

Study and Realization of the Basic Methods of the Calibration in Stereopsis For Augmented Reality

¹Fezani Mustapha, ²Batouche Chaouki, ³Benhocine Abdelhamid
¹LAIC, IUT, Department of Information, B.P. 86, 63172 AUBIERE cedex
Clermont Ferrand University, France
²Computer Department, Mentouri University of Constantine
³Department of Mathematics Ferhat Abbas University of Setif

Abstract: Augmented reality (RA) is aimed at ameliorating our perception of the real world by addition of elements which are not a priori observable by the human eye. RA is define as a system able of combining real and virtual pictures, in 3D and real-time. The virtual objects must be cast in a consistent manner in real pictures. In practice, objects are positioned in a landmark linked to the scene, and objective is to determine the point of view of the camera (position and orientation) in this landmark, which we are going to see in the next scene of calibration of camera. Camera calibration is a fundamental problem in computer vision. In this paper, we developed a semi automatic calibration method for our augmented reality system, which uses corner detection to extract pixel coordinates of projection points and uses homography to build the correspondence. So my method allows reconstruction 3D of the scene. To find resolution they use projections instead of the Euclidean methods which allow acquiring a good result with average error between the re-projected points and pixel points is round 2 pixels. Potential applications of this method are: applications real-time in the domains of medicine, service of fabricated objects, industry.

Keywords: appariement 3D-2D, calibration, geometry epipolar, augmented reality, stereopsis, homography, corner detection.

INTRODUCTION

RA is a discipline the majority of which of the advances took place these last ten years. System of composition of picture in scene of post production is used today as well by the producers of special effects as by the general public, who can obtain several software of trade (the degree of interactivity and the complicity of the taken into account applications being variable from a product to the other one).The camera is the essential tool around which develops the vision in RA.

The virtual objects must be cast in a consistent manner in real pictures. In practice, objects are positioned in a landmark linked to the stage, and objective is to determine the point of view of the camera (position and orientation) in this landmark
The projection model most fluently used is the model of perspective projection, says model pinhole which drives to a geometric model. It links the forming of pictures on the retina of the camera to a central projection ^[4].

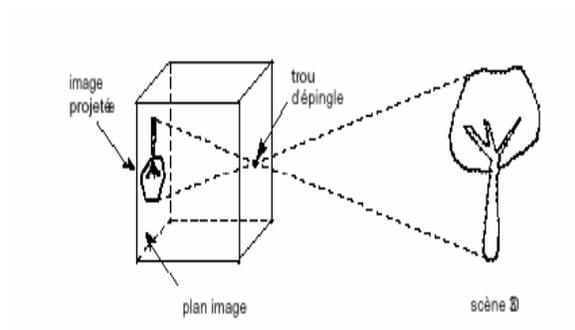


Fig. 1: The hole of pin.

In the following they are going to represent important stages which we have study and accomplish.
To cross coordinates defined in the landmark of the scene in coordinates pictures expressed in pixels (Fig. 2, three scenes are necessary ^[5]).

MATERIALS AND METHODS

All these pictures are dimension (256*256) pixels. Programs are developed under Matlab on a working station of type PC, Pentium IV.

The methods which we accomplished in this paper are:
 Methods of calibration of both cameras,
 Methods of estimate of the projection matrix,
 Methods of estimate of the fundamental matrix.

And for it they are going to introduce following notions:

A three-dimensional displacement: Points 3D expressed in a landmark of the scene are subjected to a change of landmark to pass to the landmark of the camera. This change of landmark includes six parameters therefore: three for rotation and three for translation. These parameters are others than position and orientation of the camera, they are called "*extrinsic parameters*".

$$D = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (1)$$

A projection 3D-2D: The three-dimensional points expressed in the landmark of the camera are cast on plan picture.

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (2)$$

Change of coordinates: To cross pixels to coordinates, coordinates are subjected to a familiar transformation of plan. This last, includes four parameters called "*intrinsic parameters*", they represent the internal characteristics of the camera.

This transformation can be spelled under following matrix form:

$$A = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

Where: (u_0, v_0) : represent the position of the centre of picture which is the result of the intersection of the optic axle and plan picture.

(α_u, α_v) : the real size of the following pixel horizontal and vertical direction respectively.

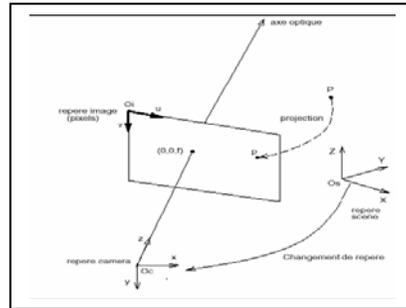


Fig. 2: Modèle géométrique.

The three named transformations c_i over can be encoded in the same matrix M (3x4) such as:

$$M = AID \quad (4)$$

Where:

D : represent three-dimensional displacement.

I : represent projection 3D-2D.

A : represent the change of coordinates.

The matrix M is said matrix of perspective projection and represents the model pinhole of the camera. The stocks of this matrix being unknown, the operation which consists in determining them is called calibration of the camera.

Calibration : The calibration is a technology allowing determining the parameters of the camera, which are two types: intrinsic and extrinsic. Intrinsic parameters model the internal characteristics of the camera, extrinsic parameters represent position and rotation of the camera in relation to the coordinates of the real world, they allow defining the relation which links the landmark picture in the landmark camera [6].

The determination of the parameters of the camera is based on the use of two techniques, (who can be combined): based techniques sensors (active vision) and based techniques vision (passive vision). Based techniques sensors require the job of appropriate equipment. They allow acquiring directly position and orientation of the camera every other minute. However, these techniques can be used only in environments of restrained size and they are besides, sensitive to the presence of materials unsettling the good functioning of the sensor (for instance of iron in the case of magnetic sensors) [7, 8]. Based techniques vision use pictures captured by the system of acquisition to find the parameters of the camera, the interest of these techniques is to require no instrumentation of the scene. Certain systems can use a hybrid approach: for instance, insert at the same time magnetic sensors and

markers positioned in the scene. This allows them to acquire the precision linked to the use of markers at the same time as the robustness given by sensors.

Our interest being carried on based techniques vision, these include two methods. The first one, said strong calibration, tries to make the research of a test pattern 3D (Fig.3.a) or of a pattern 2D (Fig.3.b) having characteristics known in the real scene.

This research can be made in a automatic or manual way. The second method tries to estimate the matrix of perspective projection by extraction of natural indications by using statistical methods from two or several pictures. It is said auto calibration or calibration weak-willed person [9].

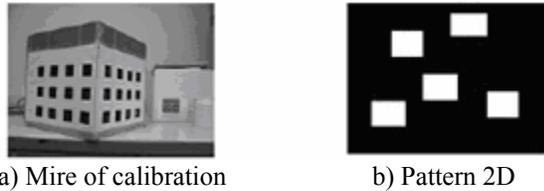


Fig. 3 : Mire and pattern calibration

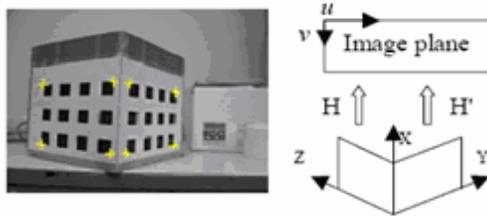


Fig. 4 : Corner used

Fig. 5: Homography

Weak calibration: The weak calibration (called also based calibration models) rests points of reference P_i on the knowledge of three-dimensional coordinates 3D of n (at least six points in the case of a test pattern and at least 4 points in the case of a pattern) and of their projections p_i in retinal plan, measured in form of coordinates pixel (for instance, the use of the detector of corners " Harris "). From this correspondence 3D-2D, we shall see that it is possible to calculate the stocks of the screening matrix what will later allow to deduct extrinsic parameters dead intrinsic parameters [7].

Appariement 3D-2D: To match coordinates 3D with their correspondents 2D, requires the appariement of coordinates 3D of corners of the windows of the test pattern with their correspondents 2D shown by HARRIS. In [1], Zonglei Huang and Boubakeur Boufama offer a half automatic method allowing finding appariements 3D-2D. Two homographies are

calculated (Fig.5), each of her represents the transformation of points 3D of the scene towards a point picture.

Equations below represent transformation:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (5)$$

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} h'_{11} & h'_{12} & h'_{13} \\ h'_{21} & h'_{22} & h'_{23} \\ h'_{31} & h'_{32} & h'_{33} \end{pmatrix} \begin{pmatrix} x \\ z \\ 1 \end{pmatrix}$$

Where: (u, v) : coordinated pictures. (x, y) and (x, z) : coordinates of points in plan XY, XZ respectively.

The determination of the stocks of homographies 2 requires the knowledge of appariements 4 for every homographie. To solve the system below they used method SVD (Singular Value Decomposition):

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -u_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -v_1x_1 & -v_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -u_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -v_2x_2 & -v_2y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -u_3y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -v_3x_3 & -v_3y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -u_4x_4 & -u_4y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -v_4x_4 & -v_4y_4 \end{pmatrix} \begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ u_4 \\ v_4 \end{pmatrix}$$

$$\begin{pmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -u_1x_1 & -u_1z_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -v_1x_1 & -v_1z_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -u_2x_2 & -u_2z_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -v_2x_2 & -v_2z_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -u_3x_3 & -u_3z_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -v_3x_3 & -v_3z_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -u_4x_4 & -u_4z_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -v_4x_4 & -v_4z_4 \end{pmatrix} \begin{pmatrix} h'_{11} \\ h'_{12} \\ h'_{13} \\ h'_{21} \\ h'_{22} \\ h'_{23} \\ h'_{31} \\ h'_{32} \end{pmatrix} = \begin{pmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ u_4 \\ v_4 \end{pmatrix}$$

Points picture calculated with the aid of this two homographies are used as a reference (an error $\delta 2$ pixels) for points discerned by an automatic detector of corners (they used HARRIS). The counting of error is established according to following expression:

$$E = \sum_{i=1}^N (\|x_i - u_i\|^2 + \|y_i - v_i\|^2) \quad (17)$$

Expression allows to calculate distance (error) between points discerned by HARRIS and re-projection of coordinates 3D.

The projection matrix: The relation between a point of the scene $P(X, Y, Z, 1)$ and its projection $p(u, v, 1)$ is given by following relation:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \lambda \text{AID} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = M \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (8)$$

λ : is the scale factor.

Equation (8) can be rewritten as follows:

$$\begin{aligned} u &= \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \\ v &= \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \end{aligned} \quad (9)$$

Where:

m_{ij} is the elements of the matrix M. To search the elements of the matrix M, six points are necessary to solve the system of equation below, by using the method of decomposition in peculiar stocks (SVD).

$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & -u_1 X_1 & -u_1 Y_1 & -u_1 Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & -v_1 X_1 & -v_1 Y_1 & -v_1 Z_1 \\ \dots & \dots \\ X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & -u_i X_i & -u_i Y_i & -u_i Z_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & -v_i X_i & -v_i Y_i & -v_i Z_i \\ \dots & \dots \\ m_{31} & m_{32} & m_{33} & m_{34} & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} = \begin{pmatrix} \dots \\ u_i, m_{34} \\ \dots \\ v_i, m_{34} \\ \dots \end{pmatrix}$$

STEREOPSIS VISION: Stereopsis is an important method to extract structure 3D from a scene. It consists in deducing the relief by looking at a scene with two cameras, disposed as the human eyes, which are going to give two pictures (stereoscopic system).

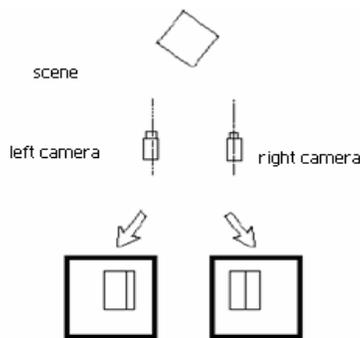


Fig 6. implement stereoscopy

Due to the distance of cameras, pictures of this pair are not without relation. So an object of the scene will see its pictures brought forward by a picture of pair on other one, by a certain number of pixels. This gap is called "difference" [10]. If they are able of allocating a difference in every pixel of picture, they will allocate a difference by extension in all points of an object on

picture. They will be then capable of replacing all points of this object in the space, therefore rebuilding the object in the scene. In effect, the distance of the object in cameras is conversely proportional to the difference of pixels picture which constitute it. All difficulty domiciles in the appariement of both pictures of pair (conscript: stereoscopic appariement). They imagine then all difficulty of problem when they think that an everyday scene can contain a big number of objects, put at various distance. The perception of this type of things at the human beings is so natural that it is very difficult to arrest functioning [10].

Geometry epipolar: This technology is founded on algebraically ownership of projective geometry. It rests on the counting of the « fundamental matrix » which imposes geometric pressures between two views, said « compelled ear polar ». The determination of this matrix requires a minimum of eight correspondences of points. In the field of the stéréovision, geometry polar ear, illustrated in the Face 8, provides following important result: Given a point p_1 in picture 1, its corresponding stereo p_2 belongs necessarily to a right of the picture 2 entirely defined by data of p_1 . This right is called *right polar ear* linked in p_1 .

Thanks to this important geometric ownership, the research of the corresponding stereo of a point of picture 1 comes down to a research 1D in picture 2 (along right epipolar linked) rather than a research 2D (exhaustive research in all picture 2).

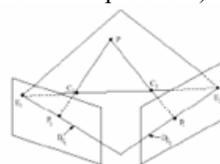


Fig 7. Epipolar geometry

Right ear polar of picture 2 form a beam of rights of centre E_2 , the picture of C_1 in the camera 2. E_2 is called the ear pole of picture 2. Situation is symmetrical and E_1 , the picture of C_2 in the camera 1, the ear is pole of picture 1. They can also consider E_1 and E_2 to be the intersections of right C_1C_2 with plans pictures P_1 and P_2 respectively.

The fundamental matrix: The fundamental matrix is the algebraically presentation of geometry polar ear. His main characteristics are translated in a pair of pictures: data of a point in the one right corresponds to the data of a right in other one, this duality correspond to the projection of the screening ray of the point. The

fundamental matrix satisfies condition that for every pair (q, q') of points matched by the stereoscopic system, they have relation:

$$q^t F q' = 0 \quad (10)$$

This translates the fact that the point q belongs to right (Fr q') and reciprocally. One adopting the following writing:

$$F = f_{ij} \text{ and } q = (u \ v \ 1)$$

The previous equation writes:

$$u' u f_{11} + u' v f_{12} + u' f_{13} + v' u f_{21} + v' v f_{22} + v' f_{23} + u f_{31} + v f_{32} + f_{33} = 0 \quad (12)$$

What allows having the system below?

$$\begin{pmatrix} u_1' u_1 & u_1' v_1 & u_1' & v_1' u_1 & v_1' v_1 & v_1' & u_1 & v_1 & 1 \\ \dots & \dots \\ u_n' u_n & u_n' v_n & u_n' & v_n' u_n & v_n' v_n & v_n' & u_n & v_n & 1 \end{pmatrix} f = Af = 0$$

The resolution of this system is accomplished by the application of SVD.

Correction: There is a particular case (in geometry epipolar) interesting which corresponds to parallel polar beams of ear in both pictures at the same time: in that case, the research of the corresponding pixels is greatly simplified. This occurs only if right C1C2 is parallel on plans pictures of both cameras. In practice, this special situation is very difficult to accomplish mechanically, but the operation of correction of pictures allows precisely coming down to it by a transformation of a posteriori pictures. To note that the scene of correction introduces distortions of pictures [11].

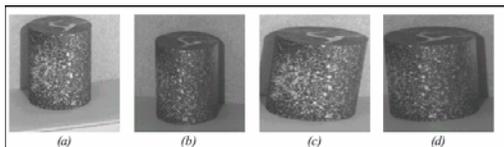


Fig. 8 : (a) left Image (b) right Image, (c) left image rectified (d) right Image rectified

Appariement by correlation of a pair of stereoscopic pictures: The scene of correction guarantees that the pixels which correspond in both pictures are on the

same line. The scene of correlation consists in determining the couples of corresponding pixels. To determine the correspondent of a pixel of the first picture in second, it is not enough of course to compare simply the intensity of pixels two - two: they measure likeness between two pixels by calculating a score of correlation determined on their neighborhood (in practice, considered neighborhood is rectangular windows centered on the examined points). Given a pixel in the first picture and his linked neighborhood, his correspondent in the second picture is the one who maximizes (or minimizes score, it depends on the criterion of used correlation) the score of correlation along the line (Face 9).

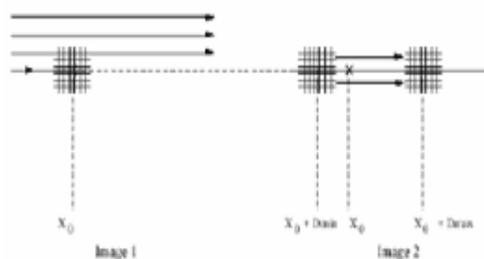


Fig 9: Scene of correlation

There are numerous criteria of correlation; the choice of criterion is often empirical [11].

RESULTS

Results acquired, given on the picture mentioned below present the stocks of the eight matrices of projection acquired with the aid of pictures presented in fig.10.



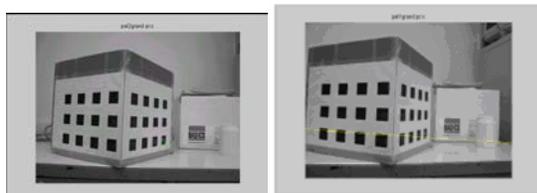
Fig. 10: Used pictures

Table 1: The projection matrices of used pictures

m_{ij}	1	2	3	4
1	0.4665	10.1725	-5.6292	213.1173
2	-10.366	4.0210	2.4108	346.6853
3	0.0003	0.0172	0.0125	1.0000
1	0.5990	9.4401	-7.7436	215.2268
2	-10.9095	4.5161	1.8683	354.2236
3	0.0008	0.0198	0.0103	1.0000
1	0.6993	13.5463	-4.4762	217.7523
2	-12.4950	4.3817	3.7306	376.6168
3	0.0008	0.0183	0.0184	1.0000
1	0.8468	15.8208	-1.6972	221.5089
2	-13.6594	4.0600	4.9324	392.8306
3	0.0009	0.0161	0.0235	1.0000
1	0.8992	16.0335	-4.1648	244.5548
2	-14.2423	4.8166	4.4899	399.4926
3	0.0013	0.0199	0.0219	1.0000
1	0.2025	12.2634	-3.8008	203.3318
2	-11.4630	3.4425	3.0748	390.6496
3	-0.0006	0.0161	0.0167	1.0000
1	0.2464	11.4637	-4.8848	205.8581
2	-11.1894	3.6064	2.7215	387.7937
3	-0.0004	0.0170	0.0150	1.0000
1	-0.164	11.4440	-5.4828	264.3412
2	-11.4670	3.3254	2.1697	362.0337
3	-0.0014	0.0182	0.0140	1.0000

Once the image coordinates of the corners are identified and the correspondences between image points and space points are established, a camera calibration is called to calculate the projection matrix. When the user clicks on the calibration button in the interface, the program performs the calibration using the data correspondences between. Table 1 contains the calibration resultat, the projection matrix (fig.11). The average error between the re-projected points and pixel points is round 2 pixels

Once they determined the stocks of both projection matrixes, they can estimate the fundamental matrix between both pictures. The fig.12.b. Show epipolar right on which the correspondent of the point is shown in fig.12.a.



a). The concerned point (b). right epipolar

Fig. 12 : Right epipolar

By using what was explained above, they could insert two objects as show it the face fig.13, while taking into account the problem of occlusion.

CONCLUSION

The epipolar geometry allows to match pictures of sequence two - two. In a system of augmented reality, the virtual objects must be cast in a consistent manner in real pictures. In practice, objects are positioned in a landmark linked to the scene, and objective is to determine the point of view of the camera (position and orientation) in this landmark. They can conclude also that besides these extrinsic parameters, we must know the intrinsic characteristics of the camera such as the focal length for instance, which define the projection of the virtual object in plan picture of the camera. Besides, to cause the position of the initial to discern in pictures of sequence, they necessarily need a specific tool. Our calibration method gives an error of about 2 pixels.

REFERENCES

1. Zonglei, H. and Boubakeur, B., 2001. A Semi Automatic Camera Calibration Method for Augmented Reality
2. Zonglei, H., 2002. Computer visionbased Registration for Augmented Reality". *A thesis for the degree of Master of science, Canada*
3. Peter Kovesi .School of Computer Science & Software Engineering. The University of Western Australia. pk at csse uwa edu au <http://www.csse.uwa.edu.au/~pk>
4. Telle, B., 2003. Méthode ensembliste pour une reconstruction tridimensionnelle garantie par Stéréovision. Thèse de doctorat
5. Boubakeur B. 1994. Reconstruction Tridimensionnelle en vision par ordinateur: cas des cameras non étalonnées .Ph D. Thesis, France
6. Zehang, S. and Christian, R.B. , 2001. Camera Calibration. Internship with UNR and Ford Motor company
7. Gilles, S., 1999. Vers un Système de réalité Augmentée Autonome. Thèse de doctorat
8. Gilles, S., Berger, M.O. , 1999. Des méthodes efficaces pour l'incrustation d'objets virtuels dans des séquences d'images. Revue traitement de signal
9. NINI Tuceryan, M., Greer, D. S., Whitaker, R. T., Breen,D. E., Crampton, C., Rose, E. & Ahlers, K. H.1995.Calibration requirements and procedures for a monitor-based augmented reality system ,IEEE Transactions on Visualization and Computer Graphics 1(3), 255–273.
10. Heikkila, J. 2000, 'Geometric camera calibration using circular control points', IEEE Transactionson Pattern Analysis and Machine Intelligence 22(10), 1066–1077.

11. Schmalstieg., D. Reitmayr, G., 2003. Location based applications for mobile augmented reality. In *Australasian User Interface Conference*. 1.4.2, 2.4.3
12. Fiorentino, M., de Amicis, R., Monno, G. & Stork, A.2002. Spacedesign: A mixed reality workspace for aesthetic industrial design, in 'International Symposium on Mixed and Augmented Reality (ISMAR'02)'.
13. Piekarski, W. & Thomas, B. H. 2003. Interactive augmented reality techniques for construction at a distance of 3d geometry, in 'Proceedings of International Immersive Projection Technology Workshop and Eurographics Workshop on Virtual Environments', pp. 461–468
14. Ryokai, K., Marti, S. & Ishii, H. 2004. I/o brush: Drawing with everyday objects as ink, in 'Proceedings of ACM CHI 2004', pp. 234–241
15. Raphael Grasset and Jean-Dominique Gascuel. , 2003. R'éalit'e augment'ee et environnement collaboratif : Un tour d'horizon. In *AFIG'03. Actes des 16emes journees de l'AFIG*. 3.1, 4
16. Dieter Schmalstieg Gerhard Reitmayr., 2003. Location based applications for mobile augmented reality. In *Australasian User Interface Conference*. 1.4.2, 2.4.3
17. M. Gross, S. Wuermlin, M. Naef, E. Lamboray, C. Spagno, A. Kunz, E. Koller-Meier, T. Svoboda, L. Van Gool, S. Lang, K. Strehlke, A. Vande Moere, and O. Staadt., 2003. Blue-c : A spatially immersive display and 3d video portal for telepresence. In *SIGGRAPH*. ACM, 2003. 2.1.3
18. Milan Ikits, Charles D. Hansen, and Christopher R. Johnson., 2003 A comprehensive calibration and registration procedure for the visual haptic worbench. In *Eurographics Workshop on Virtual Environments*, pages 247–254, Zurich, Switzerland, May 2003. 3.1, 3.2.1, 3.2.4, 3.2.5
19. Christian Kulas, Christian Sandor, and Gudrun Klinker. , 2004. Towards a development methodology for augmented reality user interfaces. In *International Workshop on Exploring the Design and Engineering of Mixed Reality Systems (MIXER)*. 3.5
20. A.J.K. Kok and Robert van Liere. , 2004. Colocation and tactile feedback for 2d widget manipulation. In *IEEE VR*, March 2004. 3.3.1
21. T.B. Moeslund, M. St'orrning, Y. Liu, W. Broll, I. Lindt, C. Yuan, and M., 2004. Wittkaemper. Towards natural, intuitive and non-intrusivehci devices for roundtable meetings. In *workshop on Multi-User and Ubiquitous User Interfaces (MU3I)*, Funchal, Portugal, 2004. 2.4.5