# DEVELOPMENT OF A COMPUTER PLATFORMFOR OBJECT 3D RECONSTRUCTIONUSING COMPUTER VISION TECHNIQUES

Teresa Azevedo
*INEGI – Inst. de Eng. Mecânica e Gestão Industrial, LOME – Lab. Óptica e Mecânica Experimental*
*FEUP – Faculdade de Engenharia da Universidade do Porto*
*Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal*

João Manuel R. S. Tavares, Mário A. Vaz
*INEGI / LOME, FEUP*
*Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal*

Keywords: Computer Vision, Active Vision, 3D Reconstruction, Structure from Motion.

Abstract: In this paper we describe the development of a Computer Platform, whose goal is to recover the three-dimensional (3D) structure of a scene or the shape of an object, using Structure From Motion (SFM) techniques. SFM is an Active Computer Vision technique, which doesn't need contact or energy projection. The main objective of this project is to recover the 3D shape of an object or scene using the camera(s)'s or object's movement, without imposing any kind of restrictions to it. Starting with an uncalibrated sequence of images, the referred movement is extracted, as well as the camera(s) parameters, and finally, the 3D geometry of the object or scene is inferred. Shortly, in the first section of this paper the goals are defined; in the second, the computer platform is presented, as well as some experimental results; in the third and last section, the conclusions relative to the study and work done are drawn and, finally, some perspectives of future work are given.

## 1 INTRODUCTION

Computer Vision is continuously trying to develop theories and methods for automatic extraction of useful information from images, in a way as similar as possible to the complex human visual system.

Contactless techniques to recover the 3D geometry of an object are usually divided in two classes: active techniques, that require some kind of energy projection or the camera's or object's movement to obtain information about the object's shape, and passive techniques, that only use ambient light and so, usually, the extraction of 3D information becomes more difficult.

The main goal of this work is to obtain 3D models of objects using a Structure From Motion (SFM) methodology, which is a Active Vision technique (Pollefeys, 1998). Along time, SFM technique has received several contributions and diverse approaches. In the present case, we do not want to impose any kind of restrictions to the movement involved: starting from an uncalibrated image sequence of an object, we intend to extract the referred movement (camera(s)'s or object's), to calibrate the camera(s) used, and finally to obtain the 3D geometry of the object in cause.

To help accomplishing our goals, a modular computer platform has been developed, with a graphical interface, into which functions, from several libraries of public domain, are being integrated.

The functions already integrated enclose several Computer Vision techniques, such as: feature points extraction and matching between images, epipolar geometry determination, rectification and dense matching (Azevedo, 2005). These techniques are usually used in 3D shape extraction of objects starting from an image sequence (Figure 1).



Azevedo T., Manuel R. S. Tavares J. and A. Vaz M. (2006).
DEVELOPMENT OF A COMPUTER PLATFORMFOR OBJECT 3D RECONSTRUCTIONUSING COMPUTER VISION TECHNIQUES.
In *Proceedings of the First International Conference on Computer Vision Theory and Applications*, pages 383-388
DOI: 10.5220/0001360003830388

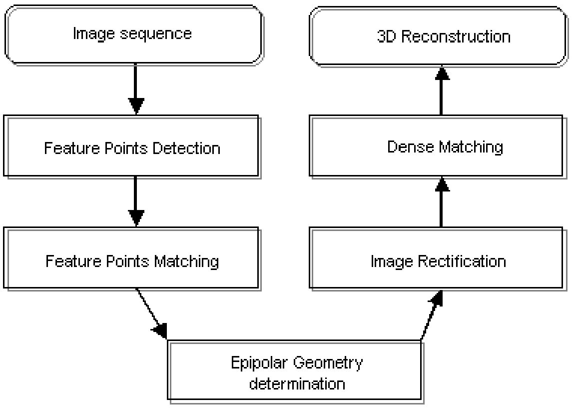


Figure 1: Adopted methodology for 3D reconstruction of objects from a sequence of images.

## 2 COMPUTER PLATFORM

The computer platform has been being developed in $C^{++}$, with the tool *Microsoft Visual Studio*, using *MFC* libraries (*Microsoft Foundation Classes*). It has a graphical interface and a modular structure allowing the user to comparatively analyse the performance of each function.

Several functions for 3D reconstruction are already available, integrated from five software programs and one computational library, all open source (for more information about them see (Azevedo, 2005)). Further on, these entities will be referred generically as *Programs 1* to *6*:

1. *Peter's Matlab Functions for Computer Vision and Image Analysis* (Kovesi, 2004): Matlab functions for image processing and analysis and 3D Vision;
2. *Torr's Matlab Toolkit* (Torr, 2002): Matlab software with a graphical interface, that applies some SFM techniques between two images;
3. *OpenCV* (2004): $C^{++}$ functions library which implement the most common algorithms in the Computer Vision domain;
4. *Kanade-Lucas-Tomasi (KLT) Feature Tracker* (Birchfield, 2004): functions in *C* that implement the KLT algorithm for feature points extraction and matching in sequences of images;
5. *Projective Rectification without Epipolar Geometry* (Isgrò, 1999): functions in *C* that perform rectification in stereo images pairs, without previous determination of the epipolar geometry;
6. *Depth Discontinuities by Pixel-to-Pixel Stereo* (Birchfield, 1999): *C* program that returns disparity and discontinuity maps between two rectified images.

In order to integrate them conveniently into the computer platform, programs originally written in *Matlab* were ported to *C*, using the *Matlab Compiler* toolbox.

In the referred computer platform, for each available Active Vision technique, the user can easily choose the algorithm to use (Figure 2), as well as conveniently define its parameters (Figure 3).

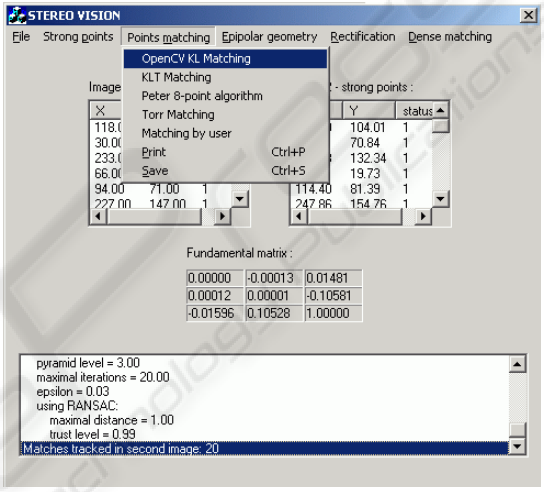


Figure 2: For each platform menu item, several algorithms are available.

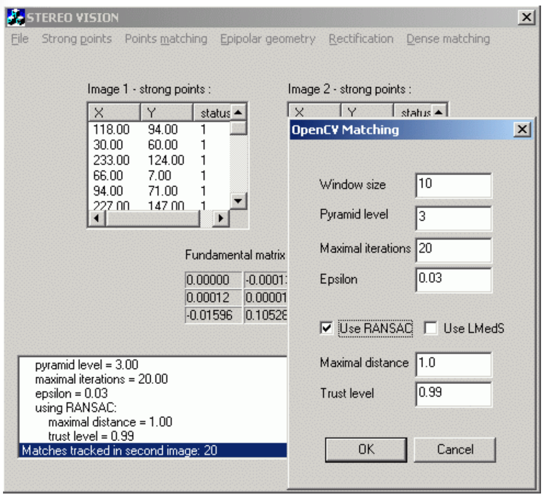


Figure 3: Example of setting the parameters for the *program OpenCV* matching algorithm.

## 3 EXPERIMENTAL RESULTS

For experimental results, stereo images where used, with dimensions $540 \times 612$ pixels, obtained from several real objects, captured with an off-the-shelf digital camera.

Some of the experimental results obtained are presented here following the sequence indicated in Figure 1.

### 3.1 Feature Points Detection

Feature or interesting points are those who reflect the relevant discrepancies between their intensity values and those of their neighbours. Usually, these points represent vertices of objects, and their detection allows posterior matching between the images of the sequences.

Only *programs 1* to *4* have algorithms for feature points detection. As an example, in Figure 4 are some results obtained with these *programs*, using one of the test images. It is possible to observe that, in this case, *program 2* presents the weakest results, especially because there are some areas with high density of detected feature points. This happens because this technique does not ensure a minimal distance between detected feature points, as happens with *programs 3* and *4*.

### 3.2 Matching

Matching is the 2D points association between sequential images, which are the projection of the same 3D object point. Automatic detection of matching points between images can be achieved using several matching measures (Azevedo, 2005).

Again, only *programs 1* to *4* have algorithms for feature points matching. Experimental results, obtained using the matching techniques integrated into the computer platform, are shown in Figure 5.

It is possible to observe that matching is a critical stage in 3D Vision. For all techniques, there are correlation mismatches, where *program 2* is again the one that presents the weakest results.

### 3.3 Epipolar Geometry

Epipolar geometry corresponds to the geometrical structure between two distinct points of view and expresses itself mathematically by the fundamental matrix $F$, of size 3x3: $m'^T F m = 0$, where $m$ and $m'$ are the matching points.

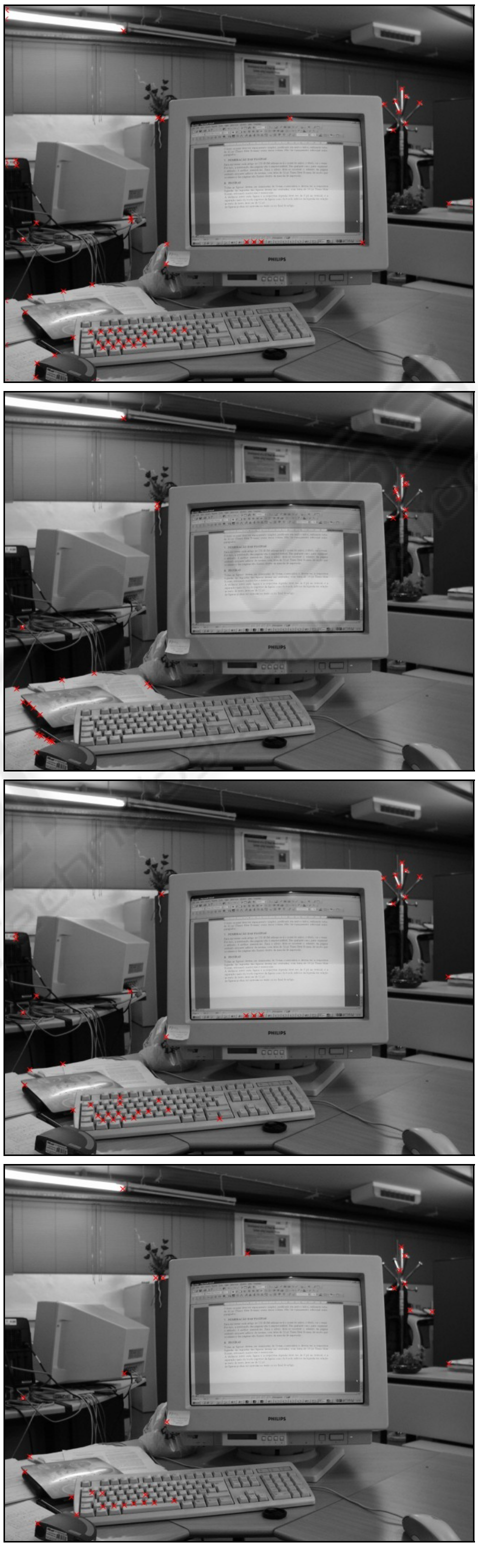

Figure 4: Results of feature points detection in one image out of the stereo pair: red crosses are the detected interesting points (top to bottom: results from *programs 1* to *4*).

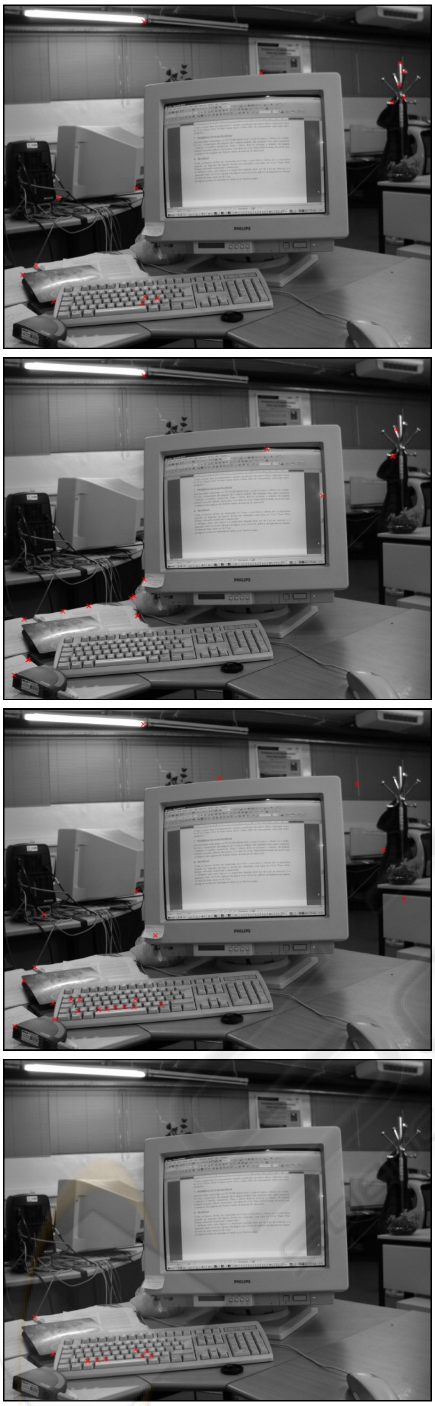

Figure 5: Red crosses are the matched interesting points in the other image of the stereo pair (top to bottom: results from *programs 1* to *4*).

Determining the epipolar geometry means getting relative pose information between two different views (images) of an object. That information allows also eliminating some previous wrong matches (outliers), as well as to make it easier to get new matching points (dense matching).

Only *programs 2* and *3* have algorithms for this purpose. Some experimental results obtained by the techniques integrated in our computer platform are given in Figure 6. Because of the weak results obtained in the prior step and presented in the previous subsection, it was not possible to determine the epipolar geometry by using *program 2*, because the number of outliers was considerably higher than the inliers's. *Program 3* (*OpenCV*) performs this step very well, although it is very important to get good matching points to correctly estimate the epipolar geometry.

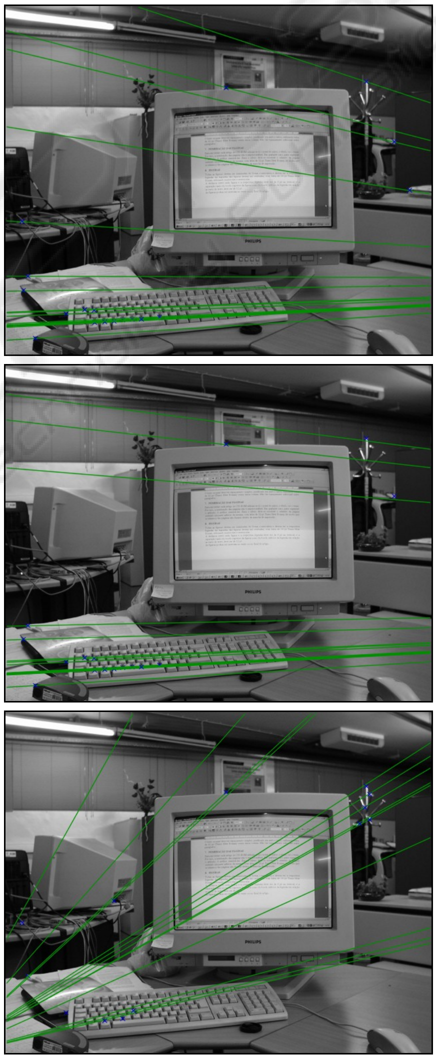

Figure 6: Epipolar lines (green) and inliers (blue) in the first image of the stereo pair, after epipolar geometry determination; top to bottom: *program*'s *1* result using *Ransac* algorithm (Fischler, 1981), *program 1*'s result using *LmedS* algorithm.

## 3.4 Rectification

The operation of changing two stereo images, in order to make them coplanar is usually known as Rectification. Performing this step makes dense matching much easier to obtain.

*Program 5* is the only one that makes rectification, without previous determination of the epipolar geometry. The results from the referred program are shown in Figure 7.

It is possible to observe that the quality of the results is proportional to the quality of the epipolar geometry determination.

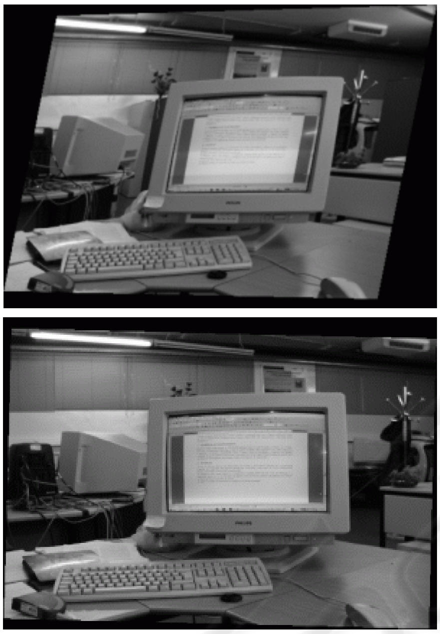

Figure 7: Rectification result obtained with *program 5*.

## 3.5 Disparity Map

A disparity map codifies the distance between the object and the camera(s): closer points will have maximal disparity and farther points will get zero disparity. For short, a disparity map gives some perception of discontinuity in terms of depth.

Only *programs 3* and *6* perform this operation, with the difference that *program 6* also returns a discontinuity map.

This step was tested with already rectified images included in the package of the *program 6* (Figure 8). The results obtained are presented in Figure 9 and Figure 10. Given two rectified images, it is possible to observe that both *program 3* and *6* perform very well on the determination of the disparity map.

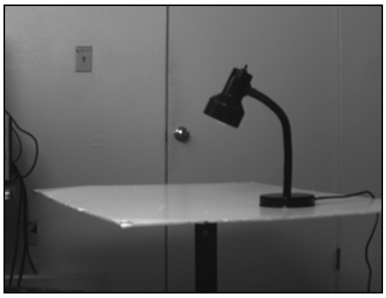

Figure 8: One of the original images used by *programs 3* and *6* to obtain the disparity map (associated to dense matching).

Figure 9: Disparity map, obtained by *program 3*.

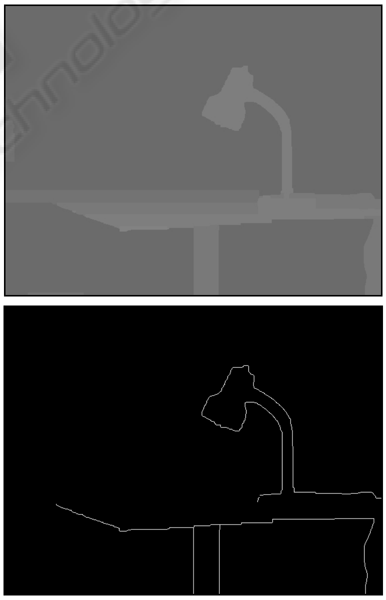

Figure 10: Disparity (top) and discontinuity (bottom) maps, obtained by *program 6*.

# 4 CONCLUSIONS

Along this work, the developing of a computer platform was initiated, with an appropriated graphical interface and a modular design, to allow the application and study of several techniques of 3D Active Vision, with the final goal of object 3D reconstruction. In short, the functions already integrated in the referred computer platform and experimentally analysed, obtain good results when applied to objects with strong characteristics. From the same used results, it is possible to conclude that low quality results are strongly correlated with strong points detection and matching, as the functions in the further steps of the 3D reconstruction methodology adopted (Figure 1) are based on those points.

## 5 FUTURE WORK

The next steps of this work will focus on improving the results obtained when the objects to be reconstructed have smooth and continuous surfaces. To do so, the approach will be:

- inclusion of space carving techniques for object reconstruction (see for example, (Kutulatos, 1998), (Sainz, 2002), (Montenegro, 2004));
- the strong points to use in the 3D space object definition will be detected with the use of a reduced number of markers added on the object;
- inclusion of a camera calibration technique, as well as pose and motion estimation algorithms; some of the techniques to consider are (Meng, 2000) and (Zhang, 2000).

Finally, the computer platform will be used in 3D reconstruction and characterization of 3D external human shapes.

## ACKNOWLEDGMENTS

This work was partially done in the scope of the project "Segmentation, Tracking and Motion Analysis of Deformable (2D/3D) Objects using Physical Principles", reference POSC/EEA-SRI/55386/2004, financially supported by FCT - Fundação de Ciência e Tecnologia in Portugal.

## REFERENCES

M. Pollefeys, R. Koch, M. Vergauwen, L. V. Gool, *Flexible acquisition of 3D structure from motion*, Proceedings of the IEEE Workshop on Image and Multidimensional Digital Signal Processing, Alpbach, Austria, pp. 195-198, 1998.

T. Azevedo, J. M. R. S. Tavares, M. Vaz, *Obtenção da Forma 3D de Objectos usando Metodologias de Reconstrução de Estruturas a partir do Movimento*, Congreso de Métodos Numéricos en Ingeniería, Granada, Espanha, 2005.

P. Kovesi, *MATLAB Functions for Computer Vision and Image Analysis*, http://www.csse.uwa.edu.au/~pk/Research/MatlabFns, 2004.

P. H. S. Torr, *A Structure and Motion Toolkit in Matlab*, http://wwwcms.brookes.ac.uk/~philiptorr/Beta/torrsam.zip, 2002.

*OpenCV - Open Computer Vision Library*, beta 4, http://sourceforge.net/projects/opencvlibrary, 2004.

S. Birchfield, *KLT: An Implementation of the Kanade-Lucas-Tomasi Feature Tracker*, http://www.ces.clemson.edu/~stb/klt/index.html, 2004.

F. Isgrò, E. Trucco, *Projective rectification without epipolar geometry*, Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Fort Collins (Colorado), USA, vol. 1, pp. 94-99, 1999.

S. Birchfield, *Depth Discontinuities by Pixel-to-Pixel Stereo*, International Journal of Computer Vision, vol. 35, no. 3, pp. 269-293, http://vision.stanford.edu/~birch/p2p/, 1999.

M. A. Fischler, R. Bolles, *RANdom SAmpling Consensus: a paradigm for model fitting with application to image analysis and automated cartography*, Communications of the Association for Computing Machinery, vol. 24, no. 6, pp. 381-395, 1981.

K. N. Kutulatos, S. M. Steiz, *A Theory of Shape by Space Carving*, Technical Report TR692, Computer Science Department, University of Rochester, New York, USA, 1998.

M. Sainz, N. Bagherzadeh, A. Susin, *Carving 3D Models from Uncalibrated Views*, Proceedings of the 5th IASTED International Conference Computer Graphics and Imaging, Hawaii, USA, pp. 144-149, 2002.

A. Montenegro, P. Carvalho, M. Gattass, L. Velho, *Space Carving with a Hand-Held Camera*, Proceedings of the SIBGRAPI'2004 - International Symposium on Computer Graphics, Image Processing and Vision, Curitiba, Brasil, 2004.

X. Meng, H. Li, Z. Hu, *A New Easy Camera Calibration Technique Based on Circular Points*, Proceedings of the British Machine Vision Conference, University of Bristol, UK, pp. 496-501, 2000.

Z. Zhang, *A Flexible New Technique for Camera Calibration*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, 2000.