presence of amphora and stones.

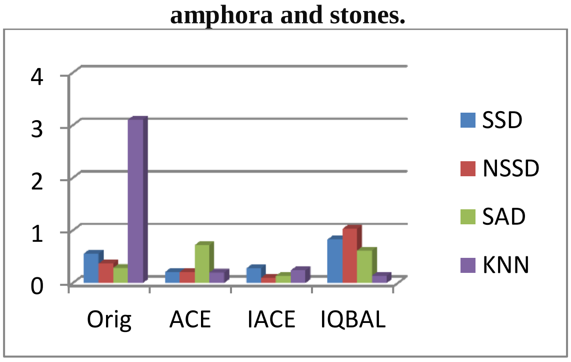


Figure 11 Average error.

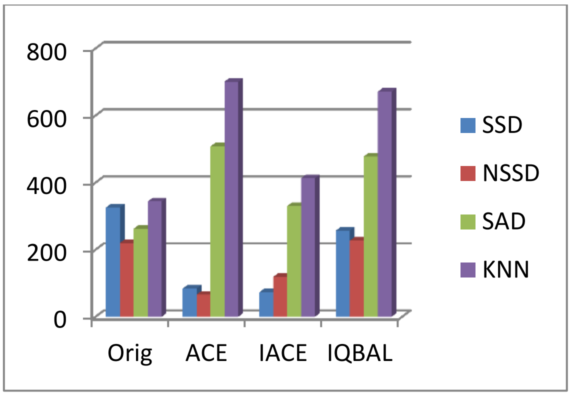


Figure 12 Matching points obtained with SIFT descriptor in the second scene with presence of amphora and stones.

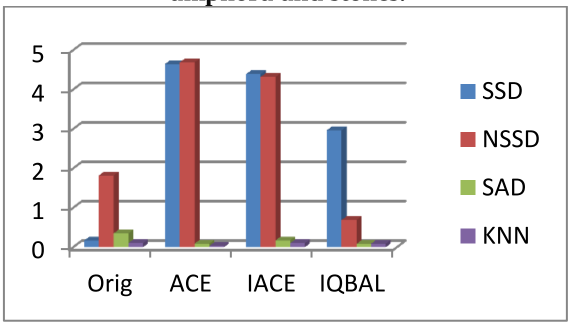


Figure 13 Average error.

SURF		SSD	NSSD	SAD	KNN
ORIG 630x425 pixels	Pairs /91	17	15	6	7
	Time (s)	51	47	26	32
	RMS	0.06	0.05	0.05	0.04
ACE 1h35	Pairs /91	11	11	12	14
	Time (s)	83	193	50	75
	RMS	0.02	0.02	0.03	0.03
IACE 0.13s	Pairs /91	11	12	12	15
	Time (s)	106	356	69	104
	RMS	0.01	0.01	0.02	0.02
IQBAL 0.15s	Pairs /91	14	13	11	14
	Time (s)	99	274	64	96
	RMS	0.03	0.02	0.02	0.02

(a)

SIFT		SSD	NSSD	SAD	KNN
ORIG 630x425 pixels	Pairs /91	15	11	11	11
	Time (s)	152	192	114	134
	RMS	0.02	0.02	0.01	0.01
ACE 1h35	Pairs /91	12	11	15	16
	Time (s)	218	742	163	214
	RMS	0.09	0.11	0.01	0.007

IACE 0.13s	Pairs /91	11	17	11	15
	Time (s)	273	1359	198	269
	RMS	0.07	0.06	0.01	0.01
IQBAL 0.15s	Pairs /91	20	16	12	16
	Time (s)	259	1216	188	260
	RMS	0.04	0.04	0.009	0.009

(b)

Table 3 Test on a set of photographs of a scene with amphora and stones, (a) results with SIFT (b) results with SURF.

As we said earlier, the number of matched points is not a sufficient factor to judge the repeatability of descriptors, for this we calculated the reprojection error to estimate the accuracy of our calculations. To minimize the reprojection error, we thought of an improvement in the quality of matched points, the idea is to correlate in a search window centered on each point matched, if the score of the correlation is less than a threshold then the point is kept if it is the nearest or it is replaced by the nearest point. Otherwise if the score is higher than threshold, the two points are deleted from the list [Kra97a].

Table 4, presents the results of a test applied to a pair of images where we find *200* matched points and a reprojection error of *1.79*. After applying the correlation, *160* points have moved and *40* points are deleted and the new reprojection error is reduced to *0.74* which means that the correlation corrects the quality of matched points. The disadvantage of this improvement is that it can be applied only on a pair of images with very weak relative rotation and a slight change of scale which is the case in our experiments and in the field of stereovision.

Features in left	Features in right	Matched points	RMS	Shifted points	Removed points	RMS minimized
749	696	200	1.79	160	40	0.74

Table 4 Table showing the result after application of correlation.

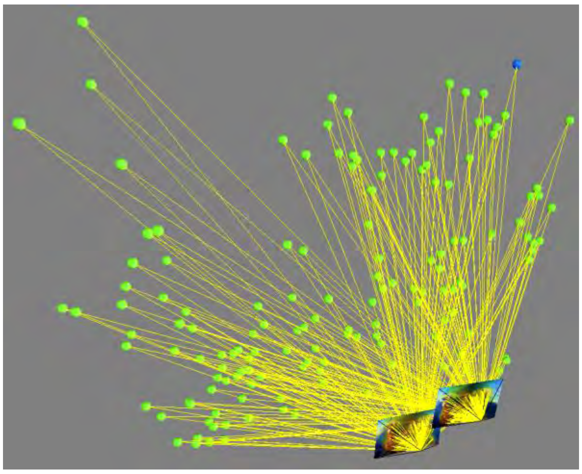


Figure 14 Example of the orientation of a pair of images.

5. CONCLUSION & FUTURE WORK

In this paper, we studied three preprocessing methods whose purpose was to improve color and contrast of underwater images and increase repeatability of descriptors compared to original images. We have also presented some methods for measuring distances where we found that the *IACE* method and the method proposed by Iqbal *et alii* give almost the same results in terms of computation time and repeatability of *SIFT* and *SURF* descriptors.

The use of one of these methods as an initial method of preprocessing with the *KNN* method for distance measurements gives good results in terms of computation time and the reprojection error compared to results obtained with images without preprocessing. Nevertheless, the *ACE* method is very slow in terms of preprocessing time, however we observed an improvement of color contrast and a brightness correction. For this reason, we plan to use the images obtained as texture after the full 3D reconstruction of the underwater scene.

We also showed in this paper, the usefulness of the correlation to minimize the reprojection error in the case of a small rotation between images. The future work is to improve the algorithm of SIFT and test Earth Mover's Distance to find the corresponding points.

6. REFERENCES

[Bay08a] Herbert Bay, Andreas Ess, Tinne Tuytelaars, & Luc Van Gool. (2008). Speeded-Up Robust Features (SURF). Comput. Vis. Image Underst., Vol.:110, pp.346-359, isbn/issn:1077-3142.

[Baz06a] Stéphane Bazeille, Isabelle Quidu, Luc Jaulin, & Jean-Phillipe Malkasse. (2006, 16 - 19 Octobre 2006). Automatic Underwater Image Pre-

Processing. Paper presented at the CMM’06 - CARACTERISATION DU MILIEU MARIN

[Cha04a] Majed Chambah, Dahbia Semani, Arnaud Renouf, Pierre Courtellemont, & Alessandro Rizzi. (2004). Underwater Color Constancy : Enhancement of Automatic Live Fish Recognition. Paper presented at the 16th Annual symposium on electronic imaging,, United States.

[Fan10a] Fan Fan, Yang Kecheng, Fu Bo, Xia Min, & Zhang Wei. (2010, 9-11 April 2010). Application of blind deconvolution approach with image quality metric in underwater image restoration. Paper presented at the Image Analysis and Signal Processing (IASP), 2010 International Conference on. pp.236-239.

[Guo09a] Chenguang Guo, Xianglong Li, Linfeng Zhong, & Xiang Luo. (2009, 21-22 Nov. 2009). A Fast and Accurate Corner Detector Based on Harris Algorithm. Paper presented at the Intelligent Information Technology Application, 2009. IITA 2009. Third International Symposium on. Vol.:2, pp.49-52.

[Iqb07a]) Kashif Iqbal, Rosalina Abdul Salam, Azam Osman, & Abdullah Zawawi Talib. (2007). Underwater Image Enhancement Using an Integrated Colour Model. IAENG International Journal of Computer Science, Vol.:34.

[Jua09a] Luo Juan, & Oubong Gwon. (2009). A Comparison of SIFT, PCA-SIFT and SURF. International Journal of Image Processing (IJIP), Vol.:3, pp.143-152.

[Kal11a] Robin Kalia, Keun-Dong Lee, Samir B.V.R., Sung-Kwan Je, & Weon-Geun Oh. (2011, 9-11 Fev.). An analysis of the effect of different image preprocessing techniques on the performance of SURF: Speeded Up Robust Feature. Paper presented at the 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV). pp.1-6.

[Kra97a] Karl Kraus. (1997). Photogrammetry vol 1 & 2 (Stewardson Peter, Trans. Dummlerbush ed. Vol. 2). Bonn, Germany.

[Lin09a] Andrea Lingua, Davide Marenchino, & Francesco Nex. (2009). Performance Analysis of the SIFT Operator for Automatic Feature Extraction and Matching in Photogrammetric Applications. Sensors, Vol.:9, pp.3745-3766, isbn/issn:1424-8220.

[Low04a]David Lowe. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, Vol.:60, pp.91-110.

[Pet08] Frédéric Petit, Philippe Blasi, Anne-Sophie Capelle-Laizé, & Jean-Christophe Burie. (2008). Underwater Images Enhancement by Light Propagation Model Reversion. Paper presented at the IS&T Fourth european conference on Color in Graphics Imaging and Vision, Terrassa : Espagne.

[Pet10] Frédéric Petit. (2010). Traitement et analyse d’images couleur sous-marines : modèles physiques et représentation quaternionique. Doctorat, Sciences et Ingénierie pour l'Information, Poitier.

[Que04a]Jose P. Queiroz-Neto, Rodrigo Carceroni, Wagner Barros, & Mario Campos. (2004). Underwater Stereo. (Conférencier invité) Proceedings of the Computer Graphics and Image Processing, XVII Brazilian Symposium.

[Riz04a] Alessandro Rizzi, & Carlo Gatta. (2004). From Retinex to Automatic Color Equalization: issues in developing a new algorithm for unsupervised color equalization. Journal of Electronic Imaging, Vol.:13, pp.75-84.

[Sch04a] Yoav Y. Schechner, & Nir Karpel. (2004). Clear Underwater Vision. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), Vol.:1, pp.536-543.

[Sch10a] Raimondo Schettini, & Silvia Corchs. (2010). Underwater Image Processing: State of the Art of Restoration and Image Enhancement Methods. EURASIP Journal on Advances in Signal Processing, Vol.:2010, pp.14.

[Ste06a] Henrik Stewenius, C. Engels, & David Nister. (2006). Recent developments on direct relative orientation. Isprs Journal of Photogrammetry and Remote Sensing, Vol.:60, pp.284-294, isbn/issn:0924-2716.