PROJECT PHOTOFLY : NEW 3D MODELING ONLINE WEB SERVICE
(CASE STUDIES AND ASSESSMENTS)

D. Abate [a], G. Furini[a], S. Migliori [b], S. Pierattini [a],

[a]ENEA Research Centre, UTICT, Bologna, Italy - (dante.abate, graziano.furini, samuele.pierattini)@enea.it
[b] ENEA Headquarter, UTICT, Roma, Italy – silvio.migliori@enea.it

Commission V, WG 4

KEY WORDS: Photogrammetry, Computer Vision, Automation, Reconstruction Algorithms

ABSTRACT: During summer 2010, Autodesk has released a still ongoing project called Project Photofly, freely downloadable from AutodeskLab web site until August 1 2011.
Project Photofly based on computer-vision and photogrammetric principles, exploiting the power of cloud computing, is a web service able to convert collections of photographs into 3D models.
Aim of our research was to evaluate the Project Photofly, through different case studies, for 3D modeling of cultural heritage monuments and objects, mostly to identify for which goals and objects it is suitable. The automatic approach will be mainly analyzed.

1. INTRODUCTION

Sometimes it is not easy to digitize cultural heritage artifacts, due to permits granted by public authority, environmental features, limited time available on site, and so on and so forth. In these cases the speed of acquisition is a key factor during field survey and photography ensures it.

The final goal of the project must also be taken into account: visualization purposes for large audiences; restoration; recording and documentation.

Indeed expensive and time consuming photogrammetric approaches or laser scanning survey are not worth if the final goal of the project is just visualization, and high accuracy of the 3D model is not required.

Photofly technology was acquired on May 2008 from Realviz. After few years of research and development conducted by AutodeskLab, Photofly was released on July 22 2010. Realviz, one of the few available commercial technologies, developed the image processing software products, which enabled the production of 3D digital content of a set of images.

Structure from motion (SfM) is the main concept which stands behind a fully automated 3D modelling approach from set of markless images. Recently some tools are available. In 2006 M. Vergauwen and L. Van Gool released a Web-based 3D reconstruction service (*Vergauwen and Van Gool, 2006*). The automatic process, running on a remote server, computes the camera calibration together with the depth maps of the images uploaded. The result is then downloaded on the user's PC. Other two projects which exploit the SfM concepts are Bundler, developed at the University of Washington, USA (*Snavely et al. 2008*) and Samantha developed at the University of Verona, Italy (*Farenzena et al. 2009*).

Both are able to deal with huge number of images. Bundler is SfM system for unordered image collections able to create sparse point clouds. An earlier version of it was used for the Photo Tourism project (Snavely *et al. 2008*). To obtain denser point clouds Bundler can be used in conjunction with PMVS2 (*Furukawa and Ponce 2007*). Samantha appears to be even faster than Bundler. It shows a lower computational complexity and introduces a local bundle adjustment procedure.

A new methodology for the automatic identification of an images correspondence, for a wide range of image datasets, is represented by ATiPE. The acronym stands for Automatic Tie Point Extraction (*Barazzetti et al. 2010*). It distinguishes itself from other software packages because of its flexibility, capability to process different kinds of images and completeness of the obtained point cloud.

1.1 Project Photofly. How it works.

Project Photofly is a web service technology .
The user has to install the "Photo Scene Editor" to upload images, and possibly to modify manually the photo-scene returned by the system.
Although Pictures can be shot with various camera and various focal length, it's recommended to use the same camera, even a commercial camera, and the same zoom ratio for the entire scene dataset.
Better results are gained using a wide angle lens (20 to28 mm or more). Fish eye lenses are not supported by the Photofly engine, whereas ultra wide lenses may create a distortion in the images, which could not be compensated during the reconstruction.
During the photographs acquisition it is important to avoid objects occluding the target, and objects in movements in front of it, as well as flat and untextured surfaces which wouldn't help to create the depth information.
Once the photographs (jpeg or tif file formats), taken with standard digital cameras and following specific guidelines, are uploaded to the server, the system computes the photo-scene and sends a sparse point cloud back to the user. He can decide to wait online until the process is completed by the server, and

the file is sent back, or leave a valid email address, where he will be contacted later on, to download the 3D model (rzi file format).

Basically the server stitches the camera positions, and as many photographs as it can together, by itself using a series of computer vision algorithms called the "Camera Factory". Additional features can be applied locally after the download, and the final model can be eventually exported. Nowadays the Photo Scene Editor is version 1.2 and the point cloud can be exported in four file formats: dwg, las, fbx, rzi.

If the web service is unable to determine where some pictures belong, the user can manually identify at least four common points for each picture, in order to have the unused photo added to the global scene.
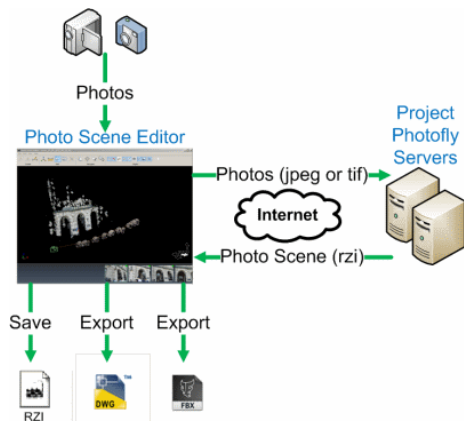
Figure 1. Project Photofly Pipeline

Once the scene is created the user can:

- add reference points to enhance the stitching process;
- add supplementary 3D points;
- add real measurements acquired on field;
- set the origin and axes of the global coordinates system.

It is also possible to draw lines directly onto the point cloud and export just them to a CAD software package for further changes and improvements.

The system itself is able to use metadata stored in the photographs to get the camera features and improve the accuracy of the model. If the exif file is available, it is automatically read and the images are grouped by same focal length. However the application does never ask for those instructions directly.

2. CASE STUDIES

Pictures datasets of different cultural heritage artifacts, environments (squares) and buildings have been acquired, according with different dimensional scales and different cameras.

Each dataset was uploaded to the Project Photofly server and the results obtained by the automated process were analyzed.

2.1 Navona Square, Rome

Navona Square is placed in the historic city centre of Rome. Its shape is an ellipse. The major axis is 250 meters and the minor axis is 50 meters.

The photographic dataset of Navona Square consists of 92 pictures.

The automatic reconstruction was unsuccessful. The engine stitched together 83 pictures on a total amount of 92. However on 83 pictures, only 44 were correctly placed whereas the other 48 were mixed with each other. It follows that the final points
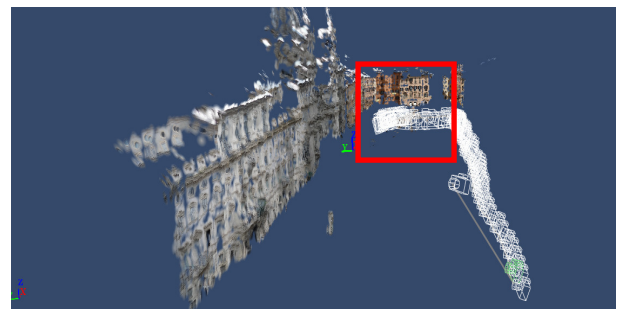

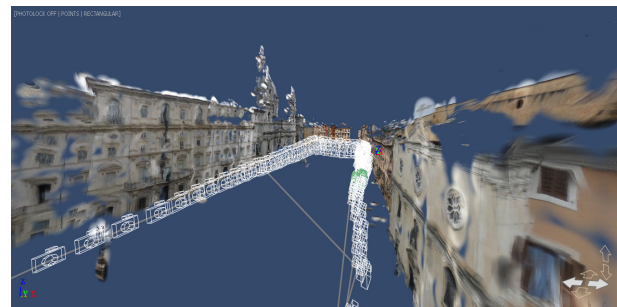Figure 2. Navona Square, Rome, Automatic Reconstruction

Figure 3. Navona Square, Rome, Manual Reconstruction

cloud produced (41452 points) contains almost 50% of wrong points.

Proceeding with the manual addition of references points the final model of the square was reconstructed, and a point cloud of 50811 points was created. This manual approach results to be really time consuming.

Camera: Samsung St45/ Vluu St45/ Samsung Tl90

2.2 Campidoglio Square, Rome

Campidoglio Square in Rome has a trapezoidal shape. Three sides are closed by buildings, while the fourth is open on a staircase. "Palazzo dei Conservatori" on the right, "Palazzo dei Musei Capitolini" on the left, together with "Palazzo Senatorio" at the bottom, feature the space. The two buildings on the sides are perfectly equal and symmetrical.

This dataset is composed by 56 photographs.

The automatic reconstruction was unsuccessful. Despite the relative low number of pictures, the Photofly engine was unable to reconstruct the model correctly, probably due to equality of the two buildings on the left and on the right side.

After the manual add of reference points it was possible to reconstruct the model correctly. The final point cloud consists of 33981 points.
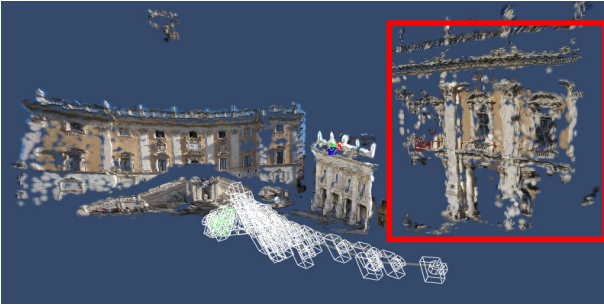
Camera: Canon Power Shot G10

Figure 4. Campidoglio Square, Rome, Automatic Reconstruction

Figure 5. Campidoglio Square, Rome, Manual Reconstruction

2.3 Maggiore Square, Bologna

Maggiore Square is the main square of the city of Bologna. It's 115 meters long and 60 meters wide, and it's surrounded by the most valuable buildings of the medieval city.
Seventy Eight pictures were shot.
After uploading all the photographs, the 3D reconstruction failed. The automatic 3D modeling was unsuccessful: only 42 photographs among 78, and not all correctly, were stitched; a building belonging to the square was placed in a wrong position; another building was modeled with wrong dimensions. The manual approach was necessary to obtain a correct model of the whole square.
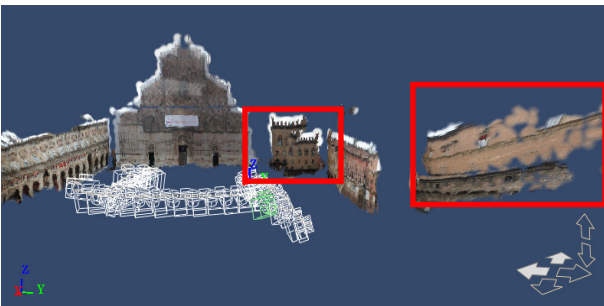Camera: Canon EOS 550D

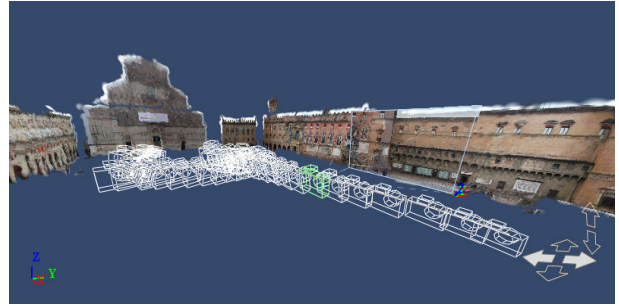Figure 6. Maggiore Square, Bologna, Automatic Reconstruction

Figure 7. Maggiore Square, Bologna, Manual Reconstruction

2.4 Buildings Facades

Palazzo d'Accursio is a palace in the city center of Bologna. It is located on the Piazza Maggiore. It includes a complex of edifices which were united during centuries.
The main two- storey facade, enriched by a clock tower, is 90 meters long and 40 meters high.
The photographs dataset is composed by 36 pictures.
The automatic reconstruction was successful. All the photographs were stitched and produced a point cloud of 34789 points.
Camera: Canon EOS 550D
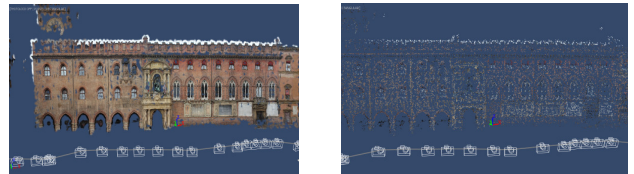
Figure(s) 8-9. Palazzo d'Accursio, Bologna, photographs view mode and points view mode, Automatic Reconstruction

Palazzo dei Diamanti in Ferrara was designed by Biagio Rossetti and its construction began in 1493. Its name is due to the shape of the elements which make up the external facades.
The two halves, separated by the main entrance door, which make up the façade, are equal.
One hundred pictures were taken.
The automatic reconstruction superimposed one half on the other. After the manual intervention the pictures were stitched together properly. However a clear distortion of the point cloud is visible. Indeed the reconstructed facade shows a circular pattern instead of a straight one.
Camera: Canon EOS 550D
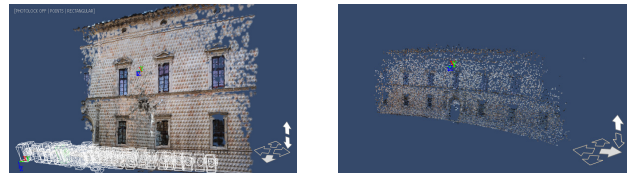
Figure(s) 10-11. Palazzo dei Diamanti, Ferrara, photographs view mode and points view mode, Automatic Reconstruction

2.5 Isolated buildings

Santa Maria in Piano is located in Loreto Aprutino (Pescara). It's original foundation dates to VIII century. Rebuilt after a fire during the XII century, it now shows the XVI century face.
Eighty six pictures were shot walking around the church.
The 3D model created by Photofly was correct and 70790 3D points were computed.
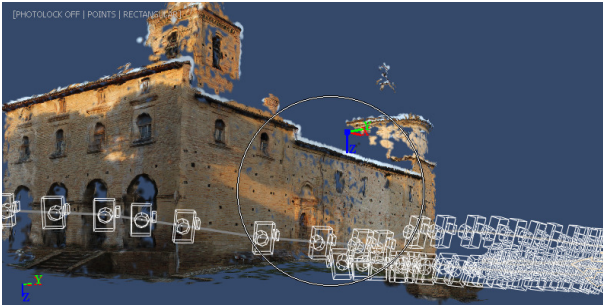Camera: Canon EOS 550D

Figure 12. Santa Maria in Piano, Loreto Aprutino (Pescara), Automatic Reconstruction, Front View
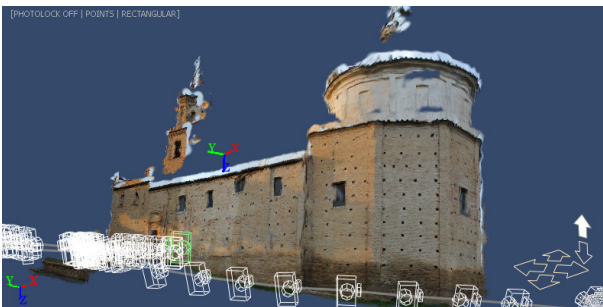

Figure 13. Santa Maria in Piano, Loreto Aprutino (Pescara), Automatic Reconstruction, Rear View

The *Estense Castle* is placed in the city centre of Ferrara. Today the castle essentially presents the appearance given to it by Girolamo da Carpi in the second half of the 16th century.
Unlike Santa Maria del Piano which shows a simple plan, the Estense Castle has a complex shape composed by different blocks.
The pictures dataset is composed by 230 photographs.
The upload of all of them together to the Photofly server resulted in a failure process, since no file was produced by the system and the Photo Scene Editor stuck. The reconstruction stopped when the amount of pictures uploaded was about half of the total amount. Even the subdivision of the dataset in smaller groups, and the uploading of each of them separately, allowed to model the building entirely.
However the half photographs dataset, computed by Photofly engine, produced a correct 3D model of the castle.
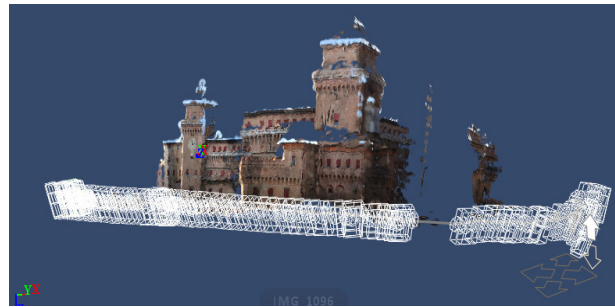Camera: Canon EOS 550D

Figure 14. Estense Castle, Ferrara, Automatic Reconstruction (50% of pictures taken)

2.6 Architectonics elements

A rib vault of the cloister of the *Archiginnasio* of Bologna was digitized in order to evaluate the potentials of Project Photofly to reconstruct single architectonic elements.
Twenty-three photographs were shot.
The model sent back by the server was properly reconstructed and the final point cloud consisted of 19950 points.
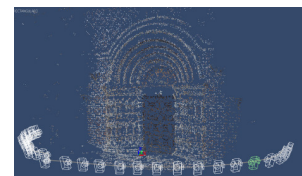Camera: Canon EOS 550D
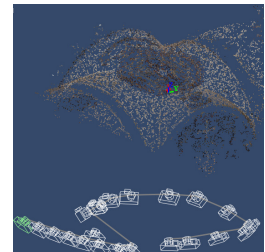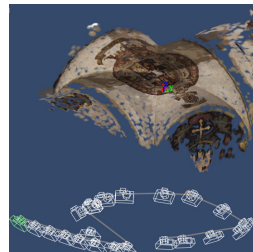
Figure(s) 15-16. *Archiginnasio* Rib Vault, Bologna, Automatic reconstruction, photographs view mode and points view mode

S. Maria of Collemaggio is a large medieval church in L'Aquila. Although the building was heavily affected by the 2009 earthquake, primarily in the rear section, the façade and the main portal didn't report any damage.
Thirty pictures of the main portal were taken. The Photofly automatic reconstruction algorithm had returned a correct model where all the photographs were properly stitched together. A point cloud of 34631 points was produced.
Camera: Canon EOS 550D

Figure(s) 17-18. Santa Maria of Collemaggio, l'Aquila, photographs view mode and automatic points view model

2.7 Sculptural objects

Thirty-one pictures of a male marble bust, placed in the cloister of the *Archiginnasio* in Bologna, were taken.
After the upload, the photographs were stitched together correctly producing a point cloud of 30966 points.
Camera: Canon EOS 550D
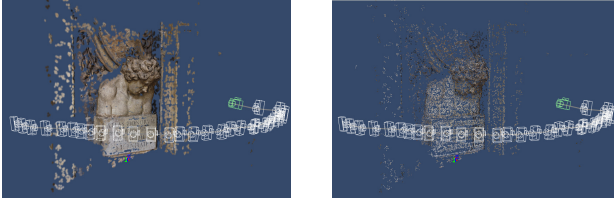
Figure(s) 19-20. *Archiginnasio* Marble Bust, Bologna, Automatic reconstruction, photographs view mode and points view mode

Another statue was used as beta test. The statue represents an Hippogriph. It is placed in front of the cathedral of the city of Ferrara.
Ninety four photographs were shot and processed trough Photofly. The pictures were automatically stitched together in a point cloud of 52846 points.
The large number of photos is due to continuous shooting setting enabled on the camera.
Camera: Canon EOS 550D
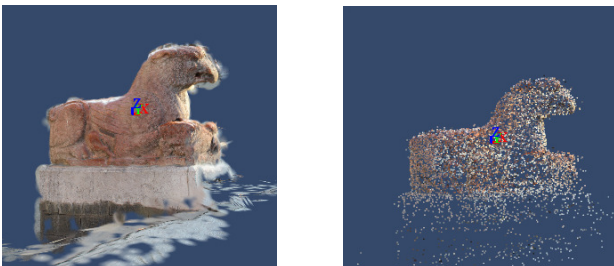
Figure(s) 21-22. Marble Hippogriph statue, Ferrara, Automatic Reconstruction, photographs view mode and points view mode

3. ASSESSMENTS AND EVALUATION

The main object of the paper was to evaluate the Project Photofly in order to understand goals, and for which objects it is suitable, and to highlight its strengths and weaknesses.
Some case studies were tested for these purposes.
Different sites and objects were digitized according with different scales:

- Open spaces (city square)
- Building facades
- Isolated buildings
- Architectonics elements
- Sculptural objects

3.1 Open Spaces (City Squares): results

All the 3D reconstructions of city squares were unsuccessful. Photofly seems to have problems when dealing with such open spaces. Manual intervention was always necessary to achieve a correct result.
In the Navona Square dataset the automatic process failed probably due to some obstructions, represented by the "Fountain of the four rivers" by Bernini, which at that time was surrounded by flat white panels. However, after marking points on building behind them, the server was able to model the entire scene correctly.
Also for the Campidoglio Square dataset the automatic process failed. In this case the wrong reconstruction was possibly caused by the equality of the two buildings on the sides of the square facing each other. One building was indeed placed wrongly behind its twin and with a bigger scale.

The last case study for this typology was the Maggiore Square in Bologna. Although the square was free of obstruction and there were no similar buildings the 3D reconstruction was incorrect.

3.2 Building Facades: results

One test based on facades (Palazzo d'Accursio, Bologna) was successful. Photofly server correctly reconstructed the model. It seems that it has no problems dealing with surface with no specular sides when the camera moves in just one direction.
However for Palazzo dei Diamanti, where the façade has mirror sides divided by the main entrance, the resulting 3D model was wrong. The two sides were superimposed. For this building also a distortion of the point cloud was noticed.

3.3 Isolated Buildings: results

Isolated Buildings, with no obstructions occluding the camera field of view, and without other structures leaning among them, were chosen.
Both the case studies showed that Photofly works correctly with isolated buildings (simple and complex shapes).
However for the reconstruction of the Estense Castle it wasn't possible to model the entire site since the engine wasn't able to process all the pictures taken on field. The upload was tried different times, even subdividing the dataset in smaller groups, always with the lack of results.
Only half of the photographs database was processed. Despite this, the model obtained was correct.
The camera path around both of them was circular (360 degrees) and the photographs were taken from different distances due to specific environmental conditions.

3.4 Architectonics Elements and Sculptural Objects: results

Simple objects, with no obstructions and easy access, were chosen for this test.
All the object digitized were modeled precisely exploiting the Photofly automatic pipeline.

For two of them (portal and bust), among four, the camera path was semicircular, covering 180 degrees in front of them. For the hippograph the camera moved 360 degrees around it, and for the forth one (rib vault) the camera path was moving along its four edges. All pictures were stitched together and none of them was missed or added manually.

4. CONCLUSIONS

Based on this study and due to lack of other documentation, besides the AutodeskLab forum, the following strengths and weakness can be highlighted.

Project Photofly is able to create quickly models (point cloud) of objects, architectonic elements, surfaces and isolated buildings, where pictures are taken without any overwhelming obstruction or elements like moving objects, which can represent noise in the camera field of view.

It shows problems to reconstruct complex and articulated spaces, as highlighted in the square typology.

It experiences problems when dealing with large dataset and when the scene includes mirror elements.

In these environments other techniques, like Laser Scanning, are preferable.

At this point Project Photofly is still a technology preview. It is not a mature technology. At this stage, failures are as interesting as successes.

Acknowledgements

References from Web-site

http://labs.autodesk.com/utilities/photo_scene_editor/

http://forums.autodesk.com/t5/Project-Photofly/bd-p/507

http://homes.esat.kuleuven.be/~visit3d/webservice/v2/index.php

http://phototour.cs.washington.edu/bundler/

http://profs.sci.univr.it/~fusiello/demo/samantha/

Reference from Literature

Vergauwen, M., Van Gool, L.J, 2006. Web-based reconstruction service. *Machine Vision and Applications*, 17(6), pp. 411-426.

Furukawa Y., Ponce J., 2007. Accurate, Dense, and Robust Multi-View Stereopsis, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*

Snavely, N., Seitz, S.M., Szelinski, R., 2008. Modelling the world from internet photo collections. *IJCV*, 80(2), pp. 189-210.

Farenzena, M., Fusiello, A., Gherardi, R., 2009. Structure and Motion pipeline on a hierchical cluster tree. Proc. *IEEE 3DIM*, Kyoto, Japan.

Barazzetti L., Remondino F., Scaioni M., 2010. Extraction Of Accurate Tie Points For Automated Pose Estimation Of Close-Range Blocks, *IAPRS*, Vol. XXXVIII, Part 3A