# COMPUTER VISION TECHNIQUE FOR 3D MODEL RECONSTRUCTION OF HISTORICAL BUILDING FROM PHOTOGRAPHS

Florin Bogdan MARIN[1], Mihaela MARIN[1]

[1]"Dunarea de Jos" University of Galati, Department of Material Science and Engineering
florin.marin@ugal.ro

## ABSTRACT

*The architectural heritage is considered part of every nation. The historical buildings represent people's identity and past. Nowadays the technology allows to produce 3D models of buildings using laser scanners or photogrammetry, but these technologies cannot be used to reconstruct from one or few past images, as in the case of lost buildings. Moreover, pictures taken might not provide a clear image of details of historical buildings. This paper presents a computer vision technique for 3d model reconstruction of historical buildings. A software proof-of-concept application is presented to reconstruct historical buildings. The new approach however requires high computational resources and might provide the solution in cases where a historical building is described by one single picture.*

**Keywords:** truck breaks, steep grades, temperature evolution, breaks overheating, SVM prediction vehicle

## 1. INTRODUCTION

Nowadays computer vision algorithms have experienced important development and the computing resources are provided by powerful processor platforms. Although computer vision and image processing algorithms require important computational power, the hardware platforms represented by computer processor as well as graphics cards processor allow implementation in engineering applications.

There are many old historical buildings which are in ruins and which were constructed centuries ago or even millenniums as in case of ancient constructions. In case of old buildings in the historical neighbourhoods canters of towns, the conservation and restoration of these buildings are both a challenge because of the cost and of the difficulty of restoration procedures. A complete 3d model of the historical building would be helpful in many ways [1]. Firstly, this technique will allow understanding the architecture history. Secondly, such a technology will be possible for a virtual view and visit of historical buildings which can be used as learning tools and tourism advertisement. Immersion in a computer generated 3d environment proved to be a very good tool for learning history or architecture. The virtual map could be presented at museums entrance to allow visitors to initially view the maps and to choose buildings of interest. The walk-through movies would also provide the visitor with further information on the region to be visited.

The virtual reconstruction of historical buildings received much attention in case of lost buldings [1].

The possibilities of taking advantages of 3d reconstruction of historical buildings are many despite the difficulties in reconstructing whose amount is important. The difficulties concern the lack of drawings, dimensions or other descriptive documents. In case of lost historical buildings the virtual reconstruction of the 3d model implies using old photos with little visual information. The process requires significant time resources and technologies [16].

A limited amount of evidence is available for existing buildings as legal conditions, high cost, occupational safety procedures will not allow measurement of many architectural features of buildings. In case of lost buildings the pictures were produced at an oblique angle, and so other technical challenges emerge as it is obviously hard to make calculations.

3D reconstruction from a single view has been researched by many scholars [2] [3] [4]. Several approaches are used by exploiting user-selected edge features [5] or by assuming geometrical constraints provided by the location of vanishing points [6].

Others used monocular depth perception and reinforcement learning to virtually build environments

[7][8]. The reconstruction using multiple images is a common approach for many research groups [9] [10] [11], but the reconstruction of a historical building, with complex architecture features, is challenge in case of one single picture as input. 3D scene reconstruction from multiple images is a well-known topic, scholars and industry provided reliable solutions. 3D reconstruction algorithm supposed to solve parameters estimations of several dozens of pictures to find an acceptable result [12][13][14].

A virtual 3d model of historical buildings would be significant from the historical research point of view.

Our goal is to create a full photo-realistic 3d model from a single image using an existing historical building in order to compare the result. A further step will mean virtual reconstruction of a lost historical building using a single picture and few or no information concerning the construction.

## 2. PROBLEM STATEMENT

The scope of this research is to identify a computer vision technique to be used in a semi-automatic procedure to allow virtual 3d reconstruction of a building using one single picture, such in case of lost historical buildings. Furthermore, in this scenario we have no technical information such as drawing, dimensions in order to achieve the goal of virtual reconstructions.



**Fig. 1.** *The picture used for 3d reconstruction*

We used as example a historical building that still exists (Fig. 1) in order to make assumptions concerning the validation of the technique. We had to have an accurate estimate for certain errors produced in the process. There are several challenges in pursuing the objective of reconstructing a historical building.

## 3. PROBLEM SOLUTION

To obtain enough historical data of building with important height and to find out architectonic details is a challenging task. This also becomes mandatory in case of historical building reconstruction of existing buildings as installing measurement devices cannot be used due to the advanced level of destruction of building or because of legal issues.

3D reconstruction techniques involve solving distance estimation problems that must be extracted from ambiguous information. Heuristics-based algorithm is used in order to deal with ambiguities provided by one single picture. In our case there are no several pictures but a single one. The main idea is to operate several virtual movements such as rotation and skew for determining amounts in order to obtain a set of images to allow SLAM (Simultaneous Localization And Mapping) [15] algorithms to be used.

In the Fig. 2 only few movement are presented to allow the reader to understand the concept, however when using high resolution pictures the "zoom in" on the buildings of interest can help a lot. This allows identifying certain details in the pictures and would also provide the SLAM algorithm to identify points and add more information concerning the building in question.
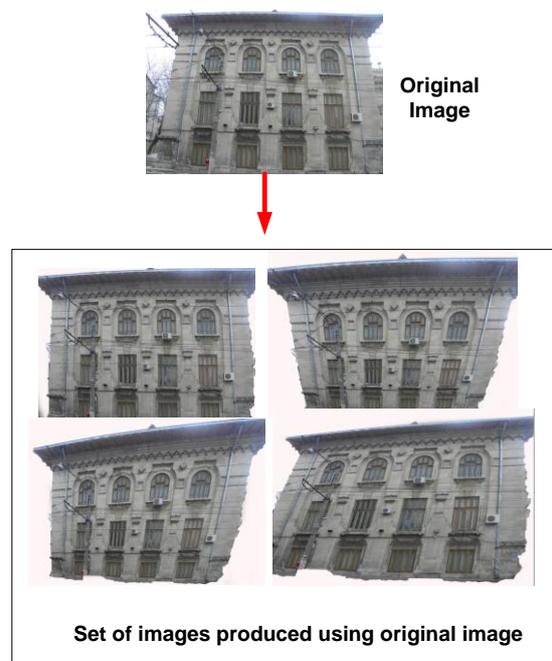


**Original Image**

**Set of images produced using original image**

**Fig. 2.** *Set of images produced*

Here we propose a 3-D reconstruction method that uses a single picture.

The virtual camera that moves in certain angles does not need to get calibrated. In fact, we only need to approximate the focal length of the lens. In case the technique will be used on a single old photo that will not be possible. By using several deformation of the image simulating in this way moving mainly along vertical axes and rotating in small amount, we simulate the existence of images to enable the identification of 3d points to build a surface. Consequently there are multiple coordinate systems in

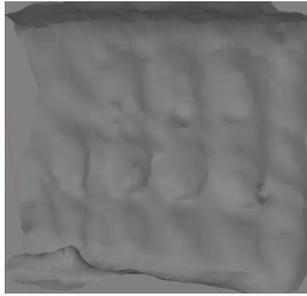the technique proposed. In Fig. 3 is shown the surface resulted.



**Fig. 3.** *Surface resulted*

We developed a method that computes the 3d points cloud formation, using orientations and the distances resulted. When the 3d points are determined, the image features correspondence is needed to identify important features of buildings such as corners or other easily identifiable points.

The main steps are depicted in Fig. number 4.?

Once we know or predict a measure compared to the building scale, we can have main buildings dimensions and we can apply the proposed 3-D reconstruction method.

The next step builds a depth map for each object based on the images produced virtually by the moving camera. This artificial camera movement generates the visual hull of the object which is a depth map of the buildings and is a good approximation of the original object concerning depth features. While the vertical and horizontal dimensions and proportions can be predicted by using a scale, the depth of some features (such as windows) cannot be predicted by a human.

In CAD terms, a fully manual method of rebuilding a 3d building can be translated in predicted dimensions that are in the some plane, considering reference dimensions. However in case of information provided by single picture, once the buildings depth map are produced by the user by manually using separate planes, creating a wireframe model from a deformed surface in order to build longitudinal and transversal features.

In Figure 4 are shown the main steps of the proposed technique.

3D shape reconstruction is done semi-automatically by using the depth map. Our approach detects occlusions explicitly. The depth map representation (surface generated), which produces the Z axes information according to the proposed technique, is able to represent partially occluded parts of the building, such as in the case of objects as trees, cables that are relatively small compared to the building overall dimensions. In this respect, we encountered none of the issues as the considered building has a lot of features that are the same. This is the case of many historical building windows, many features are basically copied in the façade of the

building. In Fig. 5. is depicted the 3d model of the historical building reconstructed.

The hardware setup we used is an AMD FX 8320 8 core processor, with 8000 points in the PerformanceTest benchmarking software, with 32 Gb RAM. The development environment is Visual Studio 6 with OpenGL for building 3d model. As computer vision library we have used the well-known OpenCV 2.4 library from Intel Inc. Is important to note that no graphic card capabilities were used during automatic processing, though CUDA features might allow future processing speed. The time needed for automatic steps is 10 hours for real scale.
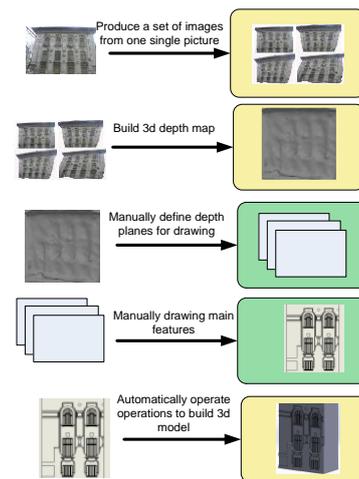


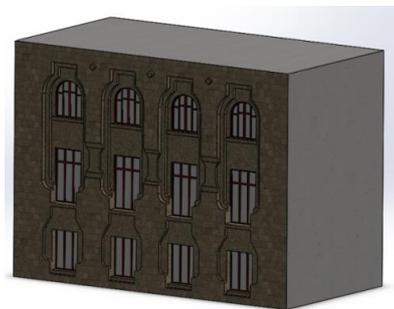**Fig. 4.** *Main Steps of the proposed technique*



**Fig. 5.** *3d Model of the building with reconstructed façade*

The only input is a single picture from one single point of view. The first step, which is automatic and done by the developed software application, is to produce a set of images that are produced by using transformations with known measures. The next step is to use the set of images to build 3d depth map information consisting of points in space that are smoothed to a surface. The surface cannot describe the real surface of the façade of the building; however it is an acceptable or good indicator of depth for certain architecture features.

The next step is to manually select planes in the 3d depth map. This is necessary in order to set

how many reference planes are needed and the distance between them. This is required for the manual step of manually drawn features for each of the plane previously determined. This is most times a consuming operation done by an operator. The operator draws by over imposing on the original image, processed in such a way to minimize parallax error. We have used FreeCAD software and the output was an IGS file to be inputted in our proof-of-concept application.

The final step, performed automatically, is to build 3d information. By inputting each plane drawing done in the free CAD software, and using previous steps information, the program is able to build 3d CAD model of the building.

Iterative matching of similar points on the 2d pictures is performed in order to find 3D discretized cloud points. We assume that there are no hidden points, such as in the case of reconstruction using pictures around the object scanned. However not all points can be identified in all scenes. The size of the points to be identified in the pictures is variable and is not about pixels but groups of pixels. The spatial resolution corresponds to the picture resolution. Not all groups of pixels are considered, but only those to be easily identified, such the  corners is taken as its projection.

One important thing is that we use no pixel intensities of the set of images to obtain estimates for 3d triangulation to obtain 3d model, but we use Harris corner detection algorithm in OpenCV 2.4 We have presented reconstruction results in Fig.5.

The algorithm converges using 30-40 images though we have used 60 generated images.

In order to assess the errors of the technique, we have measured some features and resulted errors between 3% and 10% for dimensions which are acceptable errors taking into account that one single picture has been used. As far as the curves modeling is concerned, the angular errors are about 5%. The errors depend on the relative distance from the real camera distance to the feature to be reproduced. The closer the distance, the lower the error. We presume that in case a single picture is taken from a position at a considerable distance and a high resolution, the errors will minimize.

Figure 3 gives the reconstructed 3D model building which is a reasonable representation for the look and feel of a viewer. The best look and feel for the viewer of a building is to view it on a monitor of a an important size.  For numerical evaluation, the mean 2D line error is 6% as we assessed on 50 dimensions in the façade.

## 4. CONCLUSIONS

The technique for 3d model reconstruction of historical buildings presented in this paper enabled the authors to build a proof-of-concept application. In this paper we have presented a 3d reconstruction technique for reconstruction of surfaces from a number of views that are artificially produced using one picture.

The new approach however requires high computational resources and might provide the solution in cases where a historical building is described by one single picture. However, from historical perspective, this technique might fail in some specific cases. While we have used a picture taken of an existing building in order to asses errors, in case of old sketches, postcard pictures or a painted style picture, where the authors introduce their own colours and dimensions for a better look, the developed technique will fail. The pseudo-colour images introduce errors due to the accuracy of the dimensions. The development of the software might continue to integrate different methodologies for 3D reconstruction from a single view.

## REFERENCES

[1] **Hogg**. *Automated reconstruction of 3d models from real environments*. ISPRS Journal of Photogrammetry and Remote Sensing, 54:1–22, 1999.
[2] **Debevec P.E., Taylor C.J., Malik J.**, *Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach.* Proceedings of SIGGRAPH'96, Fairbairn, 11–20 pp., 1996.
[3] **Heuvel F.A**. *3D reconstruction from a single image using geometric constraints*. Journal Photogrammetry and Remote Sensing, 1998.
[4] **Barron, J. T. and Malik J**. *Color constancy, intrinsic images, and shape estimation*. In Europ. Conf. on Computer Vision, 57–70 pp., 2012.
[5] **Hoiem D., Efros A. A. and Hebert M**. *Automatic photo pop-up*. ACM Trans. Graph., 24(3): 577–584 pp., 2005.
[6] **Horry Y., Anjyo K.-I. and Arai K.** *Tour into the picture: using a spidery mesh interface to make animation from a single image*, SIGGRAPH '97,     New York, NY, USA. ACM Press/Addison-Wesley Publishing Co., 225–232 pp.,  1997.
[7] **Prasad M., Zisserman A., and Fitzgibbon A. W.** *Single view reconstruction of curved surfaces*, CVPR, 1345– 1354 pp., 2006.
[8] **Terzopoulos D., Witkin A., and Kass M**. *Symmetry-seeking models and 3D object reconstruction*. Int. J. of Computer Vision, 1, 211–221 pp., 1987.
[9] **J. Michels, A. Saxena, and A. Y. Ng**, *High speed obstacle avoidance using monocular vision and reinforcement learning*, in International Conference on Machine Learning (ICML), 2005.
[10] **E. Delage, H. Lee, and A. Y. Ng**, *A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image*, in Computer Vision and Pattern Recognition (CVPR), 2006.
[11] **D. Hoiem, A. Efros, and M. Herbert**, *Geometric context from a single image*, in International Conference on Computer Vision (ICCV), 2005.
[12] **P. Felzenszwalb and D. Huttenlocher**, *Efficient graph-based image segmentation*, IJCV, vol. 59, 2004.
[13] **T. Nagai, T. Naruse, M. Ikehara, and A. Kurematsu**, *Hmm-based surface reconstruction from single images*, in Proc IEEE International Conf Image Processing (ICIP), vol. 2, 2002.
[14] **T. Hassner and R. Basri**, *Example based 3d reconstruction from single 2d images*, in CVPR workshop on Beyond Patches, 2006.
[15] **Caprile, B. and Torre V.** *Using vanishing points for camera calibration*. IJCV 4 (29), 1990
[16] **M.W. Wright, G.C. Watson, and R.L. Middleton**. *Towards the integration of photogrammetric and laser range data for architectural reconstruction*. International Symposium on Virtual and Augmented Architecture, Dublin, Ireland, pages 255–266, 2001