

Interactive 3D reconstruction for urban areas

An image based tool

C. Chevrier and J.P. Perrin

CRAI UMR MAP 694, School of Architecture of Nancy

Key words: geometrical modelling, architecture, urban area, virtual visit.

Abstract: Urban applications (for example arrangement, new buildings, virtual sightseeing and walkthrough) require a three dimensional (3D) geometrical model of town areas. However, most of them do not need an accurate model of reality. Such model would occupy a considerable memory space and would be too slow to handle. Architects, urban designers and civil engineers can find in our tool a medium to conceive their projects. Some types of software exist but they do not correspond exactly to our needs. Consequently we have conceived and developed an interactive tool for virtual 3D rough reconstruction of buildings. The software development has been performed in the Maya environment (ALIAS Wavefront) with C++ language and MEL (Maya Embedded Language). A constraint we set for ourselves was the use of only light devices (for easy transportation) at low price (everybody can buy such devices). The principle is to overlay on the scanned photograph of the area we want to deal with, the two dimensional (2D) cadastral plan displayed from the same viewpoint as the picture. Then each building body can be extruded from its ground polygon and the roof can be created from what the user sees on the picture. A constraint is the flatness of the polygonal surfaces. Our application context was the town of Nancy in France for which some areas have been reconstructed. Some pictures have been used as textures for polygonal surfaces, giving more reality effect to the simulation.

1. INTRODUCTION

Urban applications require a three dimensional (3D) geometrical model of town areas: urban arrangements, new buildings (augmented reality simulations with handling of the interactions between real and virtual worlds (Chevrier, 1996b), virtual sightseeing and walkthrough). However, most of

them do not need an accurate model of reality. Such a model would hold a great memory place and would be too slow to handle.

We have conceived and developed an interactive tool for virtual 3D rough reconstruction of buildings. The software development has been performed in the Maya environment (ALIAS Wavefront). Architects, urban designers and civil engineers can find in this tool a medium to conceive their projects.

The principle is to overlay on the scanned photograph of the area we want to deal with, the two dimensional (2D) cadastral plan displayed from the same viewpoint as the picture. Then each building body can be extruded from its ground polygon and the roof can be created from what the user sees on the picture. A constraint is the flatness of the polygonal surfaces.

We developed a set of commands making the creation of various roof shapes easier, taking advantage of the use of photographs. A constraint we also set for ourselves was the use of only light devices at low price: only a camera without any specific features (simpler if it is a digital camera) and if possible a hand-held distance meter are required. This lowers the cost of the devices and the time spent for the camera shots. We can use as input data, the results of the automatic rough reconstruction from Medina software (Allani and Perrin, 1998) developed by our team.

Part 2 presents the state of the art for 3D reconstruction and viewpoint recovery. The next Part (3) explains the principles of our method and the required input data. Then, we see in the following part (4) how to retrieve the photograph viewpoints. The reconstruction task can finally begin with the extrusion of the building bodies (part 5) and the building up of the roofs (part 6). Part 7 presents the results with the help of a concrete application. Finally, part 8 presents future work and concludes.

2. STATE OF THE ART

Medina (Allani and Perrin, 1998) is a program for automatic 3D simplified reconstruction from 2D cadastral plans. It utilises information stored in the plans (for instance the number of storeys in a Dxf format file), urban regulations and architectural laws to build up the global shape of the urban area being treated. The most appropriate roof is constructed for each building. This program allows one quick reconstruction of large areas. However, the results lack accuracy for some applications (augmented reality) and complex roofs are not correctly managed. Nevertheless, the resulting file from Medina can be used as an input for our module. Only the badly or non built-up houses are then dealt with.

Different types of commercial interactive software exist to solve the problem of 3D reconstruction. Let us examine some of them. Canoma (Canoma, 2000), ImageModeler (RealViz, 2000) and PhotoModeler (PhotoModeler, 2000) are software based on the principle of photograph aided modelling.

PhotoModeler is based on the principles of photogrammetry. Canoma does not require a viewpoint recovery of the digital images, but several images of the same object have to be used in order to fix the object position and shape correctly in the scene. On the contrary, ImageModeler and PhotoModeler require a viewpoint recovery. The calibration process enables the building in the scene of the 3D indices corresponding to the 2D indices selected on the pictures. Then, the user can rebuild the objects with the help of these 3D indices and the photographs. In the case of PhotoModeler, the 2D indices can be segments (not only points); one can automatically obtain a 3D model composed of segments. All these software have simple primitives such as plane, cube, cylinder, cone, and sphere, interactively positioned in the scene with manipulators. Main shapes for a rough reconstruction are more or less provided according to the software. However, no composition of simple roofs to create complex roofing is possible. An application example of Canoma is the reconstruction of parts of the town of Phoenix (Arizona) mainly composed of box-shaped buildings. Nothing ensures the flatness of the polygons after modification in most programs. In architecture, most of the roofs are composed of planar polygons: it is important to respect this constraint. These software are not based on modellers so they do not dispose of classical commands. They can render images by wrapping the photographs around the 3D primitives.

Automatic reconstruction with the help of image analysis process produces good results for indoor and simple scenes. For urban scenes, most of the research carried out on that subject uses aerial video sequences (Faugeras, Laveau, et al., 1995) (Collins, Hanson, et al., 1995). Photographs taken by a walking man are difficult to deal with: scenes are composed of lots of objects of various kinds (lots of "parasite" objects in front of the buildings).

We want to create an interactive tool for helping 3D rough reconstruction of urban areas. To simplify this task an interactive modeller is useful. Classical modellers (AutoCad, Arc+) allow neither intuitive and simple handling of the objects, nor real-time 3D visualisation. Maya contains several modules, among which such a modeller. The Open Inventor library (Wernecke, 1995) allows real time visualisation and handling with manipulators of 3D graphical objects. However, no standard modelling feature is provided. Thus, we choose to develop our prototype in the Maya environment with the C++ programming language and MEL (Maya

Embedded Language), avoiding the fastidious task of writing all the standard operations of a modeller. Nevertheless, we can note that Maya's modeller is not as complete and accurate as classical modellers but it fits our needs.

Furthermore, Maya allows us to overlay a 3D scene on a background image. This image is associated to a given camera viewpoint. Several cameras and also several images can be used at the same time for our reconstruction goal.

As far as the viewpoint recovery is concerned, one can find more and more commercial software (realViz, 3Dstudio, Mayalive) and a large amount of research has been carried out on that subject (Devernay, Faugeras, 1995)(Berger, Chevrier, et al., 1996)(Simon, Berger, 1999). Some tests with MayaLive did not satisfy us for several reasons:

1) MayaLive deals with video sequences and not with just one picture. It obliged us to film the scene instead of taking a few pictures. We had to use a tripod and to borrow an expensive digital video camera. The room taken up by such devices was too great and they were not easy to move. The time spent on the shooting was also long.

2) The recovery of the viewpoint requires several steps (as usual for video sequences, whatever the software): 2D relevant indices tracking and resolution of the system. These steps are too long for the one picture that interests us in the sequence.

3) The first tests we made were in an indoor courtyard of the school of Architecture (a very linear new building). The results were very promising. Unfortunately, trials in an urban context were not satisfactory: we do not dispose of enough relevant indices in the images and in the 3D scene. Points in the cadastral plan were often not visible in the images because they were hidden by other objects (dustbins, cars, trees, low walls, etc).

ImageModeler and PhotoModeler utilise calibrated images for the 3D reconstruction. One can use pictures or movies. Camera calibration is processed as in MayaLive with the help of relevant indices (points in ImageModeler, points and segments in PhotoModeler) seen on at least three images. These points are manually pointed with a large cover in the images and in the three main directions of the space. Six points are required. 3D calculated points can be manually adjusted in case of errors and calibration is processed subsequently. In our applications, we mainly have at our disposal points on a 2D plan (cadastral plan), and hardly any indices in the vertical direction.

Thus we have developed a simple, interactive and sufficiently effective method for our needs in order to recover the viewpoint of a scanned photograph. This method was created in Maya, ensuring compatibility with the 3D reconstruction software, as in the same environment.

3. PRINCIPLES

3.1 Steps of the reconstruction

The **first** step consists, for each picture, in recovering the camera position in the scene in order to visualise the 2D plan according to the same viewpoint. The **second** step is the creation of the building bodies: ground polygon extrusion (special case of the non-planar and non-horizontal ground). Finally the **third** step allows us to construct the roofs: Some commands make easier the creation of simple roofs (1, 2, 3 or 4 slopes, pyramidal or flat) and composed roofs (association of several simple roofs).

3.2 Input data

Geometrical data about the existing scenes are necessary for viewpoint recovery: this step utilises the matching of the 3D data and their corresponding 2D data on the images. The data we use are:

- a cadastral plan with ground point elevations. If we do not dispose of such altitudes, we can interpolate them from the altitudes of reference points (geodesic points). Cadastral data have sometimes to be cleaned and treated to create closed ground polygons.
- Coordinates of several points at other altitudes. These points are useful in order to recover the viewpoint. It reduces the risk of error in the third space direction. These points can be measured with the help of a hand laser: distance between the laser and a target point (for example the corner of the buildings at the roof gutter level). One point per picture is enough. If we do not dispose of such points, a small uncertainty will subsist on the height of the reconstructed objects. This will be minimised with the use of several images.
- Photographs of the area: pictures of a street-level pedestrian and heightened views shot from high buildings in order to have a better vision of the roofs. Images are taken with a digital camera thus avoiding losing time with development and scanning (scanning produces troubles with pixel size and position of the optical centre projection).

4. VIEWPOINT RECOVERY

Recovering the viewpoint implies identifying two sets of parameters: the intrinsic and the extrinsic parameters. This means eleven parameters in totality (Horaud, Monga, 1993)(Toscani, 1987).

Intrinsic parameters are specific to the camera and are unvarying when it moves in the scene: focal length or aperture, pixel size, position of the optical centre. The use of a calibration grid (regular grid) picture allows a reliable identification of these parameters (Toscani, 1987).

Extrinsic parameters are the position of the camera in the scene, the target and the rolling angle (orientation of the camera according to the sight axis). These parameters are specific to each image. If the intrinsic parameters have been previously determined, one can compute the extrinsic parameters from a small set of corresponding 2D and 3D relevant indices (Horaud, Monga, 1993). However, estimating these parameters for noisy and complex images is not easy (Chevrier, 1996a).

Trials carried out with MayaLive did not satisfy us, so we have developed a simple interactive method that allows us both quick and good positioning of the camera in the scene.

4.1 Principle

- We utilise a first fixed point ***I1*** chosen by the user in the image. To this point corresponds a 3D point ***P1*** in the scene. The two points are superimposed: translation of the camera (see Figure 1a).
- A second semi-fixed point ***I2*** is chosen in the image, to which corresponds the 3D point ***P2***. These two points are also stacked keeping the first superposition: rotation ***x*** around the X axis, rotation ***y*** around the transform of Y axis by rotation ***x***, scaling ***e*** of the image plane (see Figure 1b).
- The image plane can be rotated according to the two fixed points in order to correctly orientate it according to the scene: rotation ***r*** (see Figure 1c).
- The point ***P2*** can move along the line (Camera position ***C***, point ***I2*** on the image): parameter ***p***. A modification of ***p*** leads to a modification of scaling factor ***e*** and the rotation angles ***x*** and ***y*** (see Figure 1d).
- Finally the focal length ***f*** of the camera can vary. A modification of ***f*** leads to variations on the points ***I1*** and ***I2*** as the image plane dimensions change.

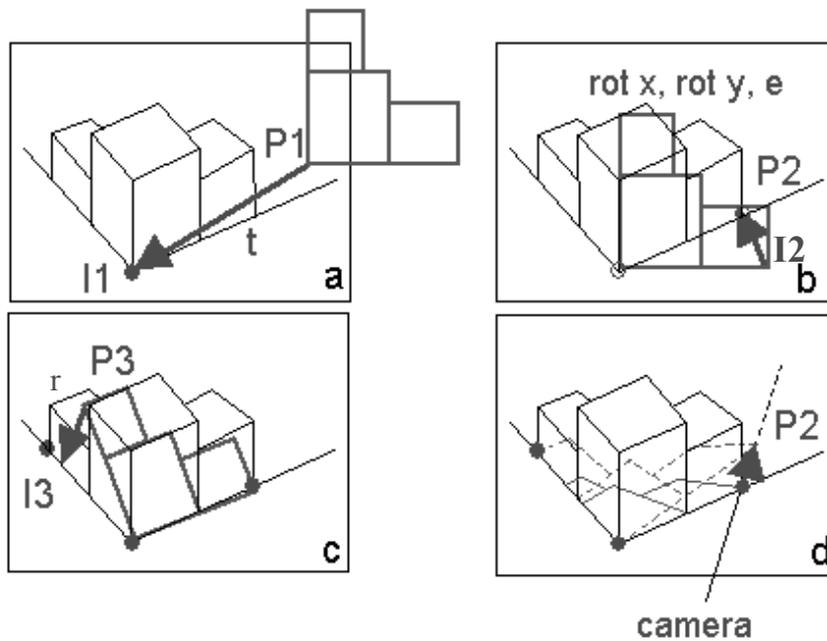


Figure 1. Estimation of the viewpoint

The first four parameters ($I1, I2, P1$ and $P2$) are fixed once at the beginning by the user. The other three parameters (p, r and f) must be adjusted together. From eleven unknown variables, we reduce that number to three. Furthermore, if the focal length is known, it can be fixed leaving us with only two unknown variables. In practice, even if the focal length has been fixed, we sometimes have to vary it a little to get better results. One can use the vanishing points to estimate the focal length value, and knowledge of the scene and the perspective of the image to approximately fix the starting value of parameter p . Finally, one can approach the value of the rotation angle r according to the high angle shot or low angle shot.

Note that as we do not have at our disposal a lot of high points, we create some vertical lines from relevant points of the cadastral plan. When several buildings have been reconstructed, these lines are not useful any more: the buildings themselves are used.

In some cases, one can use a third fixed point $I3$, corresponding to a 3D point $P3$. This point $P3$ must not be aligned with the two first points $P1$ and $P2$. This is not always possible, for instance when we have a picture of frontages in a street. Thanks to this third point, the rotation r can be

estimated from the value of the parameter p given by the user. With the rotation r , the projection of $P3$ on the image plane is on the line $(I1, I3)$ (see Figure 1c). With a variation of the parameter p , the projection of $P3$ comes on the point $I3$. This greatly simplifies the process. Only the parameter p needs to be estimated and the focal f sometimes needs to be adjusted.

4.2 Distorted images

Main distortions, which means radial distortions, are identical for all the points along a primary ray (ray shot from the eye towards a given direction). In image based aided work (medicine for example), it is often important to have non-distorted images but not necessary to know the shooting parameters. In this case, distorting the image is sufficient. It means putting it in conformity with a perfect perspective projection. In order to do that, a method of comparison between a perfect image and a photograph of a specific object (regular grid) allows one to globally correct all the distortions with no need to model them (Peuchot, 1994). As this kind of distortion depends on the focal length, this correction has to be done for each change in the focal length.

We use this method to distort our images before viewpoint recovery. For small focal lengths (less than 50 mm), one can observe distortions in a small cushion, and for long focal lengths (more than 50 mm), distortions are in a small barrel. The grid (composed of squares of 5cm) has been shot with our digital camera (Olympus) for three various focal lengths : the smallest, the greatest and one in the middle. The wide angle corresponds to a focal length of 9.2 (equivalent to 36 mm for a standard 24x36 camera). The telephoto focal length is 28mm (110 mm for a standard camera). We have developed a program to distort the images according to the focal length used. Most of the shots were made with the widest angle.

5. BUILDING BODIES

The second step is the extrusion of the ground polygons to create the building bodies. If the ground is not horizontal, we must first position the cadastral plan according to the ground altitudes: we vertically project the polygons on the 3D mesh created with the known-altitude points. With Maya, the extrusion of a polygon is performed perpendicularly to that polygon, which does not correspond to what we expected. Thus we develop a command to extrude vertically non-horizontal polygons: a bottom part is automatically created to get the horizontal level before extruding the polygon (Figure 2).

The extrusion can be performed with various options: a given height if an in-situ measure has been taken, or the number of storeys and a height per storey (estimated from architectural customs and laws of the construction period). Then, a manipulator can be used to adjust the height according to the picture.

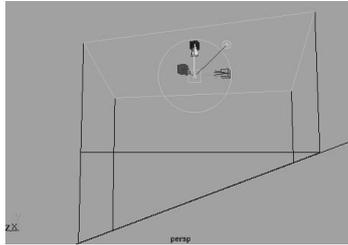


Figure 2. Polygon extrusion for a sloped ground

6. ROOF CONSTRUCTION

Thanks to the pictures, we are now able to create the roofs. We distinguish two types of roof: the simple roof and the composed roof. Simple roofs are pyramidal, flat, made of one, two, three or four slopes. Complex roofs are composed of several simple roofs with particular association rules: “A la Mansard” roofs, L, T or U shaped roofs.

A set of commands makes the roof creation easier taking into account the use of photographs. In classical modellers, roofs are created from the slope of the various sides. In our case, we do not know these data, so we have used other data: the position of the roof top. Two, three and four slope roofs are then created from the top.

6.1 Simple roofs

A flat roof is created by the raising of the top side points. The user can modify the position of the top points of the flat roof to create for example the bottom part of a “A la Mansard” roof. A simple roof (two or four slope roof) has to be constructed over that part to complete the roof.

Steeple can be modelled with pyramidal roofs. The height of the top point can be adjusted with the help of the picture. A one slope roof is created from three relevant points defining the slope plane. For this the user moves vertically one or more points of the top side of the building body.

In order to create a two, three or four slope roof, we first have to create a roof top (segment positioned thanks to its representation on various pictures). We can create a top parallel to a selected edge on the top polygon

of the building body. Otherwise, the top is arbitrarily placed in the centre of the top polygon. Then, the user interactively adjusts the roof top with the help of the background pictures and the point manipulator.

First case: the user selects the top and two points on the top polygon (one point per main slope). These points have to be at their correct place to define the corresponding slope plan. The other points on the top polygon will be vertically modified to create planar slopes. To create the sided slopes, we use the closest points to each extremity of the top roof. If other points have to be used, the second case has to be followed.

Second case: the user selects four correctly placed points on the top polygon to define the side slopes.

However, if he knows the roof slopes, the user can build the roof with these data with no help from the picture.

One can modify some points of the roof if necessary (adjustment according to other pictures), the various quadrilateral sides may not be planar any more. A command makes them planar by computing the planar shape closest to the original shape. One can notice the case of quasi-vertical polygons that can be modified to be vertical.

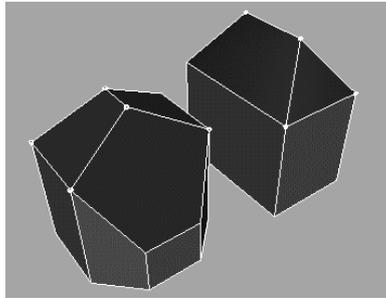


Figure 3. Examples of four slope roofs

3.2. Complex roofs

A L, U or T shaped roof is a concatenation of simple roofs according to assembling rules. In order to create a composed roof, we first have to cut the ground polygon of the building into several polygons: each part will be separately extruded and be covered with one or several simple roofs in the case of superimposing. Finally, we define the connection between two simple roofs (*Figure 4*).

Primary and secondary roof connection: the secondary roof is extended to lay on the primary roof. The top of the secondary roof was only approximately created towards the primary roof. Several cases have to be

distinguished according to the height of the top polygons and roof tops of the two parts.

Algorithm:

1. Stretch the roof top of the secondary roof until it intersects the primary roof (not possible to continue if there is no intersection)
2. **If** the top polygon of the building body of the secondary building is lower than the top polygon of the building body of the primary building then
Points have to be added to the secondary roof to fit the primary building (Figure 4a).
else
Points have to be added to the secondary body to fit the primary building (Figure 4b):
the two side edges are extended until they intersect the primary roof.

Stuck roofs along a mutual wall: the two building bodies need to have a mutual wall. The two simple roofs are then stuck together along this wall (Figure 4c).

Algorithm:

1. look for the common wall
2. stretch the two roof tops to this common wall
3. modify the two roofs according to the new tops.

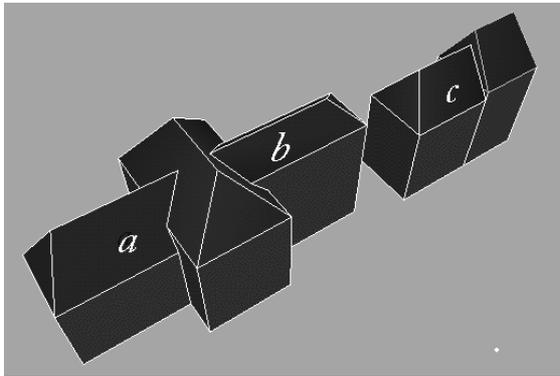


Figure 4. Composed roofs

7. RESULTS AND APPLICATIONS

Our application context was the city of Nancy in France for which some areas have been reconstructed. Pictures have been taken from the streets and from various buildings on which we were authorized to go upstairs (religious monuments, hotels, commercial centers, ...). We used our Olympus digital camera. Aerial photographs can be useful for the cutting of ground polygons and also for the horizontal placement of the roof tops. They are no helpful

for neither polygon extrusion nor for roof construction because vertical lines are not visible (except if they are taken at quite a low level). Some pictures have been used as textures for polygonal surfaces, giving more reality effect to the simulation. Pictures must be taken perpendicularly to the surface that interests us, thus avoiding perspective vision problems (balconies, window-blinds). Furthermore, pictures have been unbent, and cleaned from undesired foreground objects.

Figure 5 shows results for the area around the cathedral: (a) is the result of the viewpoint recovery, (b) is the 3D reconstruction of buildings, (c) and (d) are examples with textures.

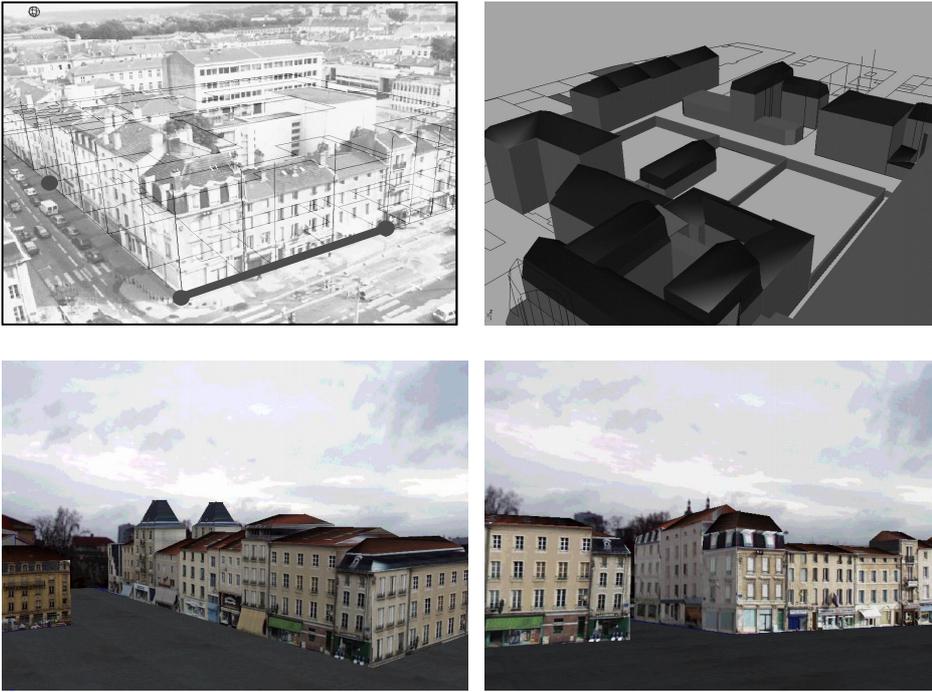


Figure 5. Results for the city of Nancy

8. FUTURE WORKS AND CONCLUSION

We have presented in this paper our system for interactive 3D rough reconstruction of urban areas from cadastral plans. However the method can be applied to any kind of known 3D data of an existing scene.

Perspectives for this work are principally to deal with curved roofs. Medina has been developed with Open Inventor library. We plan to write it

under the Maya environment. Another possibility is to develop the two modules (the automatic one and the interactive one) in another modeller environment: Maya's functionalities do not correspond to what urban designers or architects need. A solution would be to use AutoCad.

9. REFERENCES

- Allani-Bouhoula, N., and J.P. Perrin, 1998, "La Reconstruction Tridimensionnelle de Tissus urbains", *First International Conference on New Technologies for Decision Making in Civil Engineering*, Montréal, p721-732.
- Berger, M.O., C. Chevrier and G. Simon, 1996, "Compositing computer and video image sequences: Robust algorithms for the reconstruction of the camera parameters", *The International Journal of The Eurographics Association, Computer graphics forum*, p....
- Chevrier, C, 1996a, *Génération de séquences composées d'images de synthèses et d'images vidéo*. PHD Thesis, Université Henri Poincaré, Nancy.
- Chevrier, C, 1996b, "Handling interactions between real and virtual worlds", *Proceedings of the International Conference Computer Graphics International '96*, p 115-125.
- Collins, R.T., A.R. Hanson, E.M. Riseman and H. Schultz, 1995, "Automatic extraction of buildings and terrain from aerial images", *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, A. Gruen, O. Kuebler, P. Agouris, Birkhauser Verlag, Ascona, Switzerland, p.169-177.
- Devernay, F and O. Faugeras, 1995, "Automatic Calibration and Removal of Distortion from scenes of structured environments", *Proceedings of the SPIE Conference on Investigate and Trial Image Processing*, 2567, SPIE, San Diego, CA, July 1995.
- Faugeras, O, S. Laveau, L. Robert, C. Zeller and G. Csurka, 1995, "3D Reconstruction of Urban Scenes from Sequences of images", *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, A. Gruen, O. Kuebler, P. Agouris, Birkhauser Verlag, Ascona, Switzerland.
- Horand, R and O. Monga, 1993, *Vision par ordinateur, Outils fondamentaux*. Hermes, 14, rue Lantiez, 75017 Paris, first Edition.
- Peuchot, B, 1994, "Utilisation de détecteurs subpixels dans la modélisation d'une caméra", *Actes 9^{ème} Congrès AFCET Reconnaissance des Formes et Intelligence Artificielle*, Paris, pages 691-695.
- Simon, G and M.O. Berger, 1999, "Registration with a Zoom Lens Camera for Augmented Reality Applications", *Second International Workshop on Augmented Reality*, San Francisco.
- Toscani, G, 1997, *Systèmes de calibration optique et perception du mouvement en vision*. PHD Thesis, Université de Paris-Sud, Orsay.
- Wernecke, 1995, *The Inventor Mentor : programming object-oriented 3D graphics with Open Inventor*. Addison-Wesley Publishing Compagny.
- Canoma, 1999 <http://www.metacreations.com/products/canoma>
- Maya, 2000 *Using Maya Basics and MayaLive*, Alias WaveFront.
- PhotoModeler 2000 <http://www.eosystems.com>
- RealViz, 2000 <http://www.realviz.com>