The Langweil model of Prague - a challenge for state-of-the-art 3D reconstruction techniques

D. Sedlacek[†1] and J. Zara[1]

[1]Czech Technical University in Prague, Faculty of Electrical Engineering

Abstract

The challenges for current state-of-the-art 3D reconstruction algorithms covered in the Langweil model of Prague data set are summarized in this poster. The model properties are presented together with problems which bring.

Categories and Subject Descriptors (according to ACM CCS): I.4.5 [Image Processing And Computer Vision]: Reconstruction—I.3.3 [Computer Graphics]: Digitizing and scanning—I.3.7 [Computer Graphics]: Virtual reality—I.3.8 [Computer Graphics]: Applications—

1. Introduction

The Langweil model of Prague is a historical coloured paper model of the Prague city center, see Figure 1. The model itself was created in the first half of 19^{th} century and is placed in the City of Prague museum http://www.muzeumprahy.cz. The project of model digitisation started in 2006 and was finished in 2010. Almost 300 000 photographs was captured during the data acquisition process. We have prepared a selected part of the acquired data set for scientific use with the permission of City of Prague museum. The Langweil model data set is available at http://dcgi.felk.cvut.cz/langweil3d after registration.

Figure 1: *The 3D reconstruction of the Langweil model of Prague, 1/8 approximately.*

The main challenges from the point of computer vision and graphics view are presented in this poster. The poster shows that the available data set is unique in many aspects and that it should become one of sets used for verification of algorithms of different kinds.

This poster is organised as follows: firstly challenging tasks for current state-of-the-art algorithms are presented, followed by recapitulation of data set characteristics helpful for model reconstruction. The poster is closed by summarization of our own results.

2. Challenges

In this section we highlight those model properties that represent serious problems and challenges for current reconstruction methods.

Occlusions: The Prague centre, especially the parts of the old Jewish town are full of narrow streets. Due to limitation of taking photos only from over of the model some model parts are hidden on most photographs. The wall parts under the overhang of the roof are hidden completely. The texture extraction algorithms need to extract a lot of small parts and fuse them together into one big texture, the completely hidden parts need to be extrapolated from surrounding images. Similarly, the geometry extraction algorithms for example need to predict the line where the wall collide with the roof.

Low depth of field: Several problems arises from used macro shift-lens. The depth of field is low (several centimetres) and the blurred part of the image is not parallel with camera chip, but with model ground. This problem makes

† david.sedlacek@fel.cvut.cz

the camera calibration task more encouraging and also the final model textures need to be fused from more photos.

Repetitive textures: The roofs, streets and other objects with similar appearance are built up from one textured sheet of paper which was slashed on pieces of appropriate size. This results in high amount of repetitive textures which brings challenge for feature detection algorithms, descriptors and pairing algorithms.

High resolution photos: Each of 300 000 photos is a high resolution 16Mpx photo with 16bit/channel color depth. This requires the fastest and most robust algorithms for feature detection, texture extraction and camera calibration. For example the detection of SIFT features on one set with 7000 photos took approximately 34 hours and consequential features pairing was 3 days long on one CPU. Similarly the texture extraction process.
The high resolution input photos promises another challenging subject. The final texture size is for the pieces of standard size (7000 photos) around 500MB in PNG format and 1Gpx which is hard to display on current graphic hardware.

Large photo-set: Each of 57 model parts is captured separately. The smallest model part contains almost 400 photos, while the biggest more than 14 000, which brings new challenges for camera calibration algorithms. Also the reconstruction process and texture extraction process need to be properly separable to work with those enormous data. The parts provided to scientific community are built up from photo sets of the following sizes: 386, 980, 5938.

Details less than 1mm: One of the most interesting objects like statues or painted walls are originally created in the model in details less than 1mm. The capturing method was chosen to take these details even at low observing angle. The photo resolution at these cases is still 20px/mm. The geometry reconstruction tools should be able to reconstruct even those small details, in order to get maximum information from the photos.

Uniqueness: Each model part is unique in the geometry sense and walls texture, too. For this reason, the reconstruction using grammars or model fitting is not straightforward and will need special process customizations.

3. Aids and benefits

Here we present facts which are helpful for model reconstruction and make some of the problems mentioned above solvable and little bit easier.

Approximately known camera positions & data organization: All data were acquired robotically. For each photo an approximate camera position (impulses of robot arm motors) and viewing direction (viewing points in the model coordinate system) are known. The photos are ordered by viewing angles and viewing points. The approximate position is helpful for feature pairing where we can predict the possibility of observing feature in surrounding photos. Similarly with the camera calibration algorithms where we can pre-process photos to use structure from motion [HZ04] and to enable separability in other algorithms.

Ground truth: The input RAW photos are provided together with our model reconstruction in VRML format with PNG textures. The model is reconstructed with 1mm precision in the texture, while the geometry precision is higher then 0.5mm. Each model part was manually controlled and repaired, thus our reconstruction can be considered as ground truth which is necessary for algorithm validation.

Various reconstruction data sets: The provided data sets do not cover the whole Langweil model, but especially selected three parts. The parts were chosen with respect to their size and complexity.

Part 9: small, compact piece. It is suitable for study of 3D reconstruction algorithms. Terrain: **flat**.
Part 8: more exacting piece, with streets and complex geometry. Dealing with current state-of-the-art in 3D reconstruction. Terrain: **flat**.
Part 27: this piece contains a lot of occlusions caused by terrain variations and a huge number of trees. Suitable for the most robust algorithms testing. Terrain: **hilly, a lot of rough surfaces**.

4. Our results

The complete Langweil model reconstruction took almost 3 years from data acquisition to final verification. There was a lot of human work in the geometry reconstruction process and in the reparation of errors caused in the texture extraction process. The reconstruction of piece 9 took 2 weeks of 5 operators, while the pieces 8 and 27 took 6 and 4 weeks respectively.

5. Conclusions

We have presented challenges raised from the Langweil model dataset for current algorithms from the fields of computer vision and graphics. We believe, that providing these data moves forward both research fields. Download our data and prove that your algorithms are the best.

Acknowledgement

This work has been partially supported by MSMT under the research program MSM 6840770014, the research program LC-06008 (Center for Computer Graphics), and the Grant agency of the CTU Prague, grant No. SGS10/-291/OHK3/3T/13. The authors thank to the City of Prague Museum for the special data provided.

References

[HZ04] HARTLEY R., ZISSERMAN A.: *Multiple View Geometry in Computer Vision*. Cambridge University Press, March 2004.