Euclidean 3D Reconstruction of Unknown Objects from Multiple Images

Soulaiman El Hazzat
LIIAN, Dep. Of Mathematics and informatics
Faculty of Sciences Dhar-Mahraz
Fes, Morocco
soulaiman.elhazzat@yahoo.fr

Abderrahim Saaidi
LIIAN, Dep. Of Mathematics and informatics
Faculty of Sciences Dhar-Mahraz
Fes, Morocco
saaidi.abde@yahoo.fr

Khalid Satori
LIIAN, Dep. Of Mathematics and informatics
Faculty of Sciences Dhar-Mahraz
Fes, Morocco
khalidsatorim3i@yahoo.fr

***Abstract*—In this paper, we are interested in the problem of Euclidean 3D reconstruction of unknown objects by passive stereo vision method. Our method is based on the combination between Harris and Sift interest point detectors, to take advantage of the power of these two detectors, which will be useful when matching step, as a key step for 3D reconstruction, In order to have a sufficient number of matches distributed on the images. These matches will be used to estimate the 3D points (the projection matrices will be estimated after calibration using 3D Calibration Pattern). Finally, a 3D mesh is constructed by 3D Delaunay triangulation, applied to the 3D points reconstructed. Experimental results prove that this method is practical and gives satisfying results without going through the propagation step.**

***Index Terms*—Interest Points, Matching, Calibration, 3D reconstruction, 3D mesh, 3D Delaunay triangulation.**

I. Introduction

3D reconstruction from 2D images is an important problem in computer vision. Many approaches offer solutions to this problem : Stereo vision [1], Structure from motion [2 ,3], Shape from silhouettes[4], shape from shading [5, 6], and shape from texture [7].

In this paper, we are intersted in the problem of multi-view reconstruction from images by passive stereo vision methods.

In this method, the 3D information is estimated, only from images taken by cameras without any controlled light. However, the active stereo vision [8] uses a controlled light source such as a laser, or a structured light to find the 3D information.

The implementation of our method (Figure 1) involves four principal phases :

Camera calibration : Consists in estimating the cameras parameters.

Interest Points Detection : is a preliminary step in many computer vision processes, many methods have been proposed to extract points of interest. In this paper, we combined between Harris [9] and Sift [10] interest point detectors.

Matching : Finding in two images of the same scene, taken at different positions, pairs of pixels which are the projections of the same point of the scene. In this phase, the detected interest points are matched by ZNCC (Zero mean Normalised Cross Correlation) correlation measure [11].

To eliminate false matches, we used the global constraint given by the fundamental matrix F.

3D reconstruction of matched points: is to estimate the 3D coordinates from point matches and projection matrices estimated.

3D Delaunay triangulation : 3D mesh is constructed by 3D Delaunay triangulation, applied to the 3D points reconstructed. It is a triangulation that satisfies the Delaunay criterion (empty sphere)

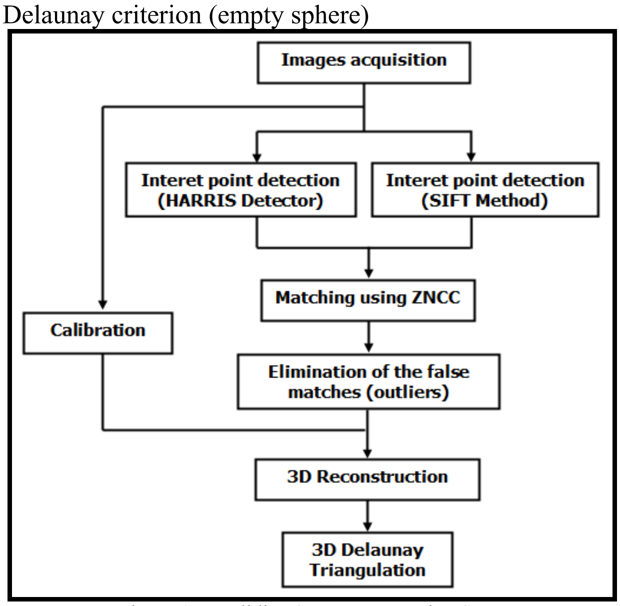

Figure 1 : Euclidien 3D Reconstruction Steps

II Camera Model and Calibration

2.1 Pinhole Camera Model

The pinhole model (Figure 2) consists of the image plane and the optical center O. A point $M_i = (X_i, Y_i, Z_i)^T$ of the 3D scene is projected onto the image plane at a $m_i = (u_i, v_i)^T$ point. This perspective projection is represented by the following formula :

© 2014 ACADEMY PUBLISHER
doi:10.4304/jetwi.6.1.59-63

$$\begin{pmatrix} \lambda u_i \\ \lambda v_i \\ \lambda \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix} \qquad (1)$$

λ : is a factor of homogeneity

$P = (p_{ij})_{i=1...3}^{j=1...4}$: is the perspective projection matrix

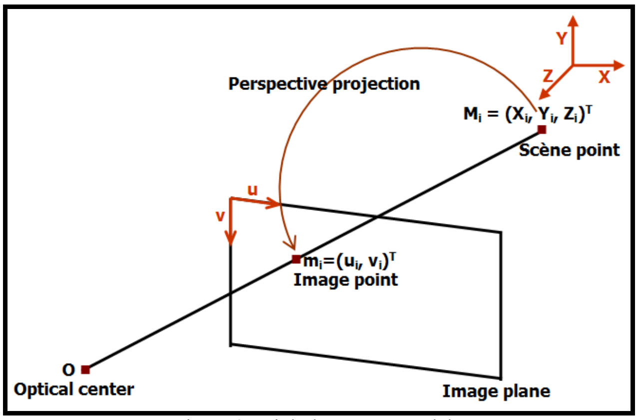

Figure 2 : Pinhole camera model

2.2 Calibration

The camera calibration consists in estimating the perspective projection matrix P that contains the parameters of the model.

3D Calibration Pattern (Figure 3) was used to estimate this matrix.

(a) (b)

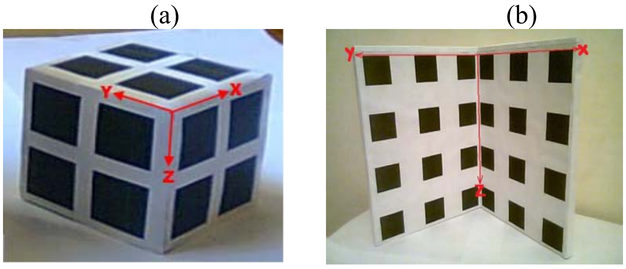

Figure 3 : 3D Calibration Pattern

From equation (1), we deduce :

$$\begin{aligned} X_i p_{11} + Y_i p_{12} + Z_i p_{13} + p_{14} - u_i X_i p_{31} - u_i Y_i p_{32} - u_i Z_i p_{33} &= u_i p_{34} \\ X_i p_{21} + Y_i p_{22} + Z_i p_{23} + p_{24} - v_i X_i p_{31} - v_i Y_i p_{32} - v_i Z_i p_{33} &= v_i p_{34} \end{aligned} \qquad (2)$$

Each point M_i of our 3D Calibration Pattern with known coordinates (X_i, Y_i, Z_i) is projected onto the image plane at a coordinate point (u_i, v_i), provides the equations (2). These equations are linear over the coefficients of P. Therefore at least 6 non-coplanar points are needed to determine P.

III Detection and Matching

3.1 Interest Points Detection

There are many methods [13] of detection point, but they do not have the same performances. In this paper, we combined between Harris [9] and Sift [10] interest point detectors. Indeed, Sift is considered as one of the best performing detectors because of its robustness to scale, rotation, translation and lighting changes. Harris Detector can find points on objects, specifically near the corner.

3.1.1 Sift keypoints

Interest Points detection takes place in two steps : (1) Scale Space Extrema Detection, (2) Keypoint Localization.

- **Scale Space Extrema Detection**

The detection is done in a space called **Scale Space** that has three dimensions : x, y and σ.

The Scale Space of an image I(x, y) is defined by:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \qquad (3)$$

Where, ***G*** is the Gaussian function.

To find the extrema, we use the function **DoG** (Difference of Gaussians) defined by :

$$\begin{aligned} D(x, y, \sigma) &= L(x, y, k\sigma) - L(x, y, \sigma) \\ D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \end{aligned} \qquad (4)$$

The extrema are pixels which present a maximum intensity, or minimum, compared with their immediate neighbours in the image as well as those in the space scale (26 neighbors).

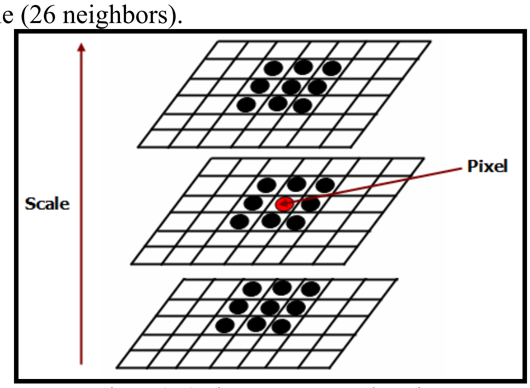

Figure 4 : Scale-space extrema detection

- **Keypoint Localization.**

In this step, key point candidates are localized and low contrast key points or those located on the image edge are removed.

3.1.2 Harris Corner Points

The detection is based on a three step process :

- Calculation of measurement of Harris R(x,y) (5) in each point of the image
- Thresholding to keep the larger values.
- Determination of local maxima in the values of R remaining.

$$R(x, y) = \det(M) - k \cdot [trace(M)]^2 \qquad (5)$$

With $k = 0.04$

M : is the matrix of autocorrlation

$$M = G(\sigma) * \begin{bmatrix} (I_x)^2 & I_x I_y \\ I_x I_y & (I_y)^2 \end{bmatrix} \qquad (6)$$

G : is the Gaussian filter

I(x, y) : is the image intensity

I_x *:* is the first derivative at x

I_y *:* is the first derivative at y

- *R < 0* : at the vicinity of an edge
- *R = 0* : in a homogeneous region
- *R> seuil* : near a point of interest

© 2014 ACADEMY PUBLISHER

3.2 Matching

The Matching of points is an important step in 3D reconstruction from images. It is to find for each point of the left image a corresponding in the right image.

In this paper, ZNCC [11] (Zero-mean Normalized Cross-Correlation) measure is used for matching interest points between different images.

The ZNCC value for a window size $(2N+1)\times(2P+1)$ is defined by :

$$ZNCC\ (m_1(u_1,v_1),m_2(u_2,v_2))=\frac{A}{\sqrt{B\times C}} \quad (7)$$

With :

$$A=\sum_{i=-N}^{N}\sum_{j=-P}^{P}(I_1(u_1+i,v_1+j)-\overline{I_1(u_1,v_1)})\times(I_2(u_2+i,v_2+j)-\overline{I_2(u_2,v_2)})$$

$$B=\sum_{i=-N}^{N}\sum_{j=-P}^{P}(I_1(u_1+i,v_1+j)-\overline{I_1(u_1,v_1)})^2$$

$$C=\sum_{i=-N}^{N}\sum_{j=-P}^{P}(I_2(u_2+i,v_2+j)-\overline{I_2(u_2,v_2)})^2$$

$$\overline{I_1(u_1,v_1)}=\frac{1}{(2N+1)(2P+1)}\sum_{i=-N}^{N}\sum_{j=-P}^{P}I_1(u_1+i,v_1+j)$$

$$\overline{I_2(u_2,v_2)}=\frac{1}{(2N+1)(2P+1)}\sum_{i=-N}^{N}\sum_{j=-P}^{P}I_2(u_2+i,v_2+j)$$

The value of the correlation ZNCC varies between -1 and 1. Consider a point in the image1 (left image), and its correlation is calculated with all points of the search box in the image2 (right image).

One retains only the maximum correlation point and above a threshold.

3.3 Removal of the False Matches

Not all extracted matches are accurate, can be of false matches (outliers). To avoid these matches, we used the global constraint given by the fundamental matrix F (matrice 3x3).

The RANSAC algorithm [12] was used to estimate this matrix and verify if the matches are correct.

This estimate is based on the equation :

$$m_2^T F m_1 = 0 \quad (8)$$

With (m1, m2) is a pair of corresponding points.

IV EUCLIDEAN 3D RECONSTRUCTION

4.1 3D Reconstruction of Matched Points

3D reconstruction is to estimate, from the matches already made and projection matrices defined during the phase of calibration P and P', the coordinates (X, Y, Z) of a point M of the scene.

By applying the projection equation (1) for a couple of matches (m1, m2), we write:

$$m_1 = P\ M$$

$$m_2 = P'M$$

By developing these equations, we find:

$$\begin{pmatrix} p_{11}-p_{31}u_1 & p_{12}-p_{32}u_1 & p_{13}-p_{33}u_1 & p_{14}-p_{34}u_1 \\ p_{21}-p_{31}v_1 & p_{22}-p_{32}v_1 & p_{23}-p_{33}v_1 & p_{24}-p_{34}v_1 \\ p'_{11}-p'_{31}u_2 & p'_{12}-p'_{32}u_2 & p'_{13}-p'_{33}u_2 & p'_{14}-p'_{34}u_2 \\ p'_{21}-p'_{31}v_2 & p'_{22}-p'_{32}v_2 & p'_{23}-p'_{33}v_2 & p'_{24}-p'_{34}v_2 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}=0 \quad (9)$$

The calculation of the coordinates (X, Y, Z) is obtained by solving the linear system (9) using the linear least squares sense. It is also possible to solve this system by a singular value Decomposition (SVD).

4.2 3D Delaunay Triangulation

3D mesh is constructed by 3D Delaunay triangulation (Figur 5), applied to the 3D points reconstructed. It is a triangulation that satisfies the Delaunay criterion (empty sphere)

- **The Delaunay criterion**

Delaunay criterion is satisfied for a couple of tetrahedra if no node of an element (tetrahedron) is contained in the Interior of the sphere circumscribing the other element.

(a) (b)

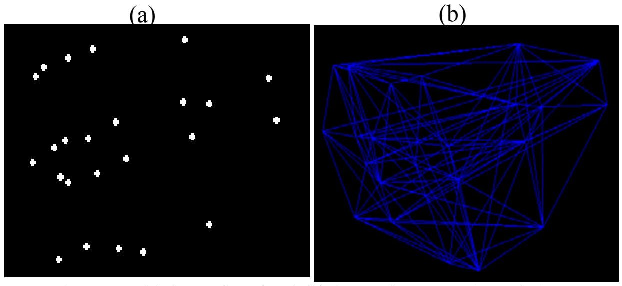

Figure 5 : (a) 3D point cloud (b) 3D Delaunay Triangulation

V EXPERIMENTS AND RESULTS

The presented algorithm was implemented in Java, using both JAMA library and The Java 3D API.

5.1 Images Acquisition

A digital camera was used for these experiments. The resolution used for the images is 640 × 480.

Image 1 Image 2

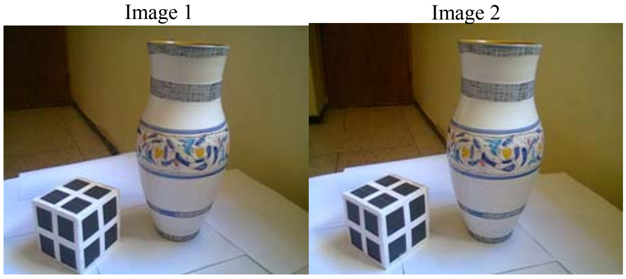

Figure 6. Images used for 3D reconstruction

5.2 Camera Calibration

In each teken view, our camera is calibrated (variable parameters). The coordinates of 10 well distributed corners in the 3D pattern are used.

© 2014 ACADEMY PUBLISHER

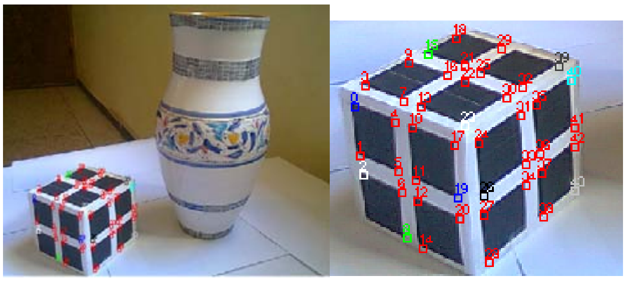

Figure 7. Detection of corners of our Calibration Pattern (Harris detector)

The first projection matrix estimated :

	14,8664	-10,74591,8653		153,8133
P1 =	-0,5906	-2,2124	16,2515	371,4603
	0,0187	0,0061	0,0060	1,0000

The second projection matrix estimated:

	14,6487	-8,6591	1,5661	146,6990
P2 =	-1,4943	-2,4432	14,7060	362,3744
	0,0136	0,0062	0,0038	1,0000

5.3 Matching

The ZNCC method was used to make the matching of points of interest detected by the Harris detector and the Sift method. The matches extracts are not all accurate, to remove false matches, we used the RANSAC algorithm [12].

Figure 8. Matching of points of interest (N = 7, P = 7 and thresholdZNCC = 0.85)

5.4 Euclidean 3D Reconstruction

We have the projection matrices and a set of matches, we can pass to the 3D reconstruction (we used a Singular Value Decomposition to solve the linear system (9)) :

5.4.1 3D Reconstruction of Matched Points

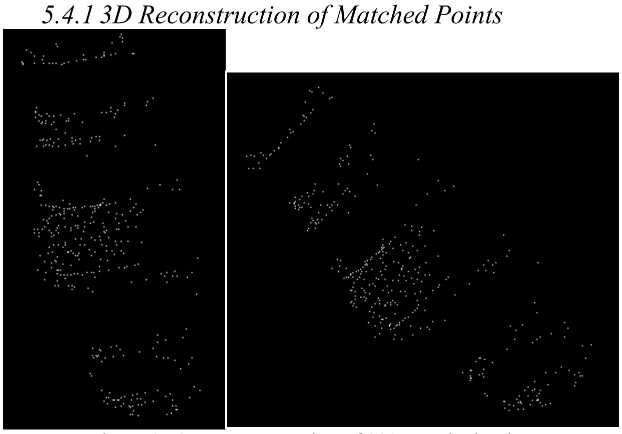

Figure 9. 3D Reconstruction of 389 Matched Points

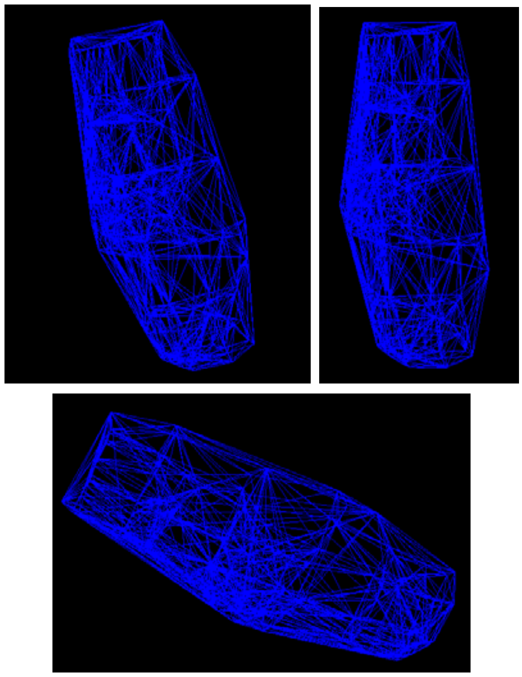

Figure 10. 3D mesh constructed by 3D Delaunay triangulation, applied to the 3D points reconstructed

5.4.2 *3D Reconstruction of Matched Points After propagation method* [14]

The propagation method is based on a set of reliable matches(germs). The germ with the best score ZNCC is removed from the current list of matches(germs), the new matches are searched in its neighborhood.

This method was applied to 167 germs for having 9622 matches.

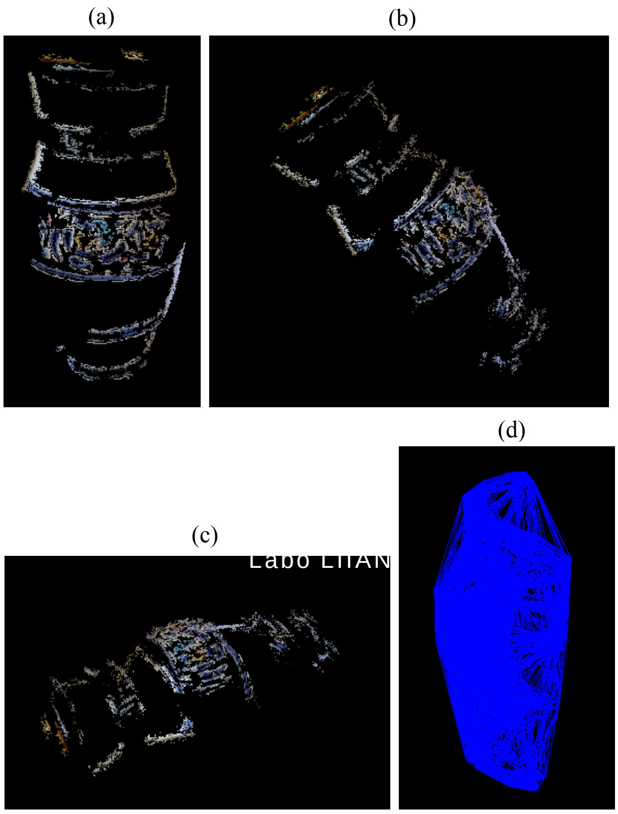

Figure 11. (a), (b) and (c) 3D Reconstruction of 9622 Matched Points (d) 3D Delaunay triangulation of 3D points reconstructed

© 2014 ACADEMY PUBLISHER

5.4.3 COMPARAISON

The results are presented in the table below :

TABLE 1.
RESULTS OF 3D RECONSTRUCTION

	Proposed method	**Propagation method [14]**
Number of reconstructed points	389	9622
Time of calculations	15 s	3 min
Reprojection error (pixel)	0.96	1.25

Based on the results of experiments, we can deduce :

TABLE 2.
COMPARAISON

	Proposed method	**Propagation method [14]**
Advantages	Rapid	The number of reconstructed points is large enough
	Quality of the resulting 3D model is accepted	Quality of the resulting 3D model is accepted
Disadvantages	Non-operational for non-textured objects	Non-operational for non-textured objects
		Slow due to the large number of reconstructed points

- **The reprojection error is defined by:**

$$Err = \frac{1}{2n}\sum_{i=1}^{n}\left(d(P_1 M_i, m_{1i})^2 + d(P_2 M_i, m_{2i})^2\right) \qquad (10)$$

n : is the number of matches (Number of 3D points).
P_1 et P_2 : are the projection matrices.
(m_{1i}, m_{2i}) : is a pair of corresponding points.
M_i : is the 3D reconstructed point from (m_{1i}, m_{2i}).

CONCLUSION

The proposed 3D reconstruction method enables to achieve satisfactory results in a short time. The combination between Harris and Sift detectors allows to have a sufficient number of matches without passed to the propagation step, What has been very useful to have a 3D model of quality in short time.

REFERENCES

[1] N.Salman, M.Yvinec, Surface Reconstruction from Multi-View Stereo of Large-Scale Outdoor Scenes. The International Journal of Virtual Reality, 2010.
[2] M.Han, T.Kanade. Multiple Motion Scene Reconstruction with Uncalibrated Cameras. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 7, 2003.
[3] S. Christy and R. Horaud. Euclidean Shape and Motion from Multiple Perspective Views by Affine Iterations. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 18, NO. 11, 1996.
[4] K.Kolev, T.Brox, D.Cremers. Fast Joint Estimation of Silhouettes and Dense 3D Geometry from Multiple Images. IEEE Transactions On Pattern Analysis And Machine Intelligence, 2012.
[5] J.D.Durou, V.Charvillat, M.Daramy and P.Gurdjos. Résolution du shape from shading par apprentissage. ORASIS - Congrès des jeunes chercheurs en vision par ordinateur, Praz-sur-Arly : France (2011)
[6] M.B.Adm, A.M.Said. 3D Reconstruction using Interactive Shape from Shading. International Journal of Computer Applications Volume 28– No.1, August 2011
[7] T.Collins, J.D.DUROU, Pierre GURDJOS and Adrien BARTOLI. Single-view Perspective Shape-from-Texture with Focal Length Estimation: A Piecewise Affine Approach. 3DPVT'10, France, May 2010.
[8] A.Dipanda, S.Woo, F.Marzani, J.M. Bilbault, 3-D Shape Reconstruction In An Active Stereo Vision System Using Genetic Algorithms, Pattern Recognition, Vol.36, pp 2143-2159, 2003.
[9] C.Harris et M.Stephens. A combined Corner and Edge Detector. Alvey vision Conference. pp. 147-151, 1988.
[10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2):91–110, 2004.
[11] S.Elhazzat, A.Saaidi et K.Satori. Reconstruction tridimensionnelle semi-automatique. WoISA'13, 2013.
[12] M. A. Fischler and R. C. Bolles. Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM, 24(6) :381–395, 1981.
[13] N.El akkad, A.Baataoui, A.El abderrahmani, A.Saaidi et K.Satori. Etude comparative des détecteurs des points d'intérêt. WCCCS, 2011.
[14] M. Lhuillier et L.Quan. Appariement dense robuste à l'aide de contraintes géométrique locales et globales. RFIA, 2000.

© 2014 ACADEMY PUBLISHER