3D Rekonstruktion des Langweil-Modell von Prag: Die Datenerfassung

3D Reconstruction of Langweil's Model of Prague: Data Acquisition

Vladimír Smutný, Petr Prášek, Petr Palatka, Tomáš Pajdla

Center for Machine Perception, Czech Technical University in Prague
Karlovo nám. 13, 121 35 Prague 2, Czech Republic
Tel: +420 224 357 280, Fax: +420 224 357 385
E-mail: {smutny,pajdla}@fel.cvut.cz, http://cmp.felk.cvut.cz

Neovision s.r.o.
Barrandova 409, 143 00 Prague 4- Modřany, Czech Republic
E-mail: {prasek,palatka}@neovision.cz, http://www.neovision.cz

Zusammenfassung:

Das Langweil-Modell von Prag wurde zwischen 1826 und 1836 von Antonin Langweil hergestellt. Das Modell stellt Prag während der ersten Hälfte des 19. Jahrhunderts dar, es schließt die Judenstadt ein, die bis Ende des 19. Jahrhunderts existierte. Das Modell ist aus Papierkarton gefertigt und mit einer außerordentlichen Realitätsnähe dekoriert. Ungefähr 2 500 Häuser bedecken im Maßstab 1:480 ca. 15 m^2.

Das Langweil-Modell ist das Hauptausstellungsstück im Stadtmuseum Prag. Es befindet sich unter Schutzatmosphäre in einer geschlossenen Glaskiste und wird in einem Raum mit schwacher Beleuchtung präsentiert. Viele wunderschönen Einzelheiten der Dekoration sind mit bloßem Auge nicht sichtbar.

Dieses Referat stellt die Technik vor, mit der 250 000 detaillierte Digitalfotos des Modells erfasst wurden, um daraus das 3D-Modell zu rekonstruieren. Das detaillierte 3D-Digitalmodell wird von Historikern, Denkmalschützern und Stadtplanern genutzt, außerdem wird es Museumsbesuchern und in vereinfachter Version im Internet gezeigt.

Abstract:

The Langweil's model of Prague was made by Antonin Langweil between 1826 and 1836. It captures Prague of the first half of 19th century including the Prague Jewish quarter which has disappeared at the end of 19th century. The model has been made of cardboard and it was hand decorated to capture the city in high realism. Approximately 2 500 houses in scale 1:480 span the area close to 15 m^2.

The Langweil's model is the main exhibit of the Prague City Museum. The model is normally placed in a closed glass box with a protective atmosphere and presented under low illumination only. Many beautiful details of its decoration are not visible by bare eye.

This paper presents the technology of capturing the Langweil's model in 250 000 detailed digital photographs in order to generate its digital three-dimensional model. The detailed digital 3D model will be used by historians, preservationists, urbanists, will be presented to museum visitors and in a simplified version on Internet.

1 Introduction

The Prague City Museum is proud to present historical model of Prague made by Antonin Langweil. The model is certainly the centerpiece of the exhibition occupying the main exhibition room. It attracts visitors from all over the world. The quality, detail, technical, and historical accuracy of the model makes it appealing not only for tourists and Praguers, but also for historians, preservationist, and city planners. The model is preserved by a display case with a controlled illumination and atmosphere. Unfortunately the same display case prevents visitors to see many houses, details, and landmarks as they are obstructed by other parts of the model. If a detail view of peripheral parts and an overall view are sufficiently impressive for visitors, the specialists miss a view to the central, most valuable parts of the model, representing the heart of the historical city.

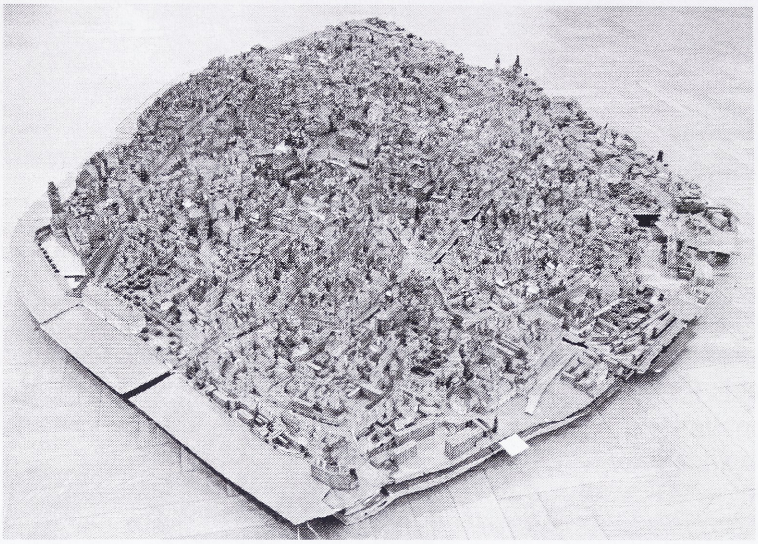


Fig. 1 Old Town on the Langweil model, about one half of the model.

As the computer vision technology progresses, the museum took the chance of regular inspection which happen every 15 years and ordered a project for preparation of a three dimensional computer model of the exhibit. Neovision together with Center for Machine Perception was selected to deliver first two phases of the project, i.e. phase 1: image acquisition and phase 2: 3-D computer model reconstruction from images. To eliminate a possible confusion in this paper, the original Langweil model will be denoted as model or exhibit, the final result of the second phase, 3D computer model will be called computer model. The final deliverable of the project, called Digital Langweil Model contains further features as a database of streets, houses, etc. and a special viewer of the model.

The model, made of cardboard, originated between 1826 and 1836. The model is made in the scale 1:480, with vertical size of the houses slightly exaggerated. The model occupies roughly 15 m^2. It covers Old Town, Lesser Town, Hradcany that is almost whole Prague at that time except of the New Town. The model is divided into about 50 main parts, which cannot be further decomposed. Some parts are not made by Langweil, e.g. Vltava River. During Langweil's life only parts of the model were exhibited and never the complete setup.

The model contains almost 2000 addresses but the number of individual buildings is probably around 5000. Simplest houses contain from several flat parts: facades and roof planes, made of painted cardboards and chimneys. Tops of the towers are usually made of lathed pieces of wood. Most of the surfaces (probably whole original model) are finished by a matt painting with detailed drawings of windows, lamps, etc. Roofs were painted by shellac during restoration, so they are quite glossy. The trees and woods are modeled by wire and hair, they are not original but they have been added during a restoration later.

2 Task Definition

3D computer model was defined in the project by two quantitative parameters. The resolution of the texture in the computer model shall be better than 0.1 mm, which we interpreted as the pixel in the resulting texture shall not correspond to the exhibit surface larger than 0.1 by 0.1 mm. The second parameter, spatial resolution 1 mm, was explained to us, that the exhibit plastic features larger than 1 mm shall be three-dimensionally reflected in the computer model.

The passive stereo was chosen as the method for reconstruction of computer model. Although some experiments were done with competing technologies like laser triangulation or off the shelf data acquisition systems like Konica-Minolta VI-9i, the results were not satisfying. The resolution of images was poor for Konica-Minolta system. Passive stereo is very mild to the exhibit and could be easily adjusted to the proper resolution.

As passive stereo was chosen for second phase of the project, image (or generally data) acquisition was tuned to satisfy its needs. The first requirement is to cover each (or practically almost every) patch of the model surface by at least two distinct focused images, which will allow reconstruction in later phase. The resolution should satisfy the texture resolution condition.

3 Image Acquisition Method

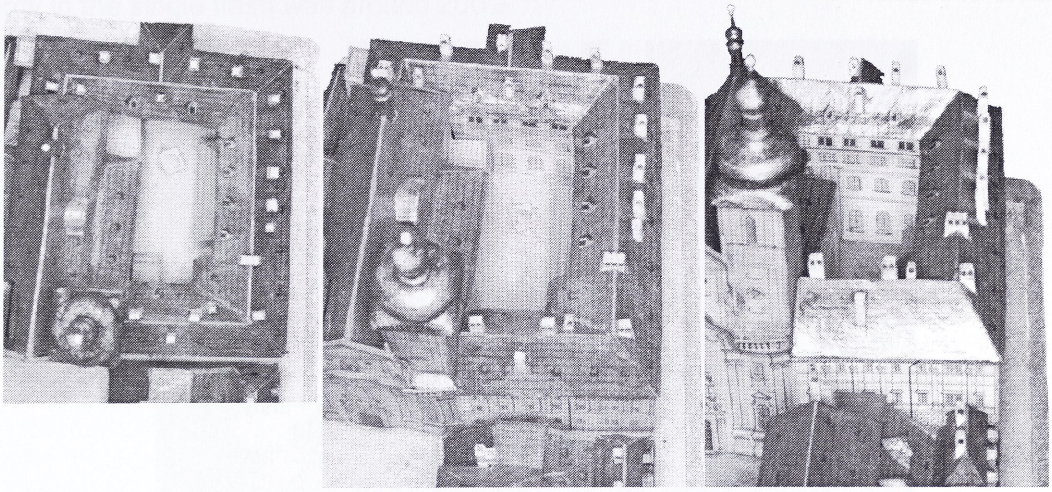


Fig. 2 Model is captured under three elevations, top view, 20°, 45° from the top view.

To satisfy the condition of visibility of each patch at least in two images, we designed the following general plan for image acquisition:

- The model will be photographed from top in regular grid satisfying condition to see into narrow streets. These images will cover mainly roofs, street and yard pavement and other horizontal structures. Some lower parts of the facades, especially in narrow streets will be also covered in periphery of the image due to the wide view angle.
- The model will be photographed with two other elevations, 20° and 45° from the top view, mainly to cover facades in great detail. There are no horizontal views planned except for St. Veit Cathedral. The vertical structures are thus reconstructed from 45° views mainly. That means that resolution along vertical directions should satisfy texture resolution condition.
- The elevated views will generally be photographed from 8 azimuths, separated by 45° degrees. Each point in the regular grid will thus be covered by 17 images: top, 8 azimuths times 2 elevations.
- As the depth of field of the camera is limited, whole 3D structure of the scene shall be covered by grid points.

The parameters of the capturing were little bit more complicated, top views were planned in a grid 30x30 mm, the elevated views were captured in alternative grid 45x45 mm. The spacing between horizontal layers, given by camera depth of field was 40 mm.

4 Capturing Setup Design

4.1 Camera

The camera, its quality, reliability, and parameters are essential for quality of the output of the first phase of the project. Several criteria were considered during camera selection:

- High dynamic range. We soon found that we will need the high bit depth, as we should be able to capture both specular reflections from glossy roofs as well as dark places in the narrow yards, well hidden to our illumination. We could make it through clever composition of several images or through the equipment which captures images with high dynamic range instantly. When we considered time of capturing, complexity of postprocessing, possibility to capture during camera motion, we found that a professional digital camera back is the right solution.

- Time of capturing. We have to find a camera which can online send the image to a data storage, could be triggered and controlled from a computer, and do all that fast.
- Spatial resolution.
- Reliability. We planned to capture several hundred thousands images. For example the best professional single lens reflex cameras from Canon promise 200 000 images per one lifetime of shutter.

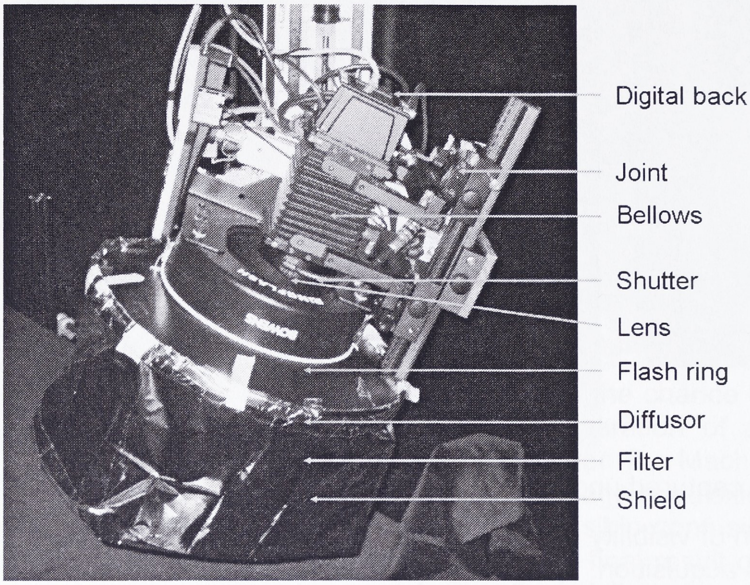


Fig. 3 Camera setup.

The criteria should not be evaluated separately, but altogether. For example industrial cameras will have no wear on shutter, they can make more images to satisfy spatial resolution on the exhibit surface, but they have rarely a high dynamic range. Also a spatial calibration of images for a good reconstruction will be difficult with an industrial camera resolution.

Finally we have chosen PhaseOne digital camera back P20. Its resolution is 4096x4096 pixels with Bayer pattern and 13 bits per pixel. The frame frequency with the transfer to the computer is 2.5 seconds per frame. Size of the chip is 36x36 mm.

4.2 Lens

To get a proper resolution when a façade is viewed under angle 20°, the field of view chosen was around 150x150 mm and magnification around 4. This leads us to a macro lens and we have chosen the Rodenstock Apo-Macro-Sironar Digital with the focal length 120 mm and aperture 5.6. The depth of field was a crucial parameter for both selecting of equipment and for planning. Although we considered various chip sizes and various field of views, it was impossible to enlarge the depth of field behind approx 40 mm when the texture resolution was given. The aperture 22 was used.

4.3 Camera Body

We needed to have either vertical or horizontal plane of focus depending on the scene even when elevated images are captured. This leaded us to the bellows body Rollei X-Act 2. Scheimpflug principle allows as to rotate plane of focus as needed. Actually the turn knob on the bellows body was a single parameter on the whole setup, which has to be adjusted manually and whose position cannot be read by computer. As a shutter we have chosen Rollei Electronic Shutter. To make the life as long as possible, the shutter time 1/30 was used.

4.4 Illumination

We were looking for a light source which could be computer triggered, which produces short flash, so the sharp images could be captured during camera motion, and which produces a strong, white, controllable illumination. Also geometrical distribution of the light was important. We have selected Bowens ring flash light with an opal glass diffuser and the Quadx 3000 generator. The flash was further shielded by an IR filter and an UV filter. The system was further modified by a shield made of aluminium foil, which both reflected the diffused light back to the photographed part and protected its surroundings. UV filter and flash unit with planned parameters were also evaluated by an independent museum advisor to approve, that the exhibit will not be damaged by incident light. The energy in the single flash was around 200 J.

4.5 Robot

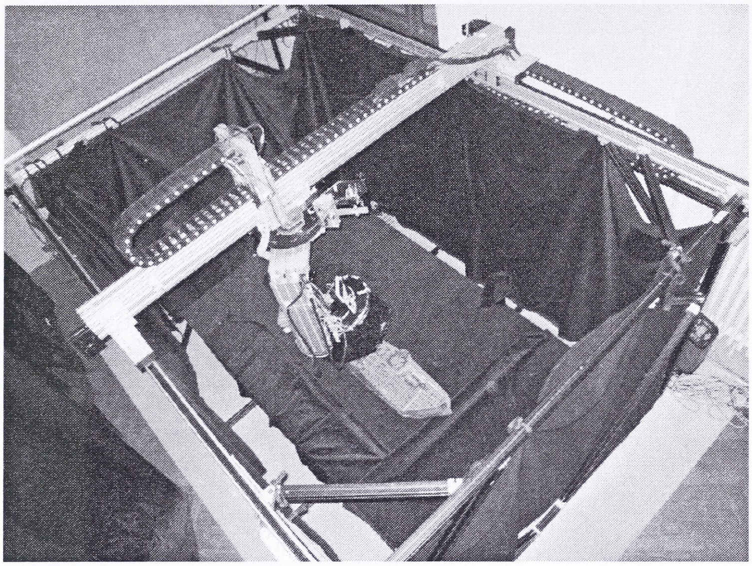


Fig. 4 Acquisition robot, surrounded by black velvet, part of the model is currently processed.

We planned to capture around 300 000 images from carefully defined positions. This could hardly be achieved by a manual capturing of photographs. We designed and build a Cartesian robot with 5 degrees of freedom. The freedoms of robot mostly reflected the required parameters of the camera, *xy* table moved in regular grid, *z* moved up and down correcting for depth of field, azimuth and elevation degrees oriented the camera. The robot was controlled by MARS 8, a microprocessor control unit.

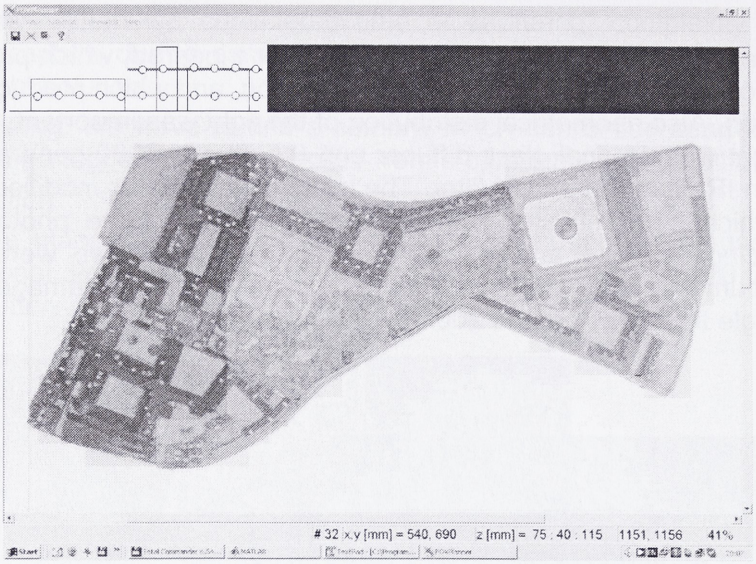
4.6 Workshop

The workshop was located in the museum building to prevent any stressful manipulations with the exhibit. The room was separated by dark plastic foils into several cubicles.

5 Data Acquisition

5.1 Planning

The model part was prepared for capturing in the first workplace. First it was photographed from the top to get data for further image acquisition planning. Then the plan was done manually using SW tools and by measuring the exhibits heights manually. Then the plan and model part was moved to next step.



32 x,y [mm] = 540, 690   z [mm] = 75 : 40 : 115   1151, 1156   41%

Fig. 5 Screenshot of the SW used for view planning.

5.2 Capturing

The model was placed under the robot in the predefined position. It was placed onto a black velvet and also the surroundings was covered by black velvet. The loop running through all planned positions was run, capturing an image in each planned position. The cubicle with robot was separated by black foil to minimize the impact of the unpleasant flash light to the operator.

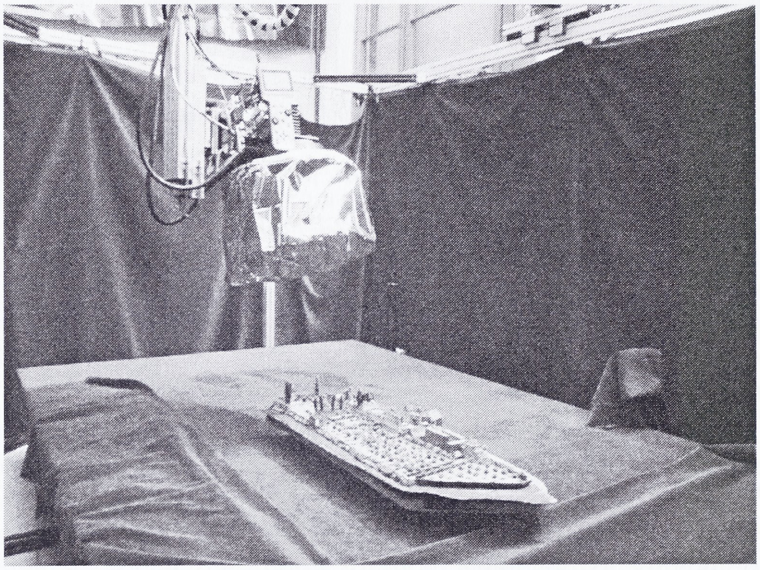


Fig. 6 Acquisition setup during capturing.

5.3 Evaluation, archival

The captured images were checked for certain properties and if correct, they were archived and sent to the customer. Bad images were reported to the operator and he manually requeued them for a later recapturing.

5.4 Software System Design

There was a lot of SW implemented for this part of project. The most complicated was a user interface, which allowed planning of views. It used orthophoto of the model part and overlaid it with planned graphics depicting planned views in different layers.

The software for the particular robot (direct and inverse kinematics, calibration) was implemented as well as control loops for positioning the robot, image capture, and image storing. The images themselves were captured from the digital back by CaptureOne SW delivered with the camera. The interface was written between CaptureOne and main loop scripts running in Matlab.

Further SW inspected images histograms and made 1024x1024 thumbnails. All steps of individual images were tracked in a database, containing for each image all its parameters (planned and captured positions, history with time stamps etc.) The database is an essential part of the deliverables of the first phase and it was invaluable in the capturing process. It will be also heavily used in further project phases.

6 Conclusions

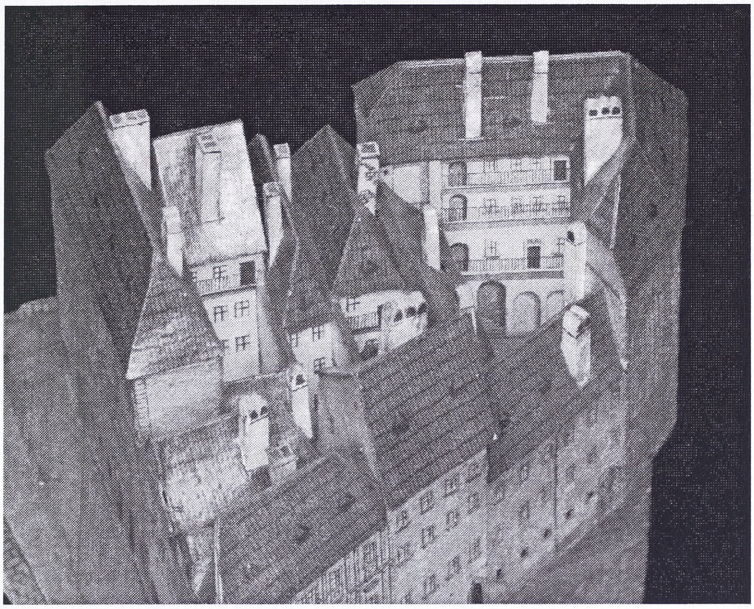


Fig. 7 Captured image.

Almost 280 000 images of the Langweil model were captured. They occupy in the compressed form around 5 TB, roughly 20 MB per image. The images have very good quality.

During capturing number of technical problems occurred. Besides transient problems one has to calculate with a limited lifetime of the Rollei Electronic Shutter. Our piece was repaired after 130 000 and again after 107 000 images. The first flash generator regularly got overheated when the frequency of capturing was under 3 seconds per frame, but Bowens send us a new version of the generator and this unit worked completely reliable. No flash bulb was burned during capturing. Originally there were problems with flash synchronizations, both an infra system (noname) and a radio system (Pulsar Radio Trigger by Bowens) were not working reliably. The infra system was missing triggers due to occlusions and black surrounding, the radio system was triggered by an unidentified radio noise sometimes. The cable solved finally all those problems. PhaseOne digital back worked without problems, but it has to be forced externally to recalibrate itself during robot idle motion and to precede regular automatic recalibration. There were some problems with robots, because the actuators were undersized and got overheated and eventually burned. We finally reached the performance 2.7 seconds per frame during continuous motion of the camera and 15 000 images per 24 hours of the aggregated performance.

We supposed that second phase, that is reconstruction of the 3D computer model from the images will be even bigger challenge than to capture so many images in about two month including debugging of the capturing prototype.