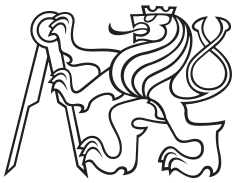
Master Thesis



Czech
Technical
University
in Prague

F3

Faculty of Electrical Engineering
Department of Computer Graphics and Interaction

Tactile models for older adults with vision impairments

Dina Chernova

Supervisor: Ing. Miroslav Macík, Ph.D.
Field of study: Open Informatics
Subfield: Computer Graphics
August 2020

ZADÁNÍ DIPLOMOVÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Chernova** Jméno: **Dina** Osobní číslo: **420050**

Fakulta/ústav: **Fakulta elektrotechnická**

Zadávající katedra/ústav: **Katedra počítačové grafiky a interakce**

Studijní program: **Otevřená informatika**

Studijní obor: **Počítačová grafika**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Hmatové modely pro seniory se zrakovým postižením

Název diplomové práce anglicky:

Tactile models for older adults with vision impairments

Pokyny pro vypracování:

Older adults with severe vision impairments often live in residential care institutions that provide specialized care. This user group has specific needs, abilities, and preferences, primarily from the perspective of orientation in space, interaction with technologies, and free-time activities. Analyze the outcomes of current research on the spatial orientation of severely visually impaired with a particular focus on older adults. Follow up the results of research conducted at the DCGI [1,2] and on solution described in the patent titled 'Three-dimensional model of a building primarily for visually impaired older adults' [3]. Specify information to be represented by geographical models of the interior as input. Design and implement a method to transform geographical models of the interior environment to (interactive) tactile models usable by visually impaired older adults. Follow the User-centered design method [4]. Evaluate the outcomes with participants representing the target user audience.

Seznam doporučené literatury:

[1] M. Macik, I. Maly, J. Balata, and Z. Mikovec. How can ict help the visually impaired older adults in residential care institutions: The everyday needs survey. In Cognitive Infocommunications (CogInfoCom), 2017 8th IEEE International Conference on, pages 000157–000164. IEEE, 2017.
[2] Gintner, V., Macik, M. and Mikovec, Z., 2019, September. Perception of Tactile Symbols by Visually Impaired Older Adults. In IFIP Conference on Human-Computer Interaction (pp. 325-334). Springer, Cham.
[3] Macík M., Malý I., Balata J., Míkovec Z, Treml L. Trojrozměrný model budovy určený především pro zrakově postižené uživatele, Aug 2019. Patent. Dostupné z: https://isdv.upv.cz/doc/FullFiles/Patents/FullDocuments/307/307912.pdf
[4] DIS, ISO. (2009). 9241-210: 2010. Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems.

Jméno a pracoviště vedoucí(ho) diplomové práce:

Ing. Miroslav Macík, Ph.D., Katedra počítačové grafiky a interakce

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **03.03.2020** Termín odevzdání diplomové práce: **14.08.2020**

Platnost zadání diplomové práce: **19.02.2022**

Ing. Miroslav Macík, Ph.D. podpis vedoucí(ho) práce	podpis vedoucí(ho) ústavu/katedry	prof. Mgr. Petr Páta, Ph.D. podpis děkana(ky)

 © ČVUT v Praze, Design: ČVUT v Praze, VIC

III. PŘEVZETÍ ZADÁNÍ

Diplomantka bere na vědomí, že je povinna vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

. _____
Datum převzetí zadání

_____
Podpis studentky

 © ČVUT v Praze, Design: ČVUT v Praze, VIC

*

Acknowledgements

I would like to thank Ing. Miroslav Macík, Ph.D., my supervisor, for his valuable advice, time and patience. Thanks to Mgr. Lukáš Treml for important insights. I thank my beloved Alessandro Alei for his support in the hardest times. Then, I owe big thanks to my family that was supporting me during my studies. Finally, I am endlessly thankful to the Czech Technical University for providing me with the opportunity to study my bachelor's and master's degree in this fine institution.

Thank you.

Declaration

I hereby declare that the presented work was made independently, and that I have listed all the sources of information used within it, in accordance with the methodical instructions for observing the ethical principles in the preparation of the university thesis.

Prague, August 2020

Abstract

Older adults with severe vision impairments often live in residential care institutions that provide specialized care. This user group has specific needs, abilities, and preferences, primarily from the perspective of orientation in space, interaction with technologies, and free-time activities. In this thesis I have proposed a pipeline for reconstruction of a building from images that satisfies the requirements of the patent 'Three-dimensional model of a building primarily for visually impaired older adults' [33] for creation of a tactile 3D map for the visually impaired older adults. The pipeline was designed, implemented and tested with a target user group using the User-centered design method. The results of the evaluation show that there's still a lot to be improved. Furthermore, I have designed and 3D printed an exterior part of the Home Palata building. The printed 3D model was then glued to a wooden model created by Ing. Miroslav Macík, Ph.D. as part of the implementation of the patent mentioned before. I have evaluated the model that I have created with a visually impaired expert. He has shared a lot of valuable insights to improve the model.

Keywords: visually impaired, blind, low vision, older adults, residential care institution, palata, navigation, orientation, indoor, outdoor, inside, outside, 3D model, 3D printed, 3D printer, external building part, facade, 3D building, tactile map, interactive

Supervisor: Ing. Miroslav Macík, Ph.D.

Abstrakt

Starší lidé s těžkým poškozením zraku často žijí v ústavních zařízeních poskytujících specializovanou péči. Tato skupina uživatelů má specifické potřeby, schopnosti a preference, především z hlediska orientace v prostoru, interakce s technologiemi a volnočasových aktivit. V této práci jsem navrhla pipeline pro rekonstrukci budovy z obrazů, které splňují požadavky patentu „Trojrozměrný model budovy primárně pro zrakově postižené starší dospělé“ pro vytvoření taktilní 3D mapy pro zrakově postižené starší dospělé. Pipeline byla navržena, implementována a testována s cílovou skupinou uživatelů pomocí metody User-centered design. Výsledky hodnocení ukazují, že je stále co zlepšovat. Dále jsem navrhla a vytiskla vnější část budovy domova Palata. Tištěný 3D model byl poté přilepen na dřevěný model vytvořený vedoucím mé práce v rámci implementace výše uvedeného patentu. Vyhodnotila jsem model, který jsem vytvořila, s nevidomým odborníkem. Odorník sdílil spoustu cenných poznatků, jak model vylepšit.

Klíčová slova: nevidomé, slepý, slabý zrak, starší dospělí, palata, navigace, orientace, vnitřní, venkovní, 3D model, 3D tisk, 3D tiskárna, vnější stavební část, fasáda, 3D budova, hmatová mapa, interaktivní

Překlad názvu: Interaktivní hmatové plány pro seniory se zrakovou vadou

Contents

1 Introduction **3**

1.1 The Goals Of This Thesis 4

1.2 Methods and terms 5

1.2.1 User-Centered Design 5

1.2.2 Usability 6

2 Analysis **7**

2.1 Statistics On Visual Impairment . 7

2.2 Perception Of Blind And Visually Impaired 8

2.2.1 Orientation And Navigation . . 9

2.2.2 Auditory Senses 10

2.2.3 Tactile Senses 11

2.3 Tactile Maps As Means Of Spacial Cognition 12

2.3.1 What Is A Map? 12

2.3.2 Tactile Maps 13

2.3.3 3D Printed Tactile Maps 15

2.3.4 Interactive Tactile Maps 16

2.4 Home Palata For Visually Impaired 18

2.4.1 The Institution 18

2.4.2 Navigation And Orientation Of Clients In Palata 18

2.5 3D printable Palata model 19

2.6 The Idea 19

2.6.1 Interactive Tactile Floor Plans 20

2.6.2 Exterior Part of Palata 22

3 3D Building Reconstruction From Images **25**

3.1 Image Association 25

3.1.1 SIFT 26

3.1.2 Other Algorithms 27

3.2 Multi-View Geometry 28

3.3 Homography 31

3.4 Non-linear Estimation Algorithms 33

3.5 3D model Reconstruction Pipeline 34

3.5.1 Correspondence Search 34

3.5.2 Sparse Reconstruction 34

3.5.3 Dense Reconstruction 35

4 Segmentation Of Building Facades 37

4.1 2D Shape Grammar 38

4.2 Learning With Randomized Forests 39

4.3 Facade Energy 40

4.4 Optimization Using Random Walk Algorithm 41

5 The Pipeline Of 3D Building Exterior Creation 43

5.1 Input And Output 45

5.2 Acquiring Building Images 45

5.3 Feature Matching & Cutting Off Redundant Part 46

5.4 3D Model Reconstruction 46

5.5 Acquiring Building Floor Plans . 47

5.6 Building A Simplified Model With Straight Walls 48

5.7 Acquiring Front-Side View Pictures/Facade Plans 49

5.8 Adjusting Pictures Manually ... 50

5.9 Picture Stitching 50

5.10 Manual Facade Elements Creation 51

5.11 Facade Element Detection 51

5.12 Mapping Of Facade Elements Onto The 3D Model 52

5.13 Result 52

6 Design & Implementation 53

6.1 Low-Fidelity Pipeline Prototype 54

6.2 High-Fidelity Pipeline Prototype 55

6.3 3D Model 56

6.3.1 Requirements 56

6.3.2 Measurements 57

6.3.3 Printing & Gluing The Model 59

7 Evaluation 61

7.1 Low-Fidelity Pipeline Prototype 61

7.1.1 Hierarchical Task Analysis .. 61

7.1.2 Cognitive Walkthrough 62

7.2 High-Fidelity Pipeline Prototype 67
7.2.1 The Goal 67
7.2.2 Target Audience 67
7.2.3 Used Hardware 68
7.2.4 Used Software 68
7.2.5 Initial Application State 68
7.2.6 Workplace Layout 69
7.2.7 Use Cases 69
7.2.8 Use Cases Action Sequence . . 70
7.2.9 Pre-Test Interview 71
7.2.10 Post-Test Interview 71
7.2.11 Testing Sessions 72
7.2.12 Overall Usability Testing Summary 76
7.3 3D Model 79
7.3.1 Front-side Part 79
7.3.2 Drawers 80
7.3.3 Corners 80
7.3.4 Texture 81
7.3.5 Summary 81

8 Conclusion 83

8.1 Future Work 85

Bibliography 87

A Measurements Of The 3D Model 91

B HTA diagram 93

C Cognitive Walkthrough 95

D Screener For Usability Testing 97

D.1 Questions 97
D.2 Answers 99
D.2.1 Pilot Test Participant 99
D.2.2 Participant 1 99
D.2.3 Participant 2 99
D.2.4 Participant 3 100
D.2.5 Participant 4 100
D.2.6 Participant 5 101

E Usability Testing Sessions **103**

E.1 Test 1 103

E.2 Test 2 107

E.3 Test 3 111

E.4 Test 4 114

E.5 Test 5 117

F Low-Fidelity Prototype **121**

Abbreviations 131

Figures

1.1 Representation of phases of UCD and their iterative nature, from [60] 6

2.1 Causes of visual impairment globally in percentage, from [40] ... 8

2.2 Example of tactile symbols tested by Macík et. al., from [31] 12

2.3 Cartographic Communication Model adapted from MacEachren, 1995, from [30] 13

2.4 Tactile Map Testing, from [59] .. 14

2.5 Tactile Navigation Map Symbols, from [30] 14

2.6 Thin raised line maps, from [54] 16

2.7 First picture: User is exploring the tactile map using one hand, the other hand obtains detailed information from the phone using special gestures, from [21]; Second picture: Linespace technology for touch detection, from [52] 16

2.8 Interactive 3D model proposed by Macík et. al., from [33] 17

2.9 Palata from the top view. Part B is newly built, part A is the old part, from [32] 18

2.10 A corridor in Palata. Red zoom shows the haptic marks, dashed green line contrast the navigation line, blue color marks the furniture for the resting, from [32] 19

2.11 First picture: Wooden model of Palata made by Macík following the patent [33]; Second picture: Control to regulate the height of the model. 20

2.12 Initial few designs of room buttons. In the end the middle one was chosen for the final prototype, from [15] 21

2.13 First picture: High-fidelity prototype of Palata floor interactive tactile map. *Blue line* are the passage that user should follow. *Green line* - segments of the route which play piano tones when are being touched. *Red line* are segments off the route, when touching "no-go" sound plays. Second picture: *a* - start/end point for navigation; *b* - high-contrasted hallway that is marked by the two raised lines; *c* -symbol for a room, *d* - stairs, from [15] 21

2.14 Picture of Palata taken from a drone by Macík 22

3.1 Local and global image description, from [49] 26

3.2 Four images generated by subtracting the consecutive Gaussians, from [50] 27

3.3 16×16 neighborhood around the keypoint is divided into 4×4 subblocks. Arrows define the magnitude that defines the descriptor. From [50]. 28

3.4 Pinhole camera with the local camera coordinate system defined by the axes *(X, Y, Z)*. Point X from the scene is projected onto the image plane at x through the camera projection center C. f is focal length. From [49]. 28

3.5 Lens distortion effects in the image, from [49] 30

3.6 Homography maps points from one plane to another in a planar scene, from [49] 30

3.7 The fundamental matrix maps points from one image to lines in the other image. This process describes the two-view geometry with non-stationary camera motion. From [49]. 32

3.8 Incremental sparse reconstruction pipeline, from [49]. 34

4.1 Original image, classification-based segmentation, and segmentation with procedural shape prior, from [55] . 38

4.2 A patch is dropped in the decision tree, then the tests are built randomly over the elements of the patch. Depending on the outcome of the test, the path is sent to one children or another. From [55]. . . . 40

4.3 First picture: Original data; Second picture: Label image (by hand). Each color represents a label; Third picture: Most probable label; Fourth picture: Probability for each pixel to be a window; Fifth picture: Probability for each pixel to be a wall. The values increase from red (p=0) to blue (p=1). From [55]. . . 41

5.1 The pipeline of 3D building exterior creation. All pastel purple rectangles are automated steps, the pastel green are manual. 44

5.2 An example of two pictures of Palata taken from Google Earth . . 45

5.3 Pictures of Palata with no background 46

5.4 Reconstructed 3D model of Palata 47

5.5 Second floor of Palata 48

5.6 Floor plans corresponding to each floor 48

5.7 Extruded wall of Palata 48

5.8 Front-side view picture of the front side of Palata. Picture taken from the drone. 49

5.9 Front-side view picture of the back side of Palata. Picture taken from the Google Earth. 50

5.10 Palata with extruded facade elements and an image of the facade located on the offset plane from the front side of the building 51

5.11 Resulting building 52

6.1 Low-fidelity prototype: step 2 completed page 54

6.2 Low-fidelity prototype: result page 54

6.3 High-fidelity prototype: step 2 completed page 55

6.4 High-fidelity prototype: result page 55

6.5 First picture: Squared window type 1; Second picture: Squared window type 2 57

6.6 First picture: Squared window type 3; Second picture: Squared window type 4 58

6.7 First picture: Arc window type 1; Second picture: Arc window type 2 58

6.8 Arc window type 3 58

6.9 First picture: Front door; Second picture: Back door 59

6.10 First picture: Back side; Second picture: Front side 59

6.11 First picture: Left side; Second picture: Right side 59

6.12 Example of a tile 60

7.1 Example of a new *"Step 2 completed"* page After round 2 of Cognitive Walkthrough 65

7.2 Examples of what users should upload and the number of uploaded pictures is added after round 2 of Cognitive Walkthrough 65

7.3 Each step that requires some time to complete now has a progression page after round 2 of Cognitive Walkthrough 66

7.4 First picture: A feature to edit pictures is added after round 2; Second picture: Reconstructed 3D model can now be downloaded in step 4&5 after round 2. 66

7.5 Pipeline schematic illustration . . 78

7.6 First picture: The floors are not on the same level due to drawer desks; Second picture: Sharp corners that exceed the wall 79

7.7 First picture: A hole between the tiles. The crossing of four tiles.; Second picture: A "good" corner . . 80

7.8 An example of how all corners should be printed 81

Tables

2.1 Categories of severity of visual impairment, from [14] 7

2.2 Global estimation of the amount of people visually impaired by age and the corresponding prevalence (%), from [40] 8

C.1 Cognitive Walkthrough Round 1 Task 1 95

C.2 Cognitive Walkthrough Round 1 Task 2 95

C.3 Cognitive Walkthrough Round 2 Task 1 96

C.4 Cognitive Walkthrough Round 2 Task 2 96

C.5 Cognitive Walkthrough Round 3 Task 1 96

C.6 Cognitive Walkthrough Round 3 Task 2 96

Chapter 1

Introduction

According to the report of the World Health Organization from 2010 there are 285 million people in the world diagnosed with different types of visual impairment, 39 million of which are blind. The majority - around 82% - of people suffering from vision impairment are over the age of 50. The main reasons for this are uncorrected refractive errors (43%) and cataract (33%). [40] In Europe there are over 30 million blind and partially sighted persons. Based on data from the Sample Survey of Disabled People [62] conducted by the Czech Statistical Office in 2013, there are approximately 65000 people with severe visual impairments in the Czech Republic.

Having visual impairment increases the risk of mobility disability especially for the older people. According to the paper [53], among people aged 65+ around 32% have a difficulty with walking three city blocks and 30% report difficulty walking up the staircase in the United States. When it comes to the people that have lost their sight the research shows that they have a bigger rate of falls [19] and smaller walking speed [41] compared to the non visually impaired.

Another problem that rises before the visually impaired is orientation in space. Visually impaired face many troubles with orientation, e.g. they are struggling with capturing important landmarks of an environment to form a mental map. While sighted people rely on their vision to orientate in unfamiliar environments, visually impaired rely on their other four senses to understand the space around them. Even though in most developed countries cities are equipped with public navigation for the visually impaired, the mobility still remains a problem, especially for the older adults. There are

several solutions on how to help this community to navigate in space, e.g. mobile applications, tactile maps and so on. Any kind of assisting orientation technology, of course, first demands training. Talking about the elderly, orientation training seems to be the hardest social-rehabilitation activity for most of them according to [6].

Presently there are no such 3D tactile maps that would allow a person with visual disabilities to explore the interior and the exterior part of the building, while being able to move from one floor to another. For this reason Miroslav Macík et. al. [33] described a new method for indoor orientation and navigation for people with disabilities in the form of the haptic 3D printed model. Their work aims to simplify and speed the process of creation of a mental model of an indoor environment for the visually impaired older adults. Subsequently, Ing. Dominika Palivcová in her master thesis followed up on the project of Macík and presented a prototype of interactive tactile map of a 3D model that allows users to explore each floor of the Palata residential building for visually impaired. [15] In Palivcová's work she was focusing on design and implementation of an internal part of Palata, but not the exterior part of the building.

This thesis follows up on the work of my colleagues ([33], [15]). That is, the aim of this thesis is to implement and print the 3D printable exterior side of Palata and to design a pipeline prototype that will allow for creation of exterior 3D printable building parts for a tactile 3D model. It is important to have the exterior part of the building as it helps the visually impaired to create a mental map of the external appearance of the building and its overall layout. Furthermore, pipeline prototype can be used for creation of other buildings' 3D tactile maps for people with visual disabilities.

1.1 The Goals Of This Thesis

There are several goals to be achieved in this thesis:

1. Analysis of the outcomes of current research on the spatial cognition of visually impaired with a particular focus on older adults

   Analyze means of spacial cognition of the visually impaired. Study current research on spacial cognition of the people with visual disabilities. Follow up the results of research conducted at the DCGI and on solution described in the patent titled 'Three-dimensional model of a building primarily for visually impaired older adults' [33]

2. Study of a method for 3D model reconstruction from images

   Research a method for reconstruction of 3D models based on a set of images. Provide with some base knowledge for this process and describe the steps.

3. Analysis of a technique for segmentation of building facades

   Study a technique for facade elements detection from images of the building. Provide with individual steps of this process.

4. Definition of a pipeline for creation of a 3D exterior building part from images

   Define a pipeline for creation of a simplified (number of triangles is minimized) exterior part of a building based on a set of images. Facade elements, like windows, doors, ledges, should have only an opening. Roof should be excluded.

5. Implementation of the low and high-fidelity prototypes of the pipeline using the UCD method

   Design and implement two generations of the prototypes for the pipeline using the UCD method.

6. Design and implementation of a 3D printable exterior part of Palata

   Design and print a 3D exterior part of Palata following the needs and requirements of the visually impaired. Attach the resulting model to the wooden model created by Macík.

7. Evaluation of the prototypes and the 3D model with the target user group

   Evaluate both low and high-fidelity prototypes with the target user group. Furthermore, evaluate the created 3D model with a visually impaired expert.

1.2 Methods and terms

1.2.1 User-Centered Design

User-Centered Design (UCD) is an iterative design process in which designers center around the clients and their needs in each period of the development procedure. In UCD designers use different research and design techniques and test their product with users in order to improve their product usability and accessibility. There's a special international standard that UCD follows: ISO 9241-210:2010. [18] The iterative process is composed of a few phases as is explained in the figure below.

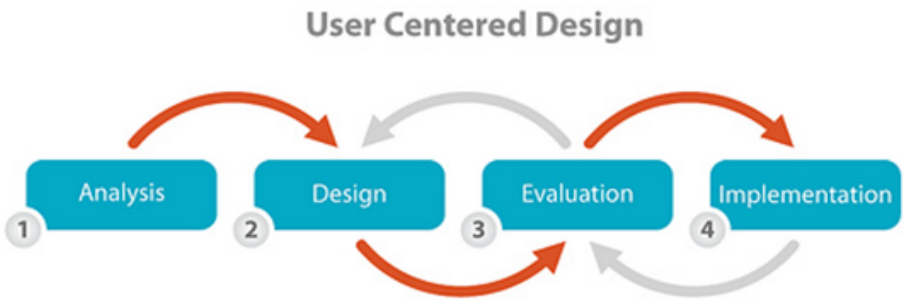


Figure 1.1: Representation of phases of UCD and their iterative nature, from [60]

1.2.2 Usability

Usability is a term used in Human-Computer Interaction (HCI) and is a quality attribute that describes how easy it is to handle user interface for a user. "Usability" as a word is also used for product improvement process during its design. This term is defined by several components [38]:

- Learnability - How easy it is for users to handle the user interface while using the system for the first time?
- Efficiency - How quickly do users complete tasks after they've learnt the interface?
- Memorability - How easily can user restore capability of handling the user interface after they have not been using this system interface for some time?
- Errors - How many errors are made by users, how serious are they and how easy is it to recover?
- Satisfaction - How pleasant is it to use the design?

It is important for a system or a product to be usable, since users do not tolerate difficult designs and will not use this product otherwise. In order to improve the usability designers must test the product with target users, observe their difficulties when using the product and improve it.

Chapter 2

Analysis

Visual impairment is a decreased ability to see which varies among different degrees of vision loss. Blindness, on the other hand, is complete or almost complete loss of vision. [13]

2.1 Statistics On Visual Impairment

The International Statistical Classification of Diseases divided the visual impairment in several groups as is shown in Table 2.1.

Category of visual impairment	Visual acuity with best possible correction		Or central visual field	Classified as
	Maximum less than	Minimum equal to or better than		
1	6/18 3/10 (0.3) 20/70	6/60 6/60 20/200		Low vision
2	6/60 1/10 (0.1) 20/200	3/60 1/20 (0.05) 20/400		Low vision
3	3/60 1/20 (0.05) 20/400	1/60 (finger counting at 1 metre) 1/50 (0.02) 5/300 (20/1200)	10 or less but more than 5	Blindness
4	1/60 (finger counting at 1 metre) 1/50 (0.02) 5/300	Light perception	5° or less	Blindness
5	No light perception			Blindness
9	Undetermined or unspecified			Unspecified

Table 2.1: Categories of severity of visual impairment, from [14]

In the year 2010 the World Health Organisation has reported statistics described in Table 2.2 that illustrate the situation in the world regarding the relationship of age to visual impairment.

Ages (in years)	Population (millions)	Blind (millions)	Low Vision (millions)	Visually Impaired (millions)
0-14	1,848.50	1.421	17.518	18.939
15-49	3548.2	5.784	74.463	80.248
50 and older	1,340.80	32.16	154.043	186.203
all ages	6,737.50	39.365 (0.58)	246.024 (3.65)	285.389 (4.24)

Table 2.2: Global estimation of the amount of people visually impaired by age and the corresponding prevalence (%), from [40]

The World Health Organisation also provides information regarding the causes of visual impairment (Figure 2.1). The main reasons are uncorrected refractive errors (RE) and cataracts. Other causes are glaucoma, age related macular degeneration (AMD), diabetic retinopathy (DR), trachoma and corneal opacities (CO). Also, a big percentage is undetermined.

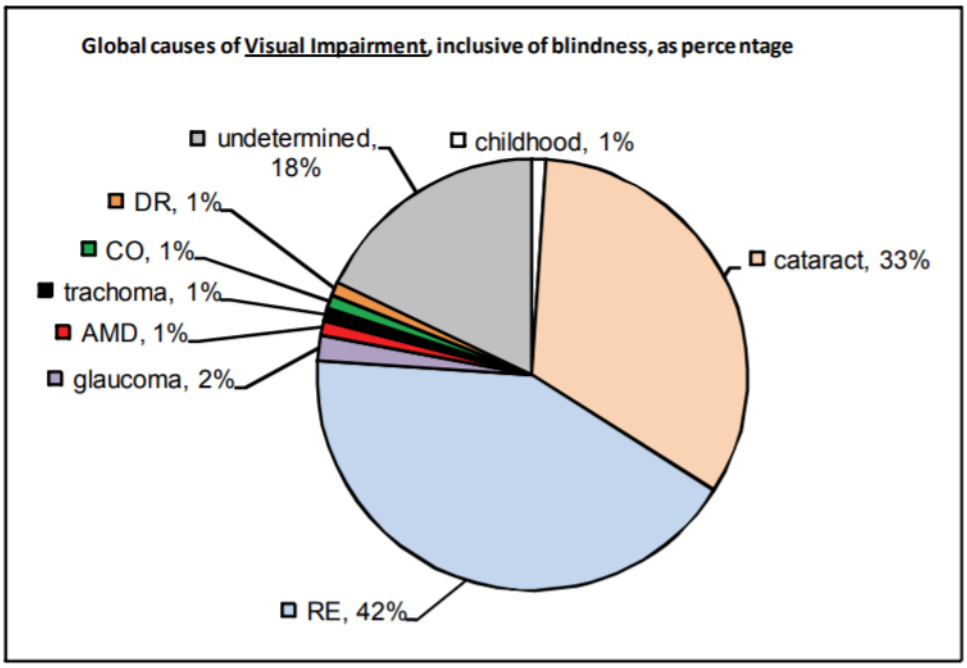


Figure 2.1: Causes of visual impairment globally in percentage, from [40]

2.2 Perception Of Blind And Visually Impaired

Orientation is "the ability to use one's remaining senses to understand one's location in the environment at any given time." [11]

Mental map or model is a phenomenon of a conscious experience of creation of a mental representation of anything based on a previous experience and knowledge. [12]

Spatial cognition is a branch of cognitive psychology. Spacial cognition studies how people acquire and use knowledge about their environment to determine their location, obtain resources and find a way home. [35]

2.2.1 Orientation And Navigation

Visually impaired and blind people have a different perception of the world comparing to the people with sight. The space becomes a set of guiding lines, markers and orientation points for the humans with vision disabilities. But again, when talking about the blind and people with a low vision, there is some differences: people who have low vision might still able to recognize high-contrast colors and shapes.

In the paper [39] Noordzij, Zuidhoek, and Postma state that according to their research early and late blind people build up spacial mental models more efficiently based on a route description, which is a description of the environment from the ground-level, whilst the sighted people tend to build it from a survey description, which is the "bird-eye" view on the environment and requires the knowledge of topology. This facts imply that a system that is designed specifically for people with a sight loss has to meet some appropriate requirements in order to satisfy the needs.

Visually impaired pursue the orientation signs in space using the following means: hearing, smell, touch and temperature. The sound feedback is very important for them if they are outside, especially in the crossroads and crossings. [7] Tarik Kapić in his paper [27] conducted a survey where he performed several interview with blind and visually impaired to find what aids are necessary to use for a navigation system. He found out several facts about the way visually impaired orientate in the outside space:

- The blind often use objects such as walls, balustrades, murals etc. in order to define directions
- Another big part in orientation of visually impaired plays the smell, e.g. bakeries, cafes and restaurants.
- Visually impaired also use their hearing senses a lot, e.g. the sound of the fountain, semaphore sound or a cafe that plays music daily.
- Configuration of the floor also plays a big role, e.g. the end of a pavement, change of the flooring, etc.

Macík, Malý, Balata and Míkovec in their paper [32] state that people with a sight loss tend to code spacial relations by reference with their body, which is proven to be an ineffective way for solving complex tasks. It is also proven that congenitally blind tend to code features in familiar routes better, than in an unknown ones. Research [15] conducted in Palata residence on elder people shows that some visually impaired navigate and orientate independently, but only along the routes that they already know. In fact, learning new routes seems problematic for the new elder clients: they prefer the help of the personnel and it takes a few months to learn the route.

2.2.2 Auditory Senses

One of the means of the world perception is sound. According to [44] hearing processes 11% of our perception of the world. For the visually impaired hearing is especially crucial, as it remains the main instrument of their experience of the world.

Researches show that blind participants show better sound localization than sighted people. Röder et. al. demonstrate in their studies [42] that for congenitally blind participants speech processing activates not only left-hemispheric language area, as for the sighted people, but also homologue right-hemispheric structures. This fact relates to Braille-reading abilities of visually impaired. It has been suggested that Braille-reading is connected more to the right hemisphere than reading printed languages.

People with visual disabilities are more sensible to speech and are more efficient in processing it [26] since their main reliance is sound. When semantic and syntactic processing in blind were tested [42], giving the participants both blind and sighted a phrase, where an adjective preceded a noun or a pseudo-word, it was found that the reaction of the blind participants was quicker than of the sighted subjects. This way, it was conducted that blind process the speech signals more effective. It is assumed that due to this fact blind might have higher working memory capacities. It has also been found, that blind perceive the speech better in the context of the noisy background. [37]

2.2.3 Tactile Senses

Tactile perception are sensations obtained through the skin. Tactile sense is one of the most important senses for the blind and visually impaired, so they perceive the space through touching the objects around them. Size, material and form are the key characteristics that help visually impaired to identify an object. The interaction with an object can be done by touching, grasping, pushing or pulling. [10]

Many studies ([20], [51], [58]) show that tactile perception is more enhanced for blind people. However, researches show that congenitally blind show worse results in shape understanding than the blind people who had a vision before, which implies that visual experience plays a big role in determining the object shapes. Although, when comparing visually impaired and sighted people both groups show the same skills in representations of haptic material properties [5].

Daniel Goldreich and Ingrid Kanics in their research [20] on tactile acuity in blindness had found out that blind women have heightened tactile acuity comparing to men. They assume the reason is their greater skin compliance. In addition, it has been found that young blind participants had higher tactile acuity than the older subjects, which is believed to be caused by peripheral and/or central neuronal loss for the older adults.

Macík et. al. in their paper "Tactile symbols for visually impaired older adults" [31] have conducted two experiments on the right shape, elevation, recognizability and intelligibility of tactile symbols and symbol preference in general. The results show, that older adults with vision disabilities show more understanding towards the outline design of symbols comparing to the inline. The authors have observed, that some symbols, like down filled arrow, heart, cross and outline symbol in the form of two lines was easily connected to real-life objects by the participants (Figure 2.2). On the other hand, clock and two lines were often mistaken. Question mark was recognized by some people as a number two. This testing also showed, that most of the visually impaired were not able to assign a symbol to its meaning. The authors assume, that the meanings of the symbols were not based on mental models relevant to the target user audience. The author of [10] assumes that tactile sense allows for a smaller level of abstraction comparing to vision and hearing.

It has been shown in research of Ungar et. al. [45], that patterns, textures, separation, and spacing of individual features are the most important aspects of tactile symbol design. Shape, type, and size of the symbol have scored less,

Figure 2.2: Example of tactile symbols tested by Macík et. al., from [31]

but above average, in their study.

Finally, an interesting fact: visual cortex is being triggered from haptic experience for blind people. Several researches [9] [47] [46] that run an experiment where they were measuring the brain activity in visual cortex in blind people have found out that the visual cortex is being activated when Braille-reading.

2.3 Tactile Maps As Means Of Spacial Cognition

2.3.1 What Is A Map?

Map is a symbolic depiction of chosen qualities of a place. The process of map creation considers the needs of the target user group, which can differ, and that is why there are different types of maps: electronic, climatic, topological, tactile, 2D, 3D etc. Human brain processes maps differently than it processes graphical information, thus map design has its own specialities. Map reading is not easy, therefore map creator needs to depict all symbols in such an efficient and effective way so the user will interpret them well.

The Figure 2.3 depicts four stages of information transformation. The communication between a map maker and a map reader is the following:

- Reality is interpreted by a map maker.
- Map maker creates a tangible map according to their internal interpretation.
- Map reader interprets the created map through their own interpretation of the world, which can be different than the interpretation of the map maker.

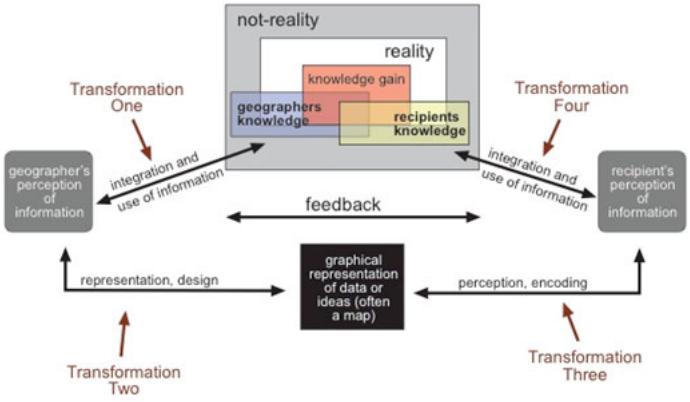


Figure 2.3: Cartographic Communication Model adapted from MacEachren, 1995, from [30]

- Map reader uses the map according to their interpretation for a navigation task.

Transformations one and two of the information transformation model are for the map maker, while transformations three and four focus on the map reader. Transformation one corresponds to a process of interpretation of the world, while the transformation two represents the process of design of the map. At the transformation three user encodes the information from the map and in the transformation four they use this information to complete some navigational task.

2.3.2 Tactile Maps

Tactile maps provide blind and people with low vision with an understanding of the environment. This kind of maps are made of accessible images, that provide user with sound, verbal description or haptic feedback.

Studies [22] have shown, that visual deprivation does not affect the amount of information acquired when touching a tactile map, nor does it affect the cognitive load, when building a mental map. In other words, the level of

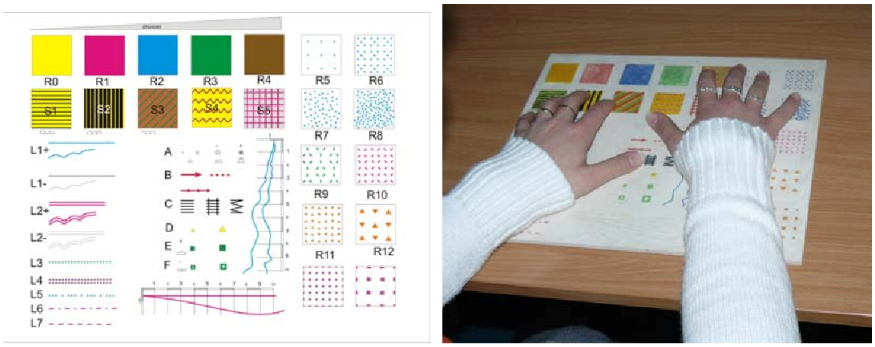


Figure 2.4: Tactile Map Testing, from [59]

difficulty of learning the tactile map is the same for persons with impaired vision and for those with a sight is the same.

After a 5-year research and years of repeated user testing that was based on previous studies of the tactile navigation map symbols Amy Lobben et. al. from Department of Geography at the University of Oregon have come up with the most usability-friendly symbols for the tactile map signs (see Figure 2.5).

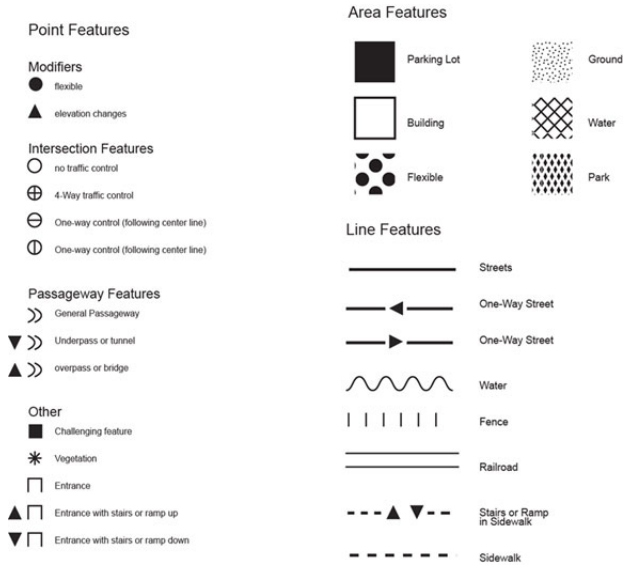


Figure 2.5: Tactile Navigation Map Symbols, from [30]

Ungar et. al. [57] have conducted several experiments to understand how visually disabled people build up their spatial knowledge using the tactile maps. Results show, that tactile maps indeed help the participants to remember the route: subjects that studied the tactile map before went out in the streets lost their way less often and walked less meters off the studied route comparing to other participants. In fact, it was discovered that the worst results had a group that only has verbal explanation of the route. This

research has detected several map learning techniques of the children with a lost sight:

- using of mnemonics (e.g. rehearsing a list of names),
- making associations with map features (e.g. "this is like my school"),
- counting items on the map,
- repeatedly reading out names from the map,
- following a route through the map,
- looking for patterns of features (e.g. finding geometric shapes and other figures),
- relating features to the map frame (e.g. by identifying an item as the one at the top of the map, or in the corner). [57]

This experiment also showed the "good" map learners techniques:

- focus more on the relationships between the objects on the map,
- find local and global patterns,
- understand locations of structures regarding to the external map framework.

2.3.3 3D Printed Tactile Maps

Holloway has shown in her research that 3D tactile maps are preferred by the visually impaired users to 2D tactile maps. [25] Participants of her studies commented that objects are much more obvious when they are 3D. Since 3D printing technology market expands every day, 3D printers became available also for an amateur computer users. This brought in many researches that focus on creation of tactile maps using 3D printers. e.g. HaptoRender [4] is a project that made a 3D printed tactile map from OpenStreetMap data. Another project is Touch Mapper, it allows users to query an address and download a 3D tactile map of that place and the surrounding area. [29]

There is still no big production of tactile maps in general and comparing to the production of software maps. Cartographic conventions are not broadly

applied for them. [30] However, 3D printing technology eases the way and cost of production of such maps. Also, it is now possible to create tactile maps individually according to the person's needs.

2.3.4 Interactive Tactile Maps

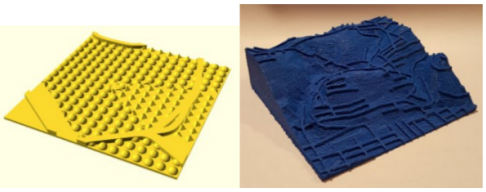


Figure 2.6: Thin raised line maps, from [54]

Interactive tactile maps is a promising approach in 3D printed tactile map design. Due to the fact that Braille language is known only to a small part of visually impaired community, interactive tactile maps that require this knowledge, are not very popular.

Brock et al. [8] used a commercially available multi-touchscreen devices combined with raised line prints on paper and examined gestures that were used to enhance interactive tactile map. This device did not require any Braille knowledge. The results were very positive - users were satisfied.

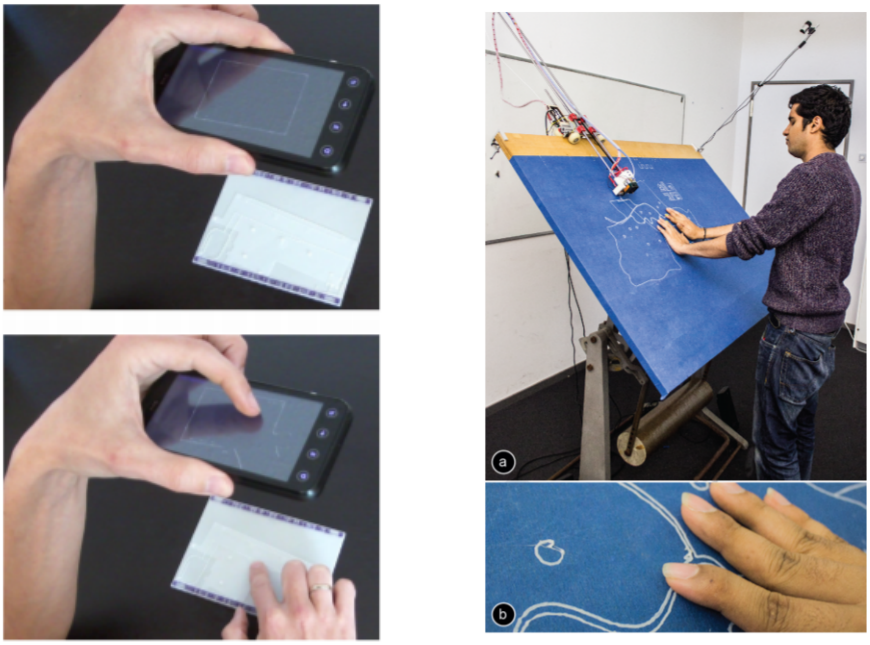


Figure 2.7: First picture: User is exploring the tactile map using one hand, the other hand obtains detailed information from the phone using special gestures, from [21]; Second picture: Linespace technology for touch detection, from [52]

Another way for interactive tactile maps implementation is using Linespace vision-based approach [52] detecting user actions while they are touching the surface of the tactile map (Figure 2.7). Linespace has more than a square meter interactive surface, which helps to catch the actions of the user. This is a promising approach, but the only problem with this system is that it is not portable.

Gotzelmann [21] has come up with another interesting approach of embedding bar codes along the edges of 3D printed map models (Figure 2.7). These bar codes can be scanned by a mobile device or a tablet. The issue with this method is that a user needs to point their cameras into a limited space frame while tapping the phone, which for visually disabled might be a problem, especially for the older adults.

Another technique that was implemented by Taylor, Dey, Siewiorek and Smailagic [54] was based on using the touchscreen overlay systems that relied on thin raised line maps (Figure 2.6). Touch screen sensitivity was increased by using a conductive filament for 3D printing. These maps work with interactive applications in common touchscreen devices.

Finally, Macík et. al. proposed an interactive 3D model for orientation and navigation of primarily visually impaired, but sighed people as well. [33] My colleague Dominika Palivcová [15] has designed and implemented interactive tactile floor plans of the building Palata following the ideas of [33], I will be talking about this project in the next section.

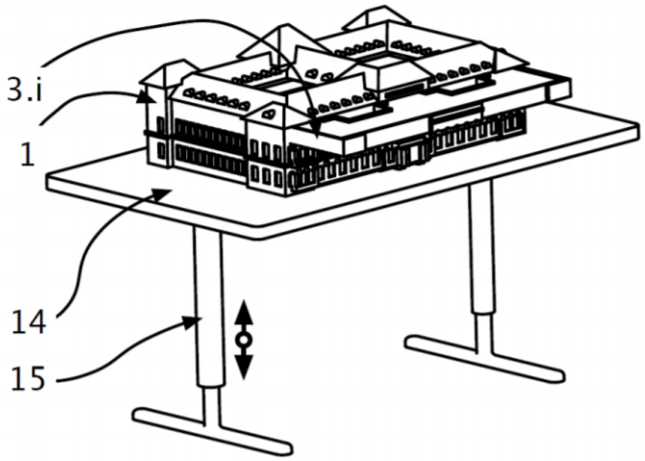


Figure 2.8: Interactive 3D model proposed by Macík et. al., from [33]

2.4 Home Palata For Visually Impaired

2.4.1 The Institution

Palata is a residential care institution for visually impaired older adults in Prague, Czech Republic. This institution was founded in 1888 and was renovated recently. The average age of their clients is 84, 75% of them are women. The Palata clients have different visual impairment types.

Palata resembles a shape of the number 8, but with right angles (Figure 2.9). It has 39 one-bed and 47 two-bedrooms. Clients of Palata are offered different physical and rehabilitation activities: there are special rooms for it in the building. This residence is built in such a way so it suites the needs of the people that live there: hand railings in the corridors to support safe movement, haptic marks for indication of the doors, high-contrast horizontal navigation lines in each corridor, furniture in each corridor in case the client need to rest (Figure 2.14).

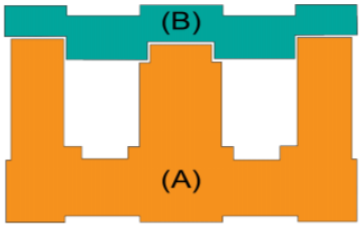


Figure 2.9: Palata from the top view. Part B is newly built, part A is the old part, from [32]

2.4.2 Navigation And Orientation Of Clients In Palata

Research in a form of interviews was conducted in Palata by Macík [32] exploring the every day life of clients of this institution. Part of his research was focusing on their orientation and navigation. During the interviews with the clients a few interesting things have been found:

- Handrails serve as a mean for clients to learn the unfamiliar paths with time.
- It is hard for the clients to learn new paths, they rather not even try because of a fear to get lost.
- Some participants do not go alone out of the room.

Figure 2.10: A corridor in Palata. Red zoom shows the haptic marks, dashed green line contrast the navigation line, blue color marks the furniture for the resting, from [32]

- For the clients with residual sight high-contrast horizontal guiding lines serve as a great aid in terms of navigation.
- Physical state of the client can affect their performance in orientation and navigation.
- Clients use signs near the door to orientate themselves.
- A lot of times clients of Palata ask for help to the staff.

2.5 3D printable Palata model

2.6 The Idea

Macík et. al. patented an interactive 3D model for orientation and navigation of visually impaired older adults. [33] Their design idea is new to the world, since no one yet implemented such an aid for the people with visual disabilities,

Figure 2.11: First picture: Wooden model of Palata made by Macík following the patent [33]; Second picture: Control to regulate the height of the model.

that is 3D printable, would allow users to examine both interior and exterior parts and to be able to move from one floor to another.

The proposed 3D model would offer its users possibilities to learn routes to the most important places of the building, like a canteen or a bathroom. The building would be presented as a set of drawers, that can be pulled out and examined. The building model is located on the adjustable stands, so user can modify the height (Figure 2.11). Each floor would have a plan of rooms, hallways, toilets etc. Each room would be indicated by a specific tactile symbol, whereas each hallway by a tactile line. The auditory support of the navigation is also planned. It is assumed this model will be 3D printed on acrylic glass.

2.6.1 Interactive Tactile Floor Plans

Dominika Palivcová as a follow up on the project [33] has created a 3D printed prototype of one part of a floor of Palata Home in her master thesis [15]. The floor map prototype consists of doors, hallways, rooms, stairs and toilets. This prototype offers a route-guidance mode, which guides users from one room to another. Here are a few characteristics of the model:

- Rooms are represented as a raised line.
- Each room has a button that is recognizable by touch as a circle (see the middle button in Figure 2.12).
- The hallway is represented as two raised lines.
- Doors are depicted as a gap in a raised line.
- Stairs are represented with a small 3D model of a staircase.
- Toilet signs are represented as spheres.

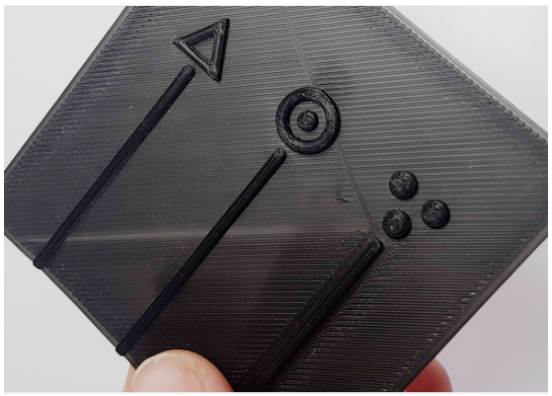


Figure 2.12: Initial few designs of room buttons. In the end the middle one was chosen for the final prototype, from [15]

- All rooms and stairs are in yellow color with a black border, which helps the low vision people to recognize it.
- The starting point is clipped to one of the room buttons as it is its own tactile symbol, it is in green color.
- Hallways are depicted with a black guiding line.

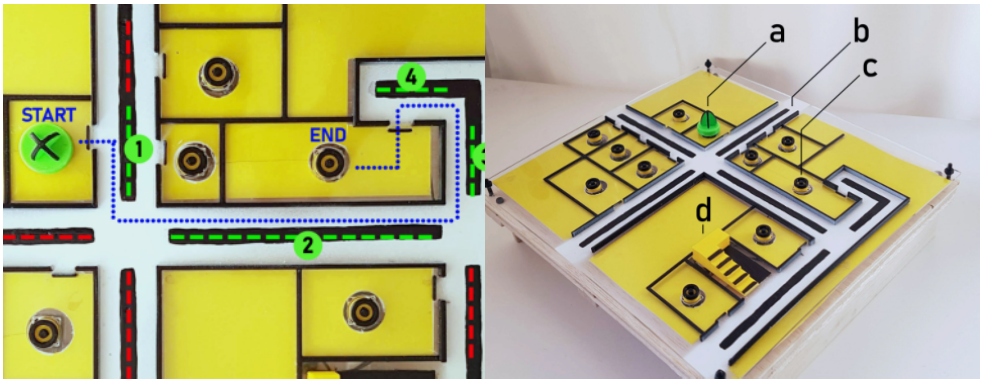


Figure 2.13: First picture: High-fidelity prototype of Palata floor interactive tactile map. *Blue line* are the passage that user should follow. *Green line* - segments of the route which play piano tones when are being touched. *Red line* are segments off the route, when touching "no-go" sound plays. Second picture: *a* - start/end point for navigation; *b* - high-contrasted hallway that is marked by the two raised lines; *c* -symbol for a room, *d* - stairs, from [15]

The tactile plan provides an auditory feedback: while testing the prototype when user touched a button the test moderator played the appropriate sound. For example, if the participant follows the right route, they will hear increasing tones of piano, or if the participant went off the main route, they will hear

a "no-go" sound. Whenever the user reaches the final part of the floor, they will hear a special sound and information about the location of the doors to the target room.

The results of the testing of this 3D model and its route-guidance mode show, that majority of participants (7/10) were satisfied with the model and seemed to be able to build a mental map, except for those who had moderate or severe dementia.

2.6.2 Exterior Part of Palata

The exterior design of Palata, which is the aim of this thesis, must provide users with an understanding of location of windows and doors in the building and help them with creation of a mental map. The design and implementation of the facade elements, like windows and doors, and the roof itself are a part of another projects supervised by Macík. It is advised in the patent [33] that the exterior part is 3D printed and then glued to the wooden model created by Macík, so the sizes of this structure must correspond with the sizes of the wooden model. The material should be ABS, PLA, PET, PETG, PTFE or similar.

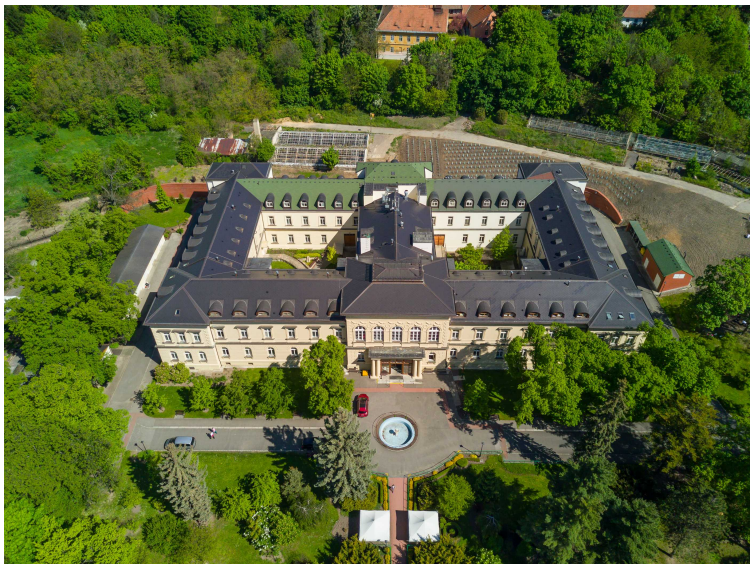


Figure 2.14: Picture of Palata taken from a drone by Macík

Exterior part can be as important as the interior. Here are a few reasons why the exterior building model is important:

1. Giving an idea of the sizes and proportions

   The exterior model of Palata gives its the users a sense of the size and proportions of the building. This is important, because most of the time visually impaired have a distorted idea of the size of the building they are in. This way, they can explore it using, e.g. some kind of a tactile line next to the model that would indicate 10 meters distance and then measure the building based on the size of the given line.

2. Creation of a mental map

   It is also important to give the clients an idea of where they live, or to extend their current idea. One client of Palata thought that the building is in a form of a round number "8". Visually impaired in Palata most of the time walk around in the building only a few meters around their room and when they make a turn they change the direction gradually, which in the end make the route in an arc form. It is necessary to transform their perception of Palata from an unknown maze into an easy map.

3. Orientation and navigation

   Further reason is to provide clients with means of orientation and navigation. For example, by counting windows they can reach a certain destination.

4. Make the clients appreciate their home

   It is important to give the users an idea that they live in a castle, to provoke good feelings towards the place they live in, so that they enjoy it more and are proud of it.

5. Increase of mobility

   Another reason is to make clients of Palata more mobile. The building must provide users with understanding, that it is not so scary to go to visit their friend on the other corner of the building.

6. PR

   It is a good PR. Even though the main purpose of this project is to make the lives of the Palata clients easier and more comfortable, another effect is that it gives the sponsors an idea that their money is well spent and that there is a modern technological improvement in quality of life and accessibility of the institution.

7. Provide visual experience for sighted users

   The visual form and physical implementation of this model will enable both tactile interaction for the visually impaired, as well as accurate display of specifics of the building for sighted users.

Chapter 3

3D Building Reconstruction From Images

For a human it is natural to build a 3D model of the environment based on what we acquire with our eyes, however it is not so easy for computers. The problem of image-based 3D model reconstruction is one of the challenges in computer vision. The objective is to acquire geometric and semantic information from 2D camera images and transform them into a 3D model. This problem is applicable in such applications as heritage preservation, digital mapping, gaming and so on. The aim of my work is also closely connected to this - it is necessary to reconstruct Palata from a series of pictures to create a tactile model for better orientation inside of the building. The goal is to create an accurate but not a highly detailed 3D model based on 2D pictures from different points of view. The following analysis was inspired by the thesis [49] of Johannes L. Schönberger, the inventor of *Colmap*, the most popular program for 3D model reconstruction from images. [48]

3.1 Image Association

One of the core principles of 3D model reconstruction is "sewing together" images. This process is about finding robust and efficient geometric correspondences between images. The association involves finding each picture's description and looking for similar descriptions in other pictures.

There are two approaches for finding a picture's description: global (e.g, "Image shows a house") and local (e.g., "One of image's regions shows a

door"). Local approach enables to associate pictures on detail level and so it is more precise. Global one is less accurate but more efficient when using a big amount of pictures.

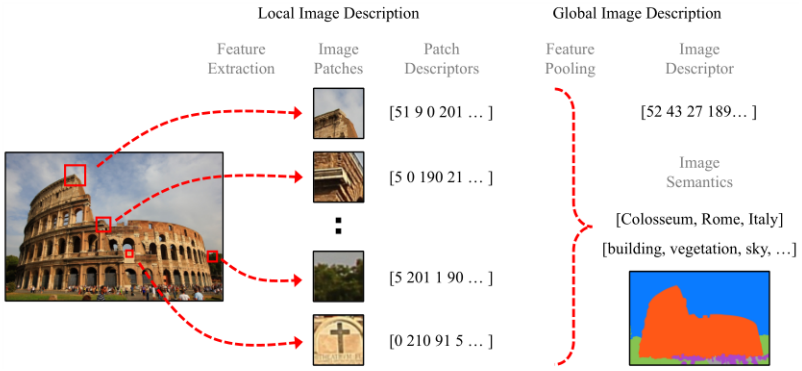


Figure 3.1: Local and global image description, from [49]

As an example local features could be points and edges which should be distinctive and detectable in different images as the same object. Two same points in two different images are distinguished based on their common geometry, e.g. size, orientation and position, or appearance, e.g. color gradient vector of pixels in a neighboring area. Local approach can encode up to hundreds of thousand local features such as points in one picture.

Global approach sums up the picture into a single representation, e.g. list of numerical values or high-level semantic category. It is much more laconic compared to a local approach, that is why image comparison using the global image description is faster. The way global description is obtained is through summarizing the local features into a histogram of a fixed size.

An example of such a feature detection algorithm is *SIFT*.

3.1.1 SIFT

SIFT is a feature detection algorithm in Computer Vision and it consists of four steps: scale space construction, keypoint localization, orientation assignment and keypoint descriptor. In the first step Gaussian Blurring technique is used to reduce noise in the image by creating different scales of the same picture and creating blur versions of each of them. Applying Difference Gaussian on each two consecutive levels of scale then leads to enhancing features for each of these images. Next step is to find keypoints by finding local maxima and minima by going through each pixel and comparing

it with nine neighbouring pixels of this level and also with the level below and above.

Further, it is necessary to calculate orientation, magnitude and construct histograms. To calculate both of these for each pixel first the gradient must be calculated in x and *y-axis*. Then using the following formulae magnitude and orientation is calculated.

$$\text{Magnitude } = \sqrt{\left[(\mathrm{G_x})^2 + (\mathrm{G_y})^2 \right]}$$
$$\Phi = \mathrm{atan}(\mathrm{Gy}/\mathrm{Gx})$$

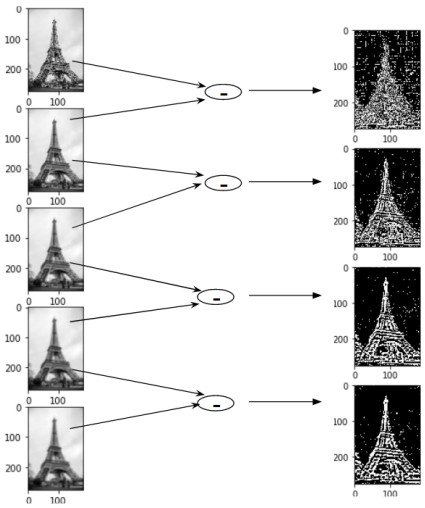


Figure 3.2: Four images generated by subtracting the consecutive Gaussians, from [50]

After creating a histogram it is possible to determine the key points - the peaks in the histogram. Each peek has its own orientation that corresponds to it. This way this step is complete.

The last step of the algorithm is to determine a unique fingerprint for a set of neighbouring pixels called descriptor. Each keypoint is taken along with it's 16×16 pixel neighbourhood and a histogram is plotted using magnitude and orientation.

After computing these steps for each of the images of the building gathered from different angles feature matching on all can be computed. [50]

3.1.2 Other Algorithms

An alternative for SIFT algorithm can be SURF, which is a SIFT approximation, but instead Gaussian averaging the image squares are used. Convolution with a square is faster for an integral image. Interest points are then found using the Hessian matrix. To find the orientation Gaussian weights are applied. Descriptors are then calculated by taking the region around the pixel,

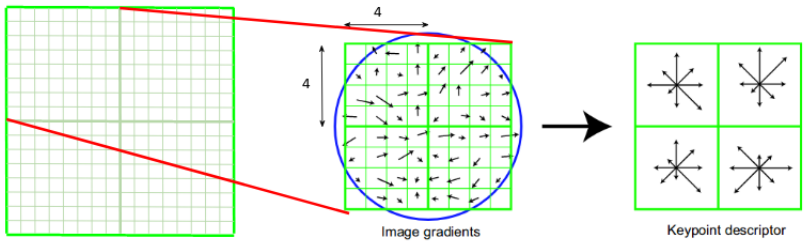


Figure 3.3: 16×16 neighborhood around the keypoint is divided into 4×4 subblocks. Arrows define the magnitude that defines the descriptor. From [50].

dividing it into smaller subregions and then looking at their Haar-wavelet responses. [56] SURF performs faster than SIFT, it doesn't reduce the quality of the detected points. [28]

3.2 Multi-View Geometry

Image capturing process is easy to describe mathematically, the inversion process is a bit more complicated. In the Figure 3.4 C is the camera center, through which all light passes from the scene point X and is captured on the image plane.

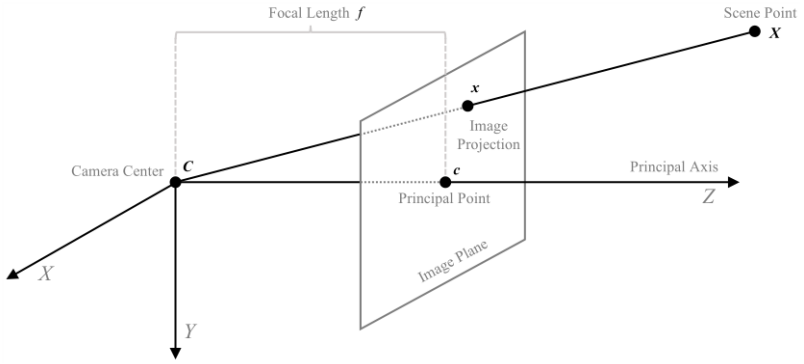


Figure 3.4: Pinhole camera with the local camera coordinate system defined by the axes *(X, Y, Z)*. Point X from the scene is projected onto the image plane at x through the camera projection center C. f is focal length. From [49].

The following formula where P is a 3×4 rank-3 matrix that projects point X in 3D space onto a point x in a 2D image plane which consists of R as a 3×3 rotation matrix, T as a 3×1 translation vector and K as a 3×3 calibration matrix.

$$\boldsymbol{x} \simeq \lambda \begin{bmatrix} u & v & 1 \end{bmatrix}^T = \boldsymbol{P}\boldsymbol{X} = \boldsymbol{K}[\boldsymbol{R} \quad \boldsymbol{T}]\boldsymbol{X} = \boldsymbol{K}\left[\boldsymbol{R} - \boldsymbol{R}^T\boldsymbol{C}\right]\boldsymbol{X}$$

Since projection matrix P consists of 12 unknowns, it is necessary to have 6 pictures to solve the following equation for matrix P estimation:

$$\begin{matrix} \boldsymbol{P}_3^T\boldsymbol{X}u = \boldsymbol{P}_1^T\boldsymbol{X} \\ \boldsymbol{P}_3^T\boldsymbol{X}v = \boldsymbol{P}_2^T\boldsymbol{X} \end{matrix} \Rightarrow \boldsymbol{0} = \boldsymbol{A}\begin{bmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \boldsymbol{P}_3 \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^T & 0 & -\boldsymbol{X}^T u \\ 0 & \boldsymbol{X}^T & -\boldsymbol{X}^T v \end{bmatrix}\begin{bmatrix} \boldsymbol{P}_1 \\ \boldsymbol{P}_2 \\ \boldsymbol{P}_3 \end{bmatrix}$$

Solving the null space of matrix A or, in other words, singular value decomposition of A, would lead to solving this homogeneous equation and finding P and its intrinsic and extrinsic components. Extrinsic parameters of a camera are those that define the location and orientation of the camera with respect to the world frame - rotation matrix and translation vector. Intrinsic parameters are those that allow a mapping between camera coordinates and pixel coordinates in the image frame - focal length, the optical center and the skew coefficient.

All of the previous theory assumes the perfect pinhole camera is in use, which in practice is not so. Real-world camera produces radial lens distortion effect, which occurs when light rays bend in the corners of the lens. [61] This is solved by using a distortion correction function.

$$\begin{aligned} x_{\text{distorted}} &= x\left(1 + k_1^* r^2 + k_2^* r^4 + k_3^* r^6\right) \\ y_{\text{distorted}} &= y\left(1 + k_1^* r^2 + k_2^* r^4 + k_3^* r^6\right) \end{aligned}$$

Where k_1, k_2, k_3 are radial distortion coefficients and $x^2 + y^2 = r^2$

It is not as simple as inverting the first equation in order to actually find the point X in the 3D scene since any two points along the same viewing ray are equal in projective space.

An equation for finding the point X in 3D space is the following:

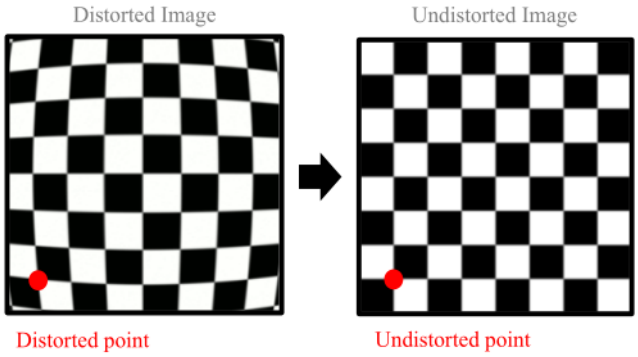


Figure 3.5: Lens distortion effects in the image, from [49]

$$\overline{\boldsymbol{X}} = \lambda \boldsymbol{R}^T \boldsymbol{K}^{-1} \boldsymbol{x} + \boldsymbol{C}$$

It assumes that the scale factor λ is known. The scale factor is the parameter that can tell what is the distance from the camera center C to the 3D point X. Point X can be found by intersecting the corresponding rays from several images and this process is called triangulation.

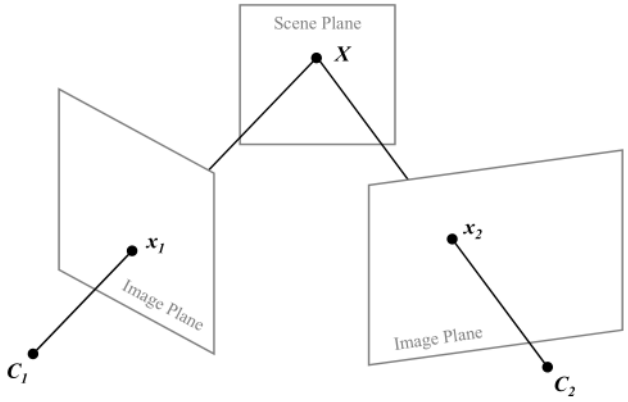


Figure 3.6: Homography maps points from one plane to another in a planar scene, from [49]

3.3 Homography

Homography is a type of transformation that maps straight lines onto another plane and preserves them. [16] Homography matrix can be estimated from at least 4 image correspondences in two views because it has 8 degrees of freedom. In the following equation point x_1 by homography onto a point x_2, where $h_{31} = h_{32} = 0$ and $h_{33} = 1$.

$$\boldsymbol{x}_2 \simeq \lambda_2 \begin{bmatrix} u_2 & v_2 & 1 \end{bmatrix}^T = h\left(\boldsymbol{x}_1\right) = \boldsymbol{H}\boldsymbol{x}_1 = \begin{bmatrix} \boldsymbol{H}_1^T \\ \boldsymbol{H}_2^T \\ \boldsymbol{H}_3^T \end{bmatrix} \boldsymbol{x}_1 = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \boldsymbol{x}_1$$

Camera can be moved purely and arbitrarily. The following formula describes homography in a case, where K_1, K_2 are the calibration matrices of two cameras, R_2 and T_2 is rotation matrix and transition vector of the second camera correspondingly, N is the unit normal vector of the scene plane and d the orthogonal distance of the scene plane to the projection center of the first camera.

$$\boldsymbol{H} = \boldsymbol{K}_2 \left(\boldsymbol{R}_2 - \frac{\boldsymbol{T}_2 \boldsymbol{N}^T}{d} \right) \boldsymbol{K}_1^{-1}$$

In case of pure rotational camera motion the projection center and the viewing rays do not change, in which case the equation is reduced to

$$\boldsymbol{H} = \boldsymbol{K}_2 \boldsymbol{R}_2 \boldsymbol{K}_1^{-1}$$

For an arbitrarily moving camera where the centers of projection of two cameras are different epipolar geometry is used. Epipole is a projection of another camera center to an image plane of the current camera.

$$e_1 = P_1 C_2 \quad \text{and} \quad e_2 = P_2 C_1$$

Epipolar line is a line created by the projection of point X in the 2D image plane and an epipole.

$$l_1 = e_1 \times x_1$$

Epipolar plane is an extension of the epipolar line along the viewing ray.

$$\Pi = \boldsymbol{P}_1^{+} \boldsymbol{l}_1$$

The following fundamental matrix that maps lines from one image to another image with minimum 7 obligatory correspondences is represented as the following:

$$\boldsymbol{F} = \boldsymbol{P}_2 \boldsymbol{P}_1^{+} \left[\boldsymbol{e}_1\right]_{\mathrm{x}} = \boldsymbol{K}_2^{-T} \left[\boldsymbol{T}_{12}\right]_{\times} \boldsymbol{R}_{12} \boldsymbol{K}_1^{-1} = \begin{bmatrix} \boldsymbol{F}_1^T \\ \boldsymbol{F}_2^T \\ \boldsymbol{F}_3^T \end{bmatrix}$$

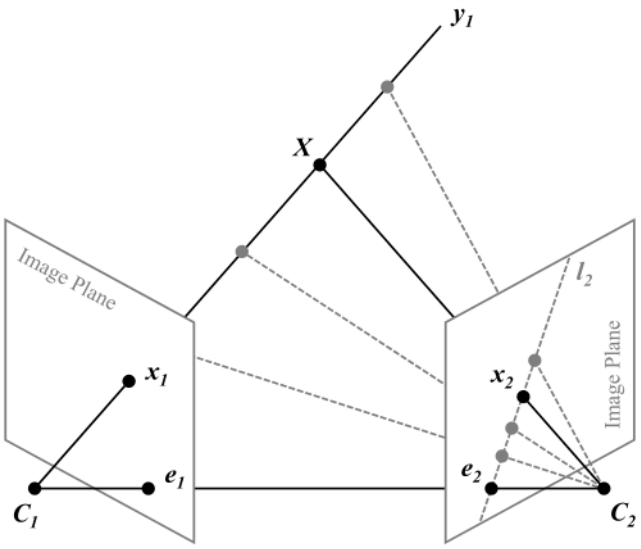


Figure 3.7: The fundamental matrix maps points from one image to lines in the other image. This process describes the two-view geometry with non-stationary camera motion. From [49].

3.4 Non-linear Estimation Algorithms

Due to observational noise the linear algorithms for reconstructing 3D scene out of 2D images presented before are not the most accurate and robust, that is why non-linear algorithms are used.

The following equation estimates structure X and motion P in a least-squares approach which minimizes the reprojection error assuming zero-mean Gaussian noise on the image observations and is called bundle adjustment.

$$\boldsymbol{X}^*, \boldsymbol{P}^* = \arg\min_{\boldsymbol{X},\boldsymbol{P}} \left\| \boldsymbol{x} - \tfrac{1}{\lambda}\boldsymbol{P}\boldsymbol{X} \right\| = \arg\min_{\boldsymbol{X},\boldsymbol{P}} \|\boldsymbol{e}_{\boldsymbol{x}}\|$$

Homorgraphy and fundamental matrices in this case would be estimated in the following way again using least-squares approach:

$$\boldsymbol{H}^* = \arg\min_{H} \left\| \boldsymbol{x}_2 - \tfrac{1}{\lambda_2}\boldsymbol{H}\boldsymbol{x}_1 \right\| + \left\| \boldsymbol{x}_1 - \tfrac{1}{\lambda_1}\boldsymbol{H}^{-1}\boldsymbol{x}_2 \right\|$$

$$\boldsymbol{F}^* = \arg\min_{\boldsymbol{F}} d\left(\boldsymbol{x}_2, \boldsymbol{F}\boldsymbol{x}_1\right)^2 + d\left(\boldsymbol{x}_1, \boldsymbol{F}^T\boldsymbol{x}_2\right)^2$$

One way to minimize all of the cost functions above is to iteratively solve approximations to the original cost function, e.g. using gradient descent method. *Gradient descent method* starts from an initial estimate and iteratively minimizes the cost by taking small steps down the local linear approximation of the cost function until it reaches an optimal value. This method is presented in the following equation, where J is a Jacobian matrix.

$$g(\boldsymbol{\theta}) = \nabla \tfrac{1}{2}\|f(\boldsymbol{\theta})\|^2 = \boldsymbol{J}^T\boldsymbol{f}(\boldsymbol{\theta})$$

If an image has repetitive scene elements then it can lead to errors in correspondence. A common approach to eliminate outliers is to down-weight their influence in the cost function by using robust kernels. Such kernels can reduce the influence of different error terms.

3.5 3D model Reconstruction Pipeline

The first stage of 3D model reconstruction pipeline is *correspondence search*, which includes features extraction, matching and geometric verification. This stage produces a scene graph which is then used in the next part - *reconstruction stage*. Reconstruction stage involves solving homography between images, scene points triangulation, outlier filtering and refining the reconstruction using bundle adjustment. The final part of the pipeline is *dense reconstruction*, which uses the estimated scene model and camera calibrations to create a more detailed scene representation like a dense point cloud or a textured surface mesh.

3.5.1 Correspondence Search

Firstly, local feature extraction takes place in this stage for each picture, it produces an output as a geometry and appearance descriptor. Feature matching step puts all local features to a vocabulary tree to create a global image description and then compares all images finding the most visually similar images, which takes linear time. The output of this stage are image pairs and their supposed correspondences. Next stage is geometric verification which verifies potentially overlapping images and their supposed correspondences. If there are enough similar features, then the pair of images is considered verified. Scene graph is then created out of all verified pairs, where overlapping images and scene points are nodes, directed edges between images specify scene overlap and undirected edges indicate visibility between images and scene points.

3.5.2 Sparse Reconstruction

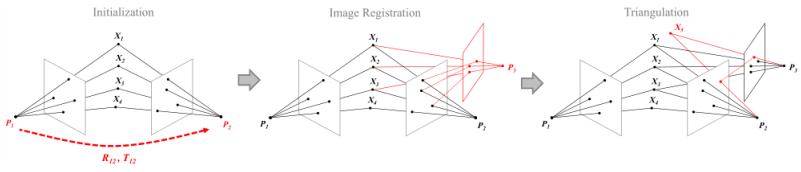


Figure 3.8: Incremental sparse reconstruction pipeline, from [49].

This pipeline stage computes the scene graph and outputs the intrinsic and extrinsic calibration and the triangulated scene structure as a set of points.

If there's no prior information about the scene orientation, location and scale remain unclear after this stage. That's why there are two approaches: global and incremental reconstruction methods.

The incremental method initializes the reconstruction from two-view reconstruction and incrementally registers new images into existing reconstruction, performs triangulation, outlier filtering and bundle adjustment refinement. Global approach on the other hand formulates the absolute camera motion as a collective optimization over all two-view geometries of the scene graph. Incremental approaches are more robust, but global are more efficient.

3.5.3 Dense Reconstruction

The previous step of the 3D reconstruction pipeline outputs a rather sparse reconstructed set of points, which hardly represents the real world. The purpose of this step is to make a richer representation, e.g. dense point cloud or a textured surface mesh. There are two main approaches that are different in their way of parametrization of the problem. First type parameterized the problem in image space and the second optimized for the surface directly in the scene space. Both of them have their own pros and cons. For example, the first type is better in terms of the fact that the number of parameters to estimate are limited by the number of pixels in the image. On the other hand, the second approach is better because it avoids redundant computations since the surface is observed from many cameras. The best solution is to mix these two approaches together, so that it leads to three steps: the recovery of dense depth maps for each image independently, the fusion of those depth maps into a globally consistent point cloud, and a final surface meshing and texturing step.

Chapter 4

Segmentation Of Building Facades

Since my work requires the knowledge of possible approaches for morphological building facade segmentation, this part will be dedicated to this topic. The main difficulty concerning this matter is the great variation of facades even in the same architectural style.

There are several methods for building facade segmentation. Two of them are model-free and model-based. The first one does not make assumptions about geometry of the scene, but rather addresses grouping the pixels according to similarities. Pixel that belong to the same facade don't have to share common characteristics, so this method can fail. Model-based approach takes into consideration prior knowledge and available image support. This method is error-prone because no static parametrization could be generic enough to deal with different facade layouts.

Another relatively new approach that involves machine learning is discussed in [55] and combines shape grammars, supervised classification and random walks. Authors of [55] use Parisian buildings which differ among each other in their architectural style. Pictures are taken in different lighting and weather conditions. This method uses procedural models like shape grammars, which help to deal with such a great variety of facades. Buildings can be represented a set of replacement rules and a dictionary of basic shapes.

Figure 4.1: Original image, classification-based segmentation, and segmentation with procedural shape prior, from [55]

4.1 2D Shape Grammar

Grammar by default definition consists of shapes and rules. Rules handle shapes using semantico-geometric building blocks, which are called basic shapes. The grammar for this concrete problem consists of a set of dictionary of basic shapes

$$\mathcal{D} = \{b_1, b_2, \ldots, b_K\}$$

and a set of rules

$$\mathcal{R} = \{r_1, r_2, \ldots, r_N\}$$

Basic shapes are defined to be squares and are positioned in the 2D plane through a bounding box called scope, which creates an atomic shape. Those are in their turn handled using replacing rules that turn a left-hand side atomic shape into some atomic shapes on the right-hand side. The only rule that is being used is a split-rule: it decomposes the scope of the LHS (left-hand side) atomic shape along a splitting direction into several chunks. The Derivative process takes an atomic shape and constantly replaces it using grammar rules. [55]

The following figure is a derivation scheme to create a facade:

Rule	LHS	$\rightarrow$	RHS	DoF
1.	IM	$\rightarrow$	GF.FA.AT.Sky	3
2.	GF	$\rightarrow$	Sh.Do.Sh	2
3.	FA	$\rightarrow$	(Ba?.FL.Wa)*	n_1
4.	FL	$\rightarrow$	Wa.$(TL.Wa)^*$	n_2
5.	AT	$\rightarrow$	Rf.(TL.Rf)*	n_2
6.	TL	$\rightarrow$	Ba?.Wi	1

Where the the vocabulary of basic shapes is defined by the symbols: IM (Image), AT (Attic), FA (Facade), GF (GroundFloor), FL (Floor), TL (Tile), Sky, Wa (Wall), Wi (Window), Sh (Shop), Do (Door), Rf (Roof) Ba (Balcony), where the bold face symbols are said to be terminals. This table represents a set of rules, all of them together represent a policy. [55]

4.2 Learning With Randomized Forests

"The random forest is a supervised learning algorithm that randomly creates and merges multiple decision trees into one "forest." The goal is not to rely on a single learning model, but rather a collection of decision models to improve accuracy." [24] Leaves contain a number of visits of input feature vectors and nodes contain a simple random test on a feature vector.

Training phase starts with running labeled feature vectors on the forest. Feature vector of a concrete label is dropped into a decision tree where it bubbles down to a leaf depending on a decision taken at an internal node. The number of visits per label is then updated in the leaf, which keeps a histogram that represents the number of feature vectors that reached this node. Histograms provide a probability of a feature vector to belong to a class C given a leaf l_τ to which this vector belongs:

$$P\left(c \mid l_\tau\right) = \frac{h_c}{\sum_i h_i}$$

The probability of the whole forest is then:

$$P\left(c \mid (l_1, \ldots, l_T)\right) = \frac{1}{T} \sum_{\tau=1}^{T} P\left(c \mid l_\tau\right)$$

In paper [55] small patches around the pixels are considered as feature vectors. Training began with manual labeling the buildings features like windows, walls, balconies, doors, roof, shop, sky and outliers.

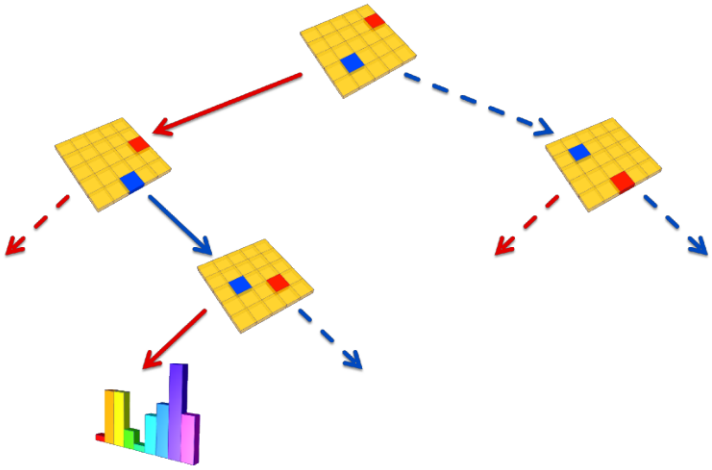


Figure 4.2: A patch is dropped in the decision tree, then the tests are built randomly over the elements of the patch. Depending on the outcome of the test, the path is sent to one children or another. From [55].

4.3 Facade Energy

In order to quantify a policy it is necessary to compute the energy. For each pixel of specific region R_s, where s is a semantic label such as door, window, balcony and so on. it is possible to compute a probability of belonging to a class c. Naturally, the belonging of the whole region R_c to a class c can be computed based on the sum of the probability of all pixels in this region:

$$p\left(R_c \in c\right) = p\left(x \in c, x \in R_c\right) = \prod_{x \in R_c} p(x \in c)$$

Using Bolltzman's transformation we can now compute the total energy of the region:

$$E\left(R_c \in c\right) = -\sum_{x \in R_c} \log p(x \in c)$$

The total energy of the whole policy is now:

$$E(\pi) = E\left(R_i \in c_i, \forall i\right) = -\sum_i \sum_{x \in R_i} \log p\left(x \in c_i\right)$$

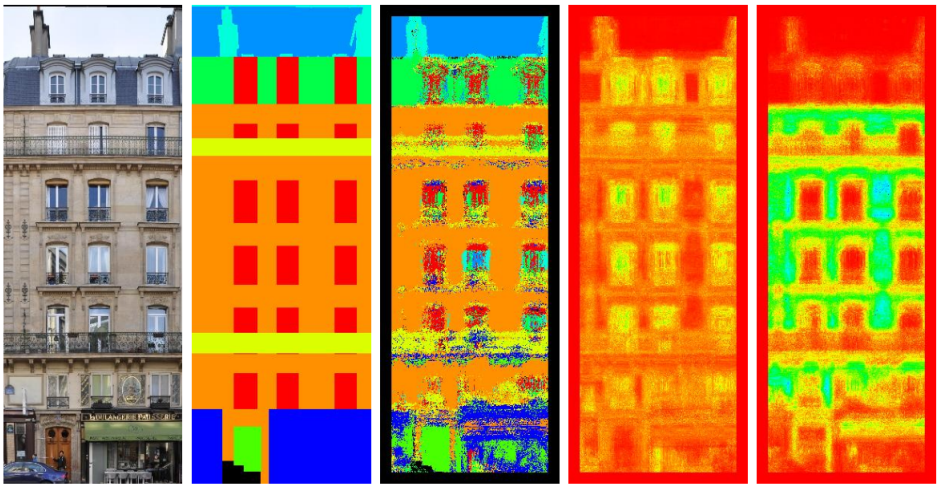


Figure 4.3: First picture: Original data; Second picture: Label image (by hand). Each color represents a label; Third picture: Most probable label; Fourth picture: Probability for each pixel to be a window; Fifth picture: Probability for each pixel to be a wall. The values increase from red (p=0) to blue (p=1). From [55].

4.4 Optimization Using Random Walk Algorithm

In order to get an optimal solution to find a minimal energy random walk algorithm is used. Algorithms start with an initial seed policy and randomly considers better policies in the neighbourhood in terms of the defined energy. If it does not find a better solution, seed is updated and continues on finding an even better solution. As the algorithm progresses the new optimal solution must be closer to the current seed, while at the beginning it must not be so. After around 50 iterations algorithm finds the optimal solution that is close to the minimum.

Chapter 5

The Pipeline Of 3D Building Exterior Creation

One of the aims of my thesis is to create prototypes of a pipeline, that would allow to reconstruct an exterior part of the building based on a set of images. This step is necessary for the future creation of 3D tactile maps of other buildings for visually impaired following the patent of Macík et. al. [33].

However, there is a requirement for the resulting building: the number of primitives should be minimized. The reason for that is the economy of material and time when printing the 3D model. And another reason is that when tiles of the 3D model are be printed, it is necessary to glue them to the wooden model, so having a model with uniform walls and a minimized number of triangles is essential. Another aim of my thesis is to impart into developing of the tactile 3D model of Palata, thus, the pipeline creation will be focused on this building.

The created pipeline consists of 8 steps (Figure 5.1). There's a possibility for some steps to be done manually, e.g. redundancy removal or picture adjustment in order to create a panorama. There's one particular technical step that can be done only manually: step 5 "Create building with straight walls". All pastel purple rectangles in Figure 5.1 are automated steps, the pastel green are manual.

The following sections will be focusing on each step individually.

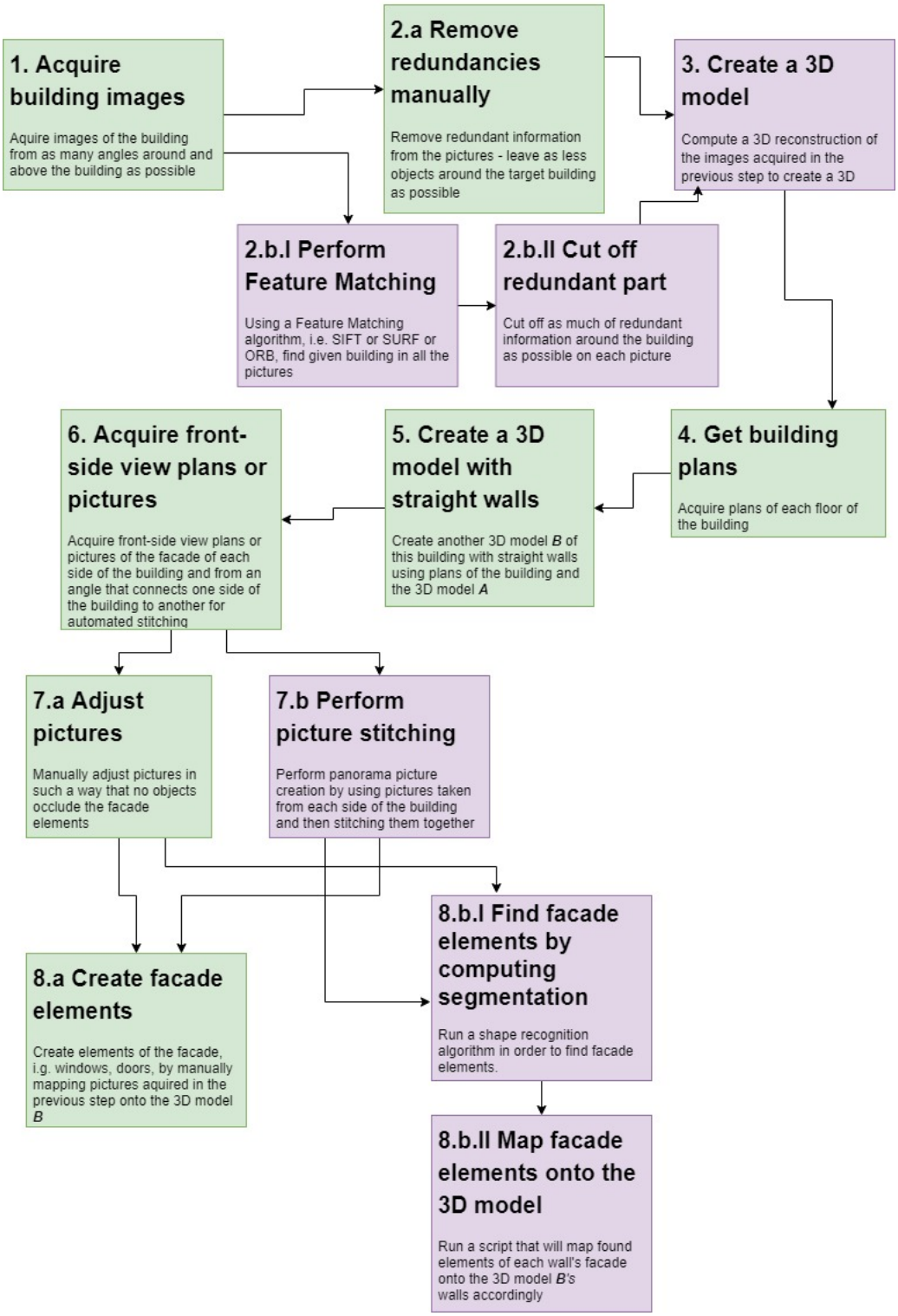


Figure 5.1: The pipeline of 3D building exterior creation. All pastel purple rectangles are automated steps, the pastel green are manual.

5.1 Input And Output

The input for this pipeline is:

- pictures of the buildings that are taken from every 3-5 degrees around the building and from the top,
- floor plans,
- front-side view pictures of each side of the building and pictures of transition from one side of the building to another, or facade plans.

The output is a 3D exterior part of the building with the minimized number of primitives (triangles) and with openings for facade elements (windows, doors, ledges) and without a roof.

5.2 Acquiring Building Images

In order to perform a 3D reconstruction of the building it is necessary to gain enough pictures from different angles around the building. I had pictures made from the drone in my disposal, but unfortunately the number of pictures was not enough, which pushed me to use *Google Earth Pro* in desktop version to collect the Palata pictures. [1] Google Earth already has every object reconstructed using the images from the satellites. As I have found it is nearly the best way to get images for building reconstruction, unless a drone is used.

I have successfully obtained 96 pictures of the building from different angles. Pictures were taken from every 3-5 degrees around the building and from the top of the building.

Figure 5.2: An example of two pictures of Palata taken from Google Earth

5.3 Feature Matching & Cutting Off Redundant Part

The aim of this step is to remove all objects, e.g. trees, other houses, pavement. etc., around the building in the pictures from the previous step. This stage is important, because in the next step, when 3D model is reconstructed, it is better to have as less additional information as possible to save time and memory.

This step can be done manually, e.g. in Photoshop. A way to automate this procedure would be to use some *Feature Matching* algorithm, like *SIFT* (Subsection 3.1.1), *SURF* or *ORB* [34]. The algorithm would find the building in all the photos taken in the previous step and remove everything around using an additional script.

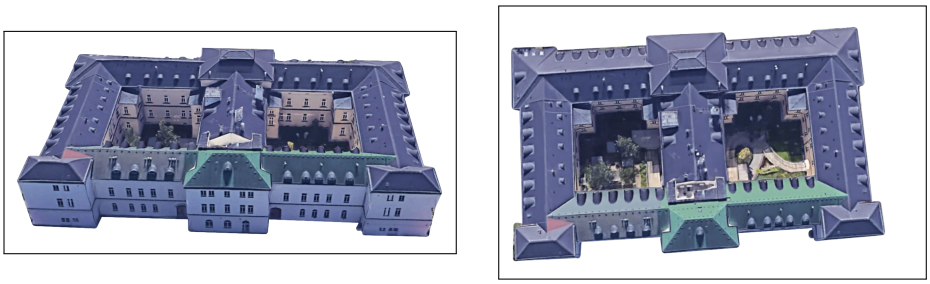


Figure 5.3: Pictures of Palata with no background

After the algorithm of feature matching is performed it is now possible to remove redundant information from the picture - preferably everything except for the building - for faster performance.

5.4 3D Model Reconstruction

After acquiring pictures if the building with no background it is now possible to create a 3D model. To make a reconstruction of Palata I used *Colmap in version 3.6-dev.3.* [48] 3D reconstruction was made on Lenovo Legion Y530 with the following hardware configuration:

- processor with Core i7-8750H (6C, 2.2 / 4.1GHz, 9MB)

- graphic card with NVIDIA GeForce with CUDA GTX 1050 Ti 4GB GDDR5
- Windows 10 Home operating system

It took around two hours to complete the reconstruction. After successfully obtaining the 3D model it might be necessary to cut off some unnecessary parts, e.g. my 3D model was occluded by a tree. Reconstructed 3D model can be found in attachments.

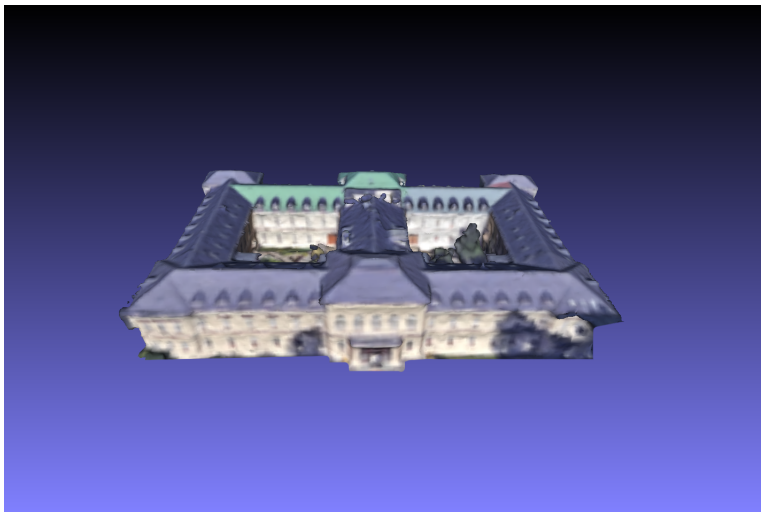


Figure 5.4: Reconstructed 3D model of Palata

5.5 Acquiring Building Floor Plans

Since the 3D model reconstruction does not provide perfectly shaped objects with a minimized number of triangles which leads to the obtained 3D model having bumps on the surface and having curved walls, it is necessary to change that. The studied possibilities of doing so using algorithms for triangle reduction in the model did not seem like a way out, since they do not give a "perfect" result, where each wall is represented as 2 triangles. In the end, it was decided to do it manually by creating a new 3D model based on the obtained reconstructed 3D model from the previous step in *Autodesk Fusion 360*. First step is to acquire building plans.

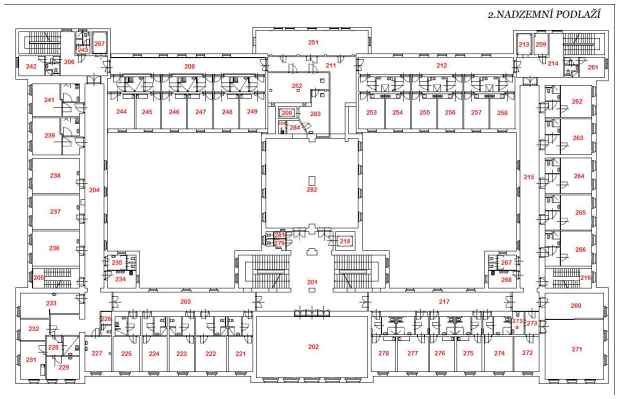


Figure 5.5: Second floor of Palata

5.6 Building A Simplified Model With Straight Walls

Next step of the building reconstruction is creating a model with straight walls and minimized number of triangles.

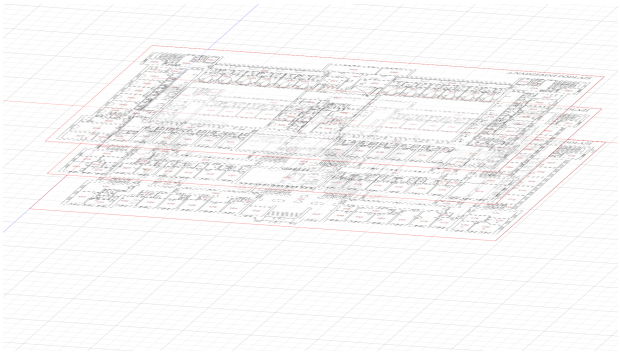


Figure 5.6: Floor plans corresponding to each floor

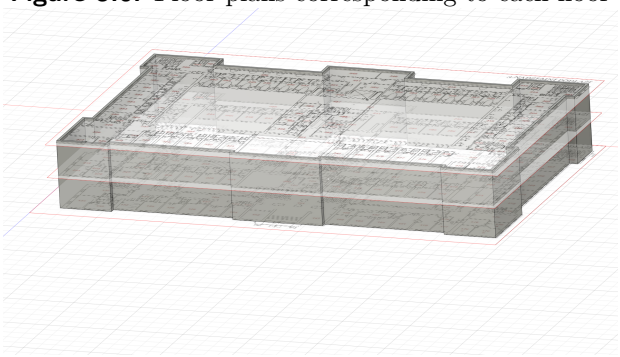


Figure 5.7: Extruded wall of Palata

I have put the building together in *Autodesk Fusion 360*. Each floor plan corresponded to each floor in a model. The mesh created earlier was used

to determine the distance between the floors and the scale. This model will further be used in the pipeline. The whole tutorial can be found in the High-fidelity prototype. For a person not familiar with *Autodesk Fusion 360* it takes around 1 hour to complete the tutorial. The Figures 5.6 and 5.7 briefly represent the workflow.

In the end of the pipeline this 3D model with a few changes (with facade elements - see next steps) will be printed and will be glued to the wooden model with the drawers created by Macík for their patent [33] implementation, so the sizes should correspond.

5.7 Acquiring Front-Side View Pictures/Facade Plans

After the new 3D model was built it is necessary to depict facade elements since these features correspond to external and internal user perspective. In order to create facade elements on the 3D model acquired in the previous step it is necessary to project a picture of each side of the building onto each side of the 3D model. Elements that repeat along the building facade should match on each side and maintain the same level. It is better if plans of the facade can be obtained, since the preciseness of the facade element's location will depend on this. Unfortunately, in my case it was impossible to gain the facade plans.

Figure 5.8: Front-side view picture of the front side of Palata. Picture taken from the drone.

If the next step is not performed manually it is necessary to acquire a picture dataset of the building facade. Pictures should be front-view of each side of the building and from the angle that connects two consecutive building sides. This is needed for stitching them together and computing a panorama. It is important to use a panorama because panorama will make the facade elements to be located at the same level, which is important for keeping the consistency of the building. If consistency of the building is not preserved, it might confuse the visually impaired users of the 3D model when exploring the model.

Figure 5.9: Front-side view picture of the back side of Palata. Picture taken from the Google Earth.

5.8 Adjusting Pictures Manually

For mapping pictures into the walls, windows and doors in this picture must not be occluded by any exterior objects. The angle at which the picture is taken should be converging to 180 degrees from the ground to the building. Fortunately, one of the pictures obtained from the drone was good enough. All I had to do was to reconstruct some of the occluded facade elements using Photoshop. Seeing that the building facade is symmetrical and the elements repeat was enough to copy an unoccluded window and align it with an occluded one on one side of the building. Then copy the reconstructed part of the building, mirror it by y-axis and paste it into the other side. This way all facade elements are detectable.

5.9 Picture Stitching

Front-side view picture of each side of the building will be stitched together. Stitching process uses an algorithm which is described before in the Section 3.1.1 for finding key points in the pictures. After that it is necessary to find a homography using 4 key points that will then be able to tell which key points are matched and which ones are not based on the error between the projection of a point in one picture to the point in another. Based on the threshold some matches are considered inliners and some other outliners. Then, each image couple is checked to find the percentage of inliners. If the percentage is greater than some number, e.g. 50, then this couple is considered to be a candidate for stitching. This way the adjustancy matrix is computed. Using strongly connected components in the adjustancy matrix clusters were created that can possibly form a panorama afterwards. In each cluster pictures are found with the most amount of keypoints that match, which in the end form a chain that results into a panorama. [23]

If it is not possible to take one whole picture of the facade, then you can

take many picture of the facade and stitch them together using e.g. the tool [43]. Then, if the building is very occluded by other objects, like trees, other houses etc., you can edit the resulting picture manually as it is explained in the Section "Adjusting Pictures Manually".

5.10 Manual Facade Elements Creation

Figure 5.10: Palata with extruded facade elements and an image of the facade located on the offset plane from the front side of the building

Next step is to map the gained picture into the walls. This step can be done either manually or automatically. The manual version can be performed by creating an offset plane (Figure 5.10) from one of the target walls from a small distance, e.g. *1mm* and inserting the modified picture there. After a few modifications for the picture to precisely fit the object the mapping was done. All facade element sketches were made in a rectangular pattern fashion so that each window of the same type does not differ in size or form and are symmetrical. The final step was to extrude the created sketch of the windows and doors to create openings in the 3D model. Each wall facade was performed in this same way while aligning the windows from one wall to another to keep the same level.

5.11 Facade Element Detection

In order to create the facade elements automatically it is necessary to perform a facade segmentation and detect elements using e.g. the algorithm explained in the previous Chapter 4. Windows, doors and ledges will be highlighted in the panorama or facade plans as vectors in different colors corresponding to

that specific type of facade element, e.g. windows are blue, doors are red and ledges are green.

5.12 Mapping Of Facade Elements Onto The 3D Model

Then, a script will be used that would take the facade plans or a panorama created from the front-side view images of each side of the building, slice it into parts that would correspond to the walls. Subsequently it would map each picture of a wall to a corresponding wall of the 3D model as a texture. After that doors, windows and ledges will be extruded from the building creating openings in the 3D model.

5.13 Result

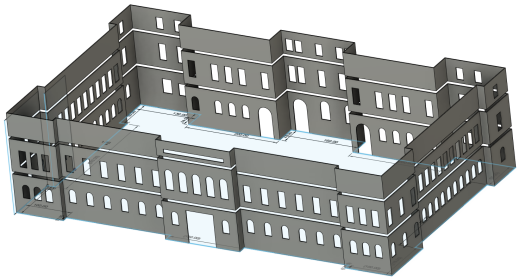


Figure 5.11: Resulting building

The result of the pipeline is a printable 3D model with a minimized number of primitives. The model has openings for the windows, doors, the sign in the upper central part and ledges. Detailed facade elements should be done manually, for those are important parts of the building that will be explored by the visually impaired and need to be done in detail and separately.

Chapter 6

Design & Implementation

One of the aims of this thesis is to design and implement a prototype for reconstruction of a 3D building exterior part from images. Thus, this chapter is dedicated to creation of low and high-fidelity prototypes following the UCD method.

The process of 3D building reconstruction from images is a pipeline, which means the output of one step is an input for the next one. There are several steps of the pipeline:

1. Prepare pictures of the building
2. Perform the removal of the background for each picture
3. Create a ununiform 3D model based on the pictures acquired from the previous step - *Model A*
4. Prepare floor plans
5. Manually create a model with minimized number of triangles based on the previous 3D model and floor plans - *Model B*
6. Prepare front-side view images of each side of the building and transition from one side of the building to another/facade plans
7. Find facade elements (doors, windows, ledges) on these pictures
8. Map the elements onto the *Model B*

6.1 Low-Fidelity Pipeline Prototype

Prototype was done in Balsamiq Wireframes [3] application as a pdf file. Pdf format allowed for a prototype to be semi-interactive, since it allowed links from one page to another. Each step in the created low-fidelity prototype is made with a backtracking option, so that users can change some data in the previous steps. The prototype can also simulate load of user's data in step 3, 8, 9 and the download of the data that was created by the program. In step 6 users can decide whether to use facade plans or their own pictures, so the pipeline forks. There are breadcrumbs located on the top of each of the pages to get the user a sense of progress. Each button is relatively big and is labeled with a symbol that lets the user associate the action with its outcome. Each title has a size of *40ptx*. The smallest font size is *13ptx*. Important information for a user is colored in red, so it catches the eye. See the final version of the low-fidelity prototype in Appendix F.

Figure 6.1: Low-fidelity prototype: step 2 completed page

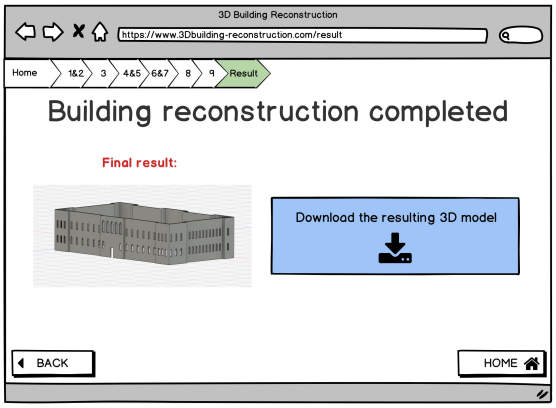


Figure 6.2: Low-fidelity prototype: result page

6.2 High-Fidelity Pipeline Prototype

High-Fidelity prototype is an interactive prototype of the final product, that can seem like a real program to its users. In my case, the high-fidelity prototype is a website, where certain interactive features are available, e.g. data can be downloaded and uploaded, 3D model is visualized in 3D and is rotatable. Buttons that indicate the normal flow of the program are green, and the ones that represent a non-obligatory action are grey. The website was made in SitePad. [2] Certain scripts, e.g. adding a feature with rotation of 3D model or upload and download, were written by me, since SitePad offers a very limited service.

Visit the web `alessandroalei.com/palata/` to see in detail.

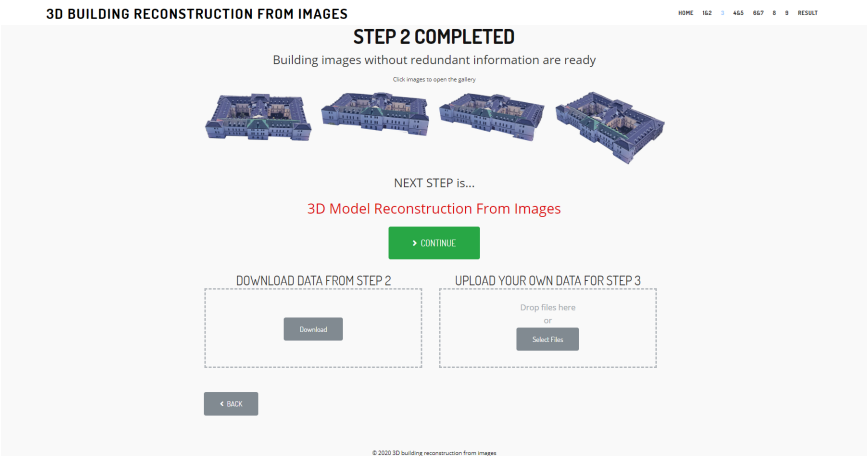


Figure 6.3: High-fidelity prototype: step 2 completed page

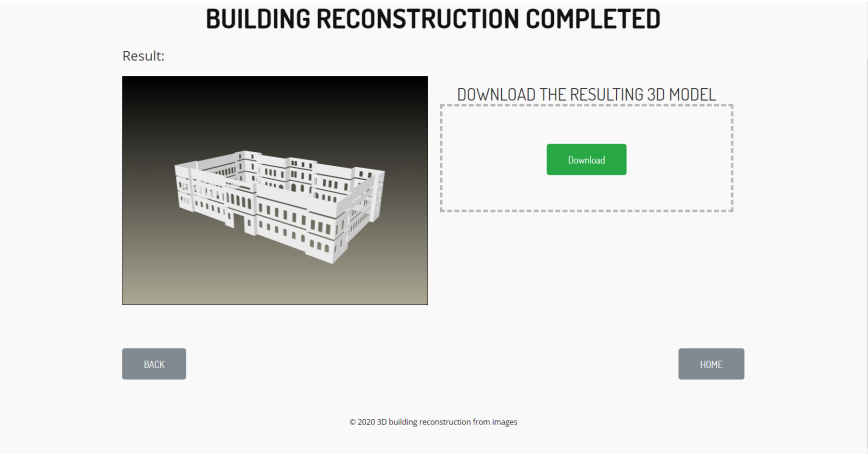


Figure 6.4: High-fidelity prototype: result page

6.3 3D Model

6.3.1 Requirements

The printed 3D model that is then glued to the wooden model should follow a few requirements:

1. Safety *The model should not have sharp edges. There should not be parts that can harm people. Users should not be afraid to use it.*
2. Stability *The 3D model should not shake or shiver when is touched. Tiles of the model should be glued well to the wooden model and to each other.*
3. Endurance *The material and its thickness should be chosen in such a way that the model is not be easily damageable.*
4. Tactile readability *The model should be readable by users: all parts should be detectable, all ledges should be pointing in user's direction, the ratio of the real building to the model doesn't have to be strict, e.g. windows can be proportionally bigger, than in actual building, some building attributes, like rainwater drain or lightning rod shouldn't be depicted in a 3D model.*
5. Universalisability *The model should be made in such a way that not only people with different types visual disability (blind, low vision, etc.) can use it, but also the sighted people. The color of the 3D printed model should be similar to the real building's color, same goes to its details, like windows, doors, etc. For the low vision and sighted users the windows could have an option to be lighten up.*
6. Hygiene *The model should use such a coating that would allow for its easy washing. The coating should not be poisonous.*
7. Usability *The model should satisfy the needs, limitations and requirements of the visually impaired users.*

After the model is ready: the exterior of the building is glued to the wooden model, all the interactive floors are glued to the drawers, the audio system is installed, etc., and overall it satisfies all the patent requirements [33] it is ready to be released for exploration for the target users, e.g. in a residential care institution Palata. The created tactile 3D map now is ought to satisfy not only the requirements mentioned above, but also the following:

8. Ergonomical accessibility *The tactile map should be accessible to any user from any side. There should be enough space around the model for one or even multiple people to be exploring it at once. If there's some kind of an auditory system, it should not be blocked by people accidentally touching some part of the model.*

6.3.2 Measurements

The 3D model created has an offset equal to *1.5mm* from the sizes of the drawers, which will indicate the wall width. The sizes are listed in Appendix A, the building is symmetrical.

The detailed measurements of the facade elements are displayed in the following tables.

The square windows are represented by the four types. Type 1 is located across the the whole perimeter of the second floor of the whole building. Type 2 is located on the back side of the building in the third floor. Type 3 and 4 are in the back side left and right parts in the first floor. See the Figures 6.5 and 6.6, where the windows are highlighted in blue.

Object	Width	Height
Square windows type 1	15mm	35mm
Square windows type 2	15mm	25mm
Square windows type 3	20mm	30mm
Square windows type 4	20mm	25mm

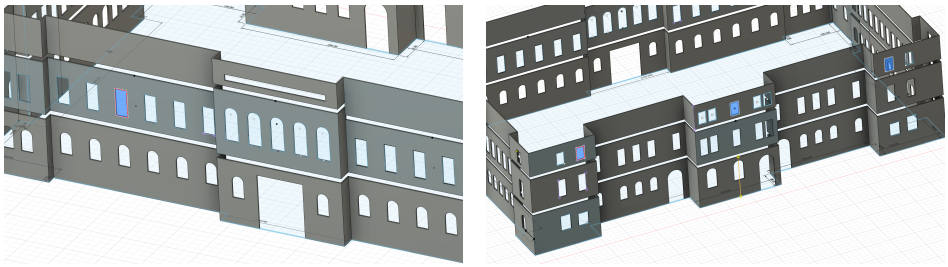


Figure 6.5: First picture: Squared window type 1; Second picture: Squared window type 2

The arc windows are of 3 types. The type 1 is all around the first floor of the building, except for the back central side, where type 3 is located. Type 2 is in the front center second floor.

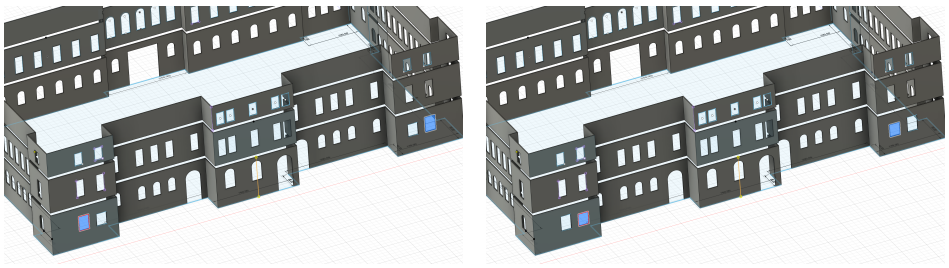


Figure 6.6: First picture: Squared window type 3; Second picture: Squared window type 4

Object	Width	Height	Arc Radius
Arc window type 1	15mm	21mm	7.5mm
Arc window type 2	17mm	35mm	8.5mm
Arc window type 3	20mm	30mm	10mm

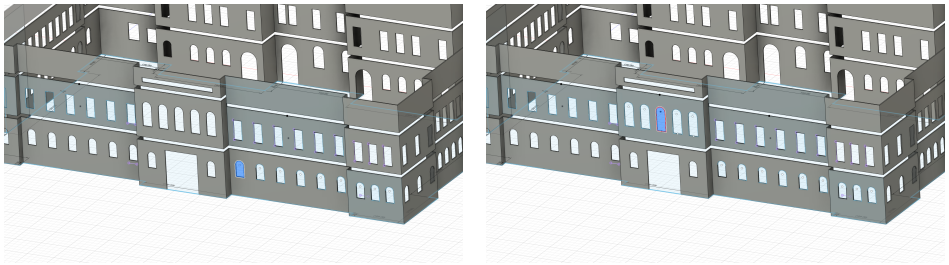


Figure 6.7: First picture: Arc window type 1; Second picture: Arc window type 2

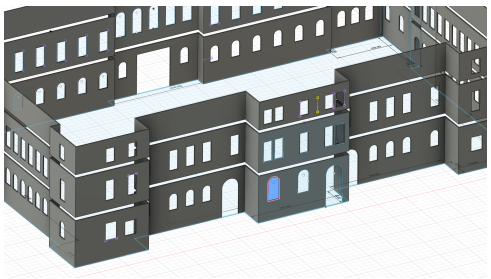


Figure 6.8: Arc window type 3

There are two types of doors. The front door, that is located in the front central side. And the doors in the back that have an arc.

Object	Width	Height	Arc
Front door	60mm	70mm	–
Back door	30mm	50mm	15mm

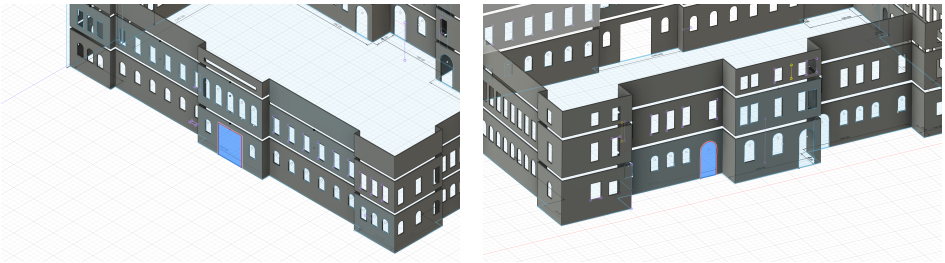


Figure 6.9: First picture: Front door; Second picture: Back door

6.3.3 Printing & Gluing The Model

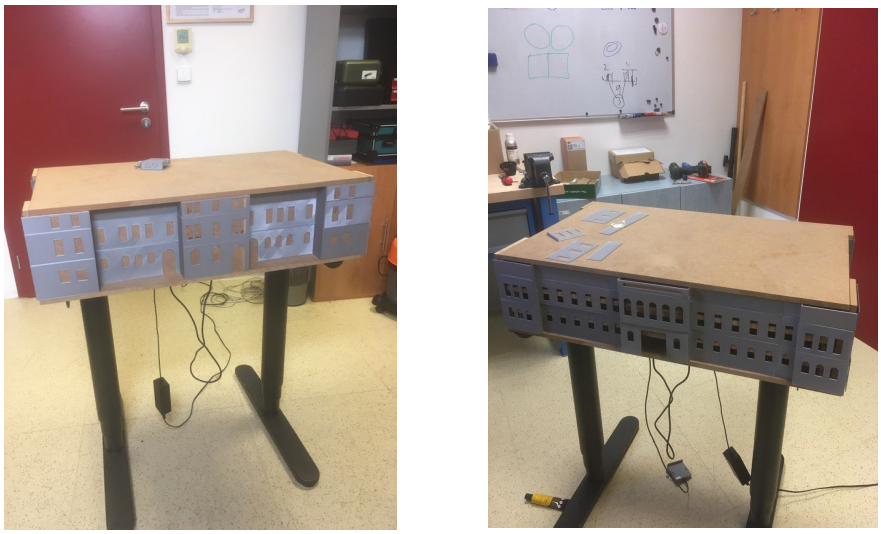


Figure 6.10: First picture: Back side; Second picture: Front side

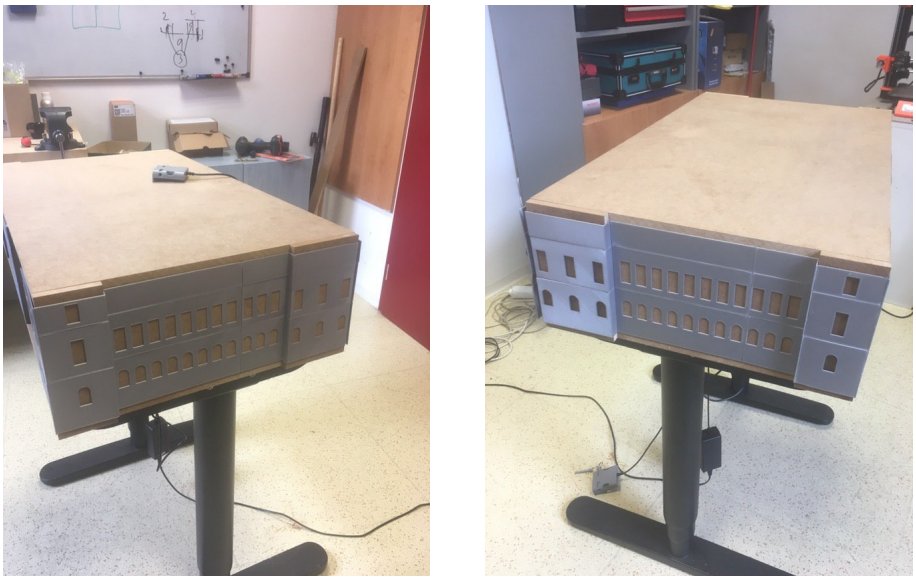


Figure 6.11: First picture: Left side; Second picture: Right side

3D printing consists of the following steps:

1. modelling the object,
2. slicing the model into tiles to fit the 3D printer's bed,
3. printing.

For 3D printing I have used Prusa i3 MK3S 3D printer with a *0.4mm* nozzle diameter and *0.30mm DRAFT* print settings. Since the bed is limited in size, I have sliced the model into tiles of maximum $20cm \times 20cm$ sizes. I have used Polyactic Acid (PLA) as a filament. The temperature of the extruder was 230°, and the bed was 90°. The 3D model was printed in *HW lab* in Karlovo Náměstí building E of the Czech Technical University in Prague.

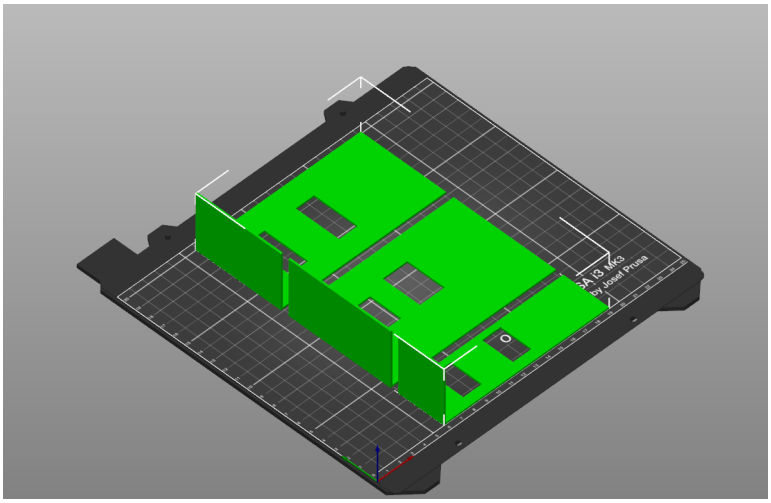


Figure 6.12: Example of a tile

There were 20 tiles to print. The time for printing one tile was from *1h 30min* to *3h 40min*, so, in the end, I spent 5 days, 10 hours each day in the laboratory.

For gluing the tiles to the wooden 3D model I have used SuperGlue of the brand *Herlitz*. The gluing took around 6 hours.

Find the 3D model in fd3 and obj formats in attachments.

Chapter 7

Evaluation

7.1 Low-Fidelity Pipeline Prototype

7.1.1 Hierarchical Task Analysis

Hierarchical Task Analysis (HTA) is a method to analyze the flow of the program, how tasks are accomplished. Goals are divided into subgoals and are assigned a certain level. The highest level is 0. HTA allows detection of flaws and unnecessary steps. This method allows us to illustrate more than one way of the program flow. Please see the diagram in Appendix B.

Looking at the diagram produced due to Hierarchical Task Analysis it is clear that some steps can be made in two different ways. Looking at point 3 in the hierarchy we can see that users can follow with the natural program flow or they can download their own data to proceed in step 3. In point 5 users can decide whether they use facade elements or pictures, which also creates a fork in the hierarchy.

7.1.2 Cognitive Walkthrough

Cognitive walkthrough is a usability walkthrough technique that can be used in the development stage of a product when design of a product is known. This technique is based on a theory that users can understand the product while exploring it rather than having to attend some special course. The users of the product are IT developers and architects. The whole pipeline will be tested with several tasks that will cover the usability of the whole process. Each task is being evaluated by me using the following questions:

1. Will users understand how to start the task?
2. Are the controls conspicuous?
3. Will users know the control is the correct one?
4. Was the user able to complete the task? [36]

For the detailed evaluation see Appendix C.

First round of evaluation

Task one

Create a model of the building with facade elements using the given a reconstructed model A and sets of pictures:

a pictures of the building taken from different angles around,

b plans of each if the floors,

c facade plans

Task sequence:

1. Upload images of the building from as many angles around as possible *a*)
2. Perform feature matching & cutting off redundant information

3. Perform a 3D model reconstruction
4. Upload facade plans *b)*
5. Perform picture stitching
6. Perform searching of facade elements in the panorama created in the previous step
7. Perform mapping facade elements from the picture to a 3D model with straight walls on data created in the previous step
8. Download the final result

Task two

You are given:

a pictures produces in step 2,

b pictures of the building taken from different angles around

From the Home page go to Step 3 and download the 3D model. Then go back and upload pictures *a)* and compute Step 3 again.

Task sequence:

1. Upload images of the building from as many angles around as possible *a)*
2. Perform feature matching & cutting off redundant information
3. Perform a 3D model reconstruction
4. Go back and upload pictures *a)*
5. Perform a 3D model reconstruction

Summary

After a thorough analysis of round one of cognitive walkthrough the following mistakes have been found:

1. Task two was wrongly formulated. Task formulations are not supposed to consist of detailed explanation.
2. When step n is completed in the low-fidelity prototype it is not clear whether actions such as *"Download data produced in step n"* or *"Upload your own data for step (n+1)"* are mandatory to complete by the user. Is it necessary to download my own data each time step like that is complete or is it enough to just click *"CONTINUE"*?
3. Visual representation of what to expect after each completed step is missing.
4. In step 6 there's a fork. *"YES"* is located on the left side, which is not a usual location for a positive user response.
5. Could be nice to provide users with some colorful buttons along the pipeline so that it is easier to orientate.
6. *"BACK"* button in each step is too big and is located in the middle of the page. Backtracking is an option, but not the main element on the page.

Second Round Of Evaluation

Following the notes regarding what to improve the following changes came into place:

1. Task two was reformulated to:
   "You are given:
   a pictures produces in step 2,
   b pictures of the building taken from different angles around
   Download the reconstructed 3D model, then reupload a new set of pictures and download a new reconstructed 3D model."
2. When step n is completed in the low-fidelity prototype now *"CONTINUE"* button is in the middle and is colored blue so that it is more clear where the default program flow is. An additional information was also added to the button indicating that no additional actions are needed to proceed with the default program flow. *"DOWNLOAD"* and *"UPLOAD"* buttons are moved below and are left uncolored to indicate that those are optional. Representation pictures were added as well as their label. *"BACK"* button is smaller and is located in the left down corner to reduce its importance.

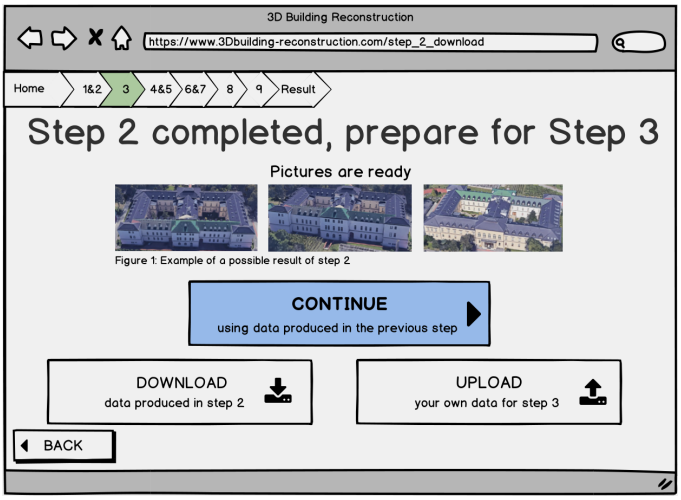


Figure 7.1: Example of a new *"Step 2 completed"* page After round 2 of Cognitive Walkthrough

Figure 7.2: Examples of what users should upload and the number of uploaded pictures is added after round 2 of Cognitive Walkthrough

Summary After a thorough review of the round two the following notes have been taken:

1. Visualization of data to be uploaded next to upload bars.
2. Give feedback on how many pictures were uploaded next to upload bars.
3. Process progression when subtask is computed should be added.
4. Instead of an example of what is some step's result should be the actual result.

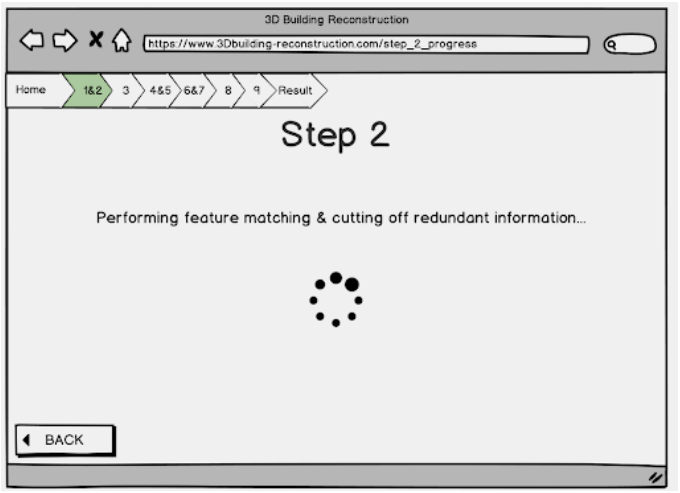


Figure 7.3: Each step that requires some time to complete now has a progression page after round 2 of Cognitive Walkthrough

5. After step 2 completed there should be a possibility for a user to manually edit the pictures.

6. Step 3 completed page should only contain the visualization of a reconstructed 3D model, the actual download should take place in the next step, so the user understands why they download it.

7. There was a mistake found in step 4&5: after the manual creation of the model with straight walls there was no way to upload it in the program for further usage. There should be an upload page and further visualization of the 3D model.

8. In step 1 it is necessary to give user more specific information on what they should upload, e.g. instead of *"Upload images of the building from as many angles around as possible a)"* there should be *"Upload images of the building taken from every 5 degrees around it a)"*.

Figure 7.4: First picture: A feature to edit pictures is added after round 2; Second picture: Reconstructed 3D model can now be downloaded in step 4&5 after round 2.

Third Round Of Evaluation

After the round 3 the following problems have been found:

1. Stop asking for user's permission to compute a step.
2. Make main buttons different color, e.g. green.
3. Change the manual, add more detailed information, add pictures.
4. Step 6 should give user detailed information of what he is going to work with in the next step, rather than *"Yes"* or *"No"*.

The issues above will be fixed in the high-fidelity prototype.

7.2 High-Fidelity Pipeline Prototype

The high-fidelity prototype that I have created will be tested using a qualitative approach of Usability Testing technique which will involve 6 participants including the pilot test participant. While testing the prototype there will be a moderator and a participant. I will be in the role of the moderator.

7.2.1 The Goal

The goal of this testing is to find out whether the created high-fidelity prototype works properly, as the target user would expect it to work. It should be clear to a user how to start, proceed and finish the task of the 3D building exterior reconstruction from images. Another goal would be to find the flaws of the prototype and fix them.

7.2.2 Target Audience

The target audience for this testing would be programmers, 3D model designers, architects, civil engineers, or students studying these fields, and anyone

with basic to proficient computer skills. The age can vary from 15 to 60. The person should have English language level knowledge from B1 and up, as the prototype is written in English. The screener and participant's answers are attached in the Appendix D.

7.2.3 Used Hardware

For this usability testing I have used my own computer and mouse.

Computer model : Lenovo Legion Y530
Operating system : Windows 10 Home
Processor : Core i7-8750H (6C, 2.2 / 4.1GHz, 9MB)
Memory : 8GBx2 DDR4-2666 MHz
Hard drive : 1TB 7200rpm
Hard drive 2: 256GB SSD M.2 PCIe NVMe
Optical drive : Without optical drive
Display : 15.6 "FHD (1920x1080) IPS 250nits 60Hz
Graphics card : NVIDIA GeForce with CUDA GTX 1050 Ti 4GB GDDR5
Pointing device : TouchPad

Mouse Model: HP Wireless Mouse Z3700
Connection: Wireless
Interface: USB, Radio frequency (with USB receiver)
Sensitivity: 1,200 DPI

7.2.4 Used Software

At a certain step of the testing participants will use *Autodesk Fusion 360 2.0.8624* installed on the computer. In order to use the high-fidelity prototype, that is in the form of a web site, users will use *Chrome 3.0.4103.116 (Official Build) (64-bit)*.

7.2.5 Initial Application State

Users will be provided with an opened Home page of the prototype in Chrome and a running Autodesk Fusion 360 with an empty design page. Each

participant will be provided with a paper with the tasks and prepared data, such as:

1. Images of the building taken every 5 degrees around the building and from the top - *package a)*
2. Plans of each of the floor - *package b)*
3. Front-side view images of each side of the building and pictures of the transition from one side of the building - *package c)*
4. Facade plans - *package d)*
5. 3D model of the building with the straight walls - *package e)*

7.2.6 Workplace Layout

The software that is to be created out of the high-fidelity prototype will most likely be used in company offices, but due to the inability of testing in such an environment the testing will take place in the houses of the participants or in a quiet public study cafe, which is a good alternative.

7.2.7 Use Cases

1. Create a 3D model of the building with facade elements(windows, doors, ledges) using the given
   a. Images of the building taken every 5 degrees around the building and from the top - *package a)*
   b. Plans of each of the floor - *package b)*
   c. Front-side view images of each side of the building and pictures of the transition from one side of the building - package *c)*
2. Upload your own data for Step 3. Perform this task using:
   a. Images of the building taken every 5 degrees around the building and from the top - *package a)*
   b. Images of the buildings with no redundant information - *package f)*
3. Open the gallery of the pictures of the buildings that have the redundant information removed. Perform this task using:

a. Images of the building taken every 5 degrees around the building and from the top - *package a)*

7.2.8 Use Cases Action Sequence

Use Case #1

1. Upload images of the building taken every 5 degrees around the building and from the top - *package a)*
2. Perform feature matching & cutting off redundant information
3. Perform 3D model reconstruction
4. Upload 3D model of the building with the straight walls - *package e)*
5. Upload front-side view images of each side of the building and pictures of the transition from one side of the building - *package c)*
6. Perform picture stitching
7. Perform searching of facade elements in the panorama created in the previous step
8. Perform mapping facade elements from the picture to a 3D model with straight walls
9. Download the final result

Use Case #2

1. Upload images of the building taken every 5 degrees around the building and from the top - *package a)*
2. Perform feature matching & cutting off redundant information
3. Download data produced in the step 2
4. Upload this data for the step 3

Use Case #3

1. Upload images of the building taken every 5 degrees around the building and from the top - *package a)*
2. Perform feature matching & cutting off redundant information
3. Open the gallery

The perfect progress: the participant will follow each use case step by step as they are being told and will be able to figure out which data they should use in which step. Participants will have no problem navigating the web.

7.2.9 Pre-Test Interview

1. Have you ever encountered/heard of an application for a 3D model reconstruction from images?
2. If yes, did you like it? What was good and bad about it?
3. What do you expect from such an application?
4. Would you ever want/need to use this kind of application?
5. If you have ever modeled 3D objects in a 3D modelling environment, tell more about this experience. How many times have you done it and which program have you used? Do you still do it now?

7.2.10 Post-Test Interview

1. What did you like and what you did not like about this application?
2. If you are able to compare this application to something similar that you already know what you like and did not like about this application?
3. Can you comment on the user interface design of the application and the controls?
4. What would you change about the application and why?
5. Would you recommend the program to your colleagues or friends?

7.2.11 Testing Sessions

Each participant was asked to be recorded with a voice and screen recorder. Each of them was also asked to think out loud. The participants were explained what this application is about and were notified that it is a high-fidelity prototype, not a real application. Each participant was provided with a computer with opened Chrome with a high-fidelity prototype web page, 3 folders with necessary data to complete the three tasks and an opened Fusion 360 with a blank designing page.

In total there were 6 participants. The minimum age was 24, the maximum was 39. The average is 28.8 years. There were two females. Most of the participants came from different areas of IT: front-end development, computer vision, machine learning. There was also an architect and a movie producer. Two of the participants use macOS, everyone else uses Windows operating system. Half them were familiar or heard of 3D model reconstruction from images, and half had an experience with modelling of 3D objects. Everyone of the participants speaks English on the level higher than B1. Only one participant did not have a higher education degree. All of them use computer every day.

The pilot session is provided below. The prototype was slightly changed after the pilot testing, since it was found that the manual page of the prototype does not contain enough information for a successful task completion. Find all testing sessions in the Appendix E.

Pilot Test

Participant's profile

Participant is a male, 27 years old and is Spanish. He speaks English with the level C1 and is a proficient computer user - he uses a computer every day. The participant is a front-end designer. He has a Bachelor's degree in Multimedia. The participant had an experience with creating 3D models a long time ago. He has heard of an application for 3D model reconstruction from images.

Testing environment

The testing took place at participant's home on Friday 17.07.2020 at 16:00.

He was alone at home, so it was a quiet environment. Testing took 1 hour and 8 minutes to complete. The participant agreed to his speech being recorded with a voice recorder and for the screen of the computer to be recorded while testing.

Pre-test questionnaire

The participant has heard of an application for 3D model reconstruction from images, but never used one. He expects some kind of a virtual reality application that based on the images can recreate how an object looks like and he can visit it. He would use this kind of application for his personal purposes e.g. for recreating some ancient buildings to see how they look like in 3D. He has modeled 3D objects long time ago a few times and he doesn't remember the name of the modeling environment.

Session

Use case #1

Step 1: The participant did not know if he should upload all pictures that he was given in the *folder a)* for step 1. He wasn't sure if pictures were uploaded - he didn't see the indication.

Step 2: In step 2 he thought that the pictures were transformed into the 3D model. In step 2 completed page he thought he had to download the data, and upload his own data for step 3.

Step 3: Participant complained that he doesn't understand what the step 3 is. He thought that the result of step 3 has nothing to do with the pictures he uploaded before.

Step 4&5: He didn't understand what is the final step of the whole application. He didn't know what he is doing. The participant didn't see a link to the manual - "doesn't look like a button". He didn't understand that the reconstructed model is already good enough and why does he need to create a new model.

Manual page 1: I navigated him to the manual. The participant didn't read anything or didn't understand.

Manual page 2: The participant didn't read pro tips - maybe this info is too soon. He didn't know which plane is *XZ*, didn't understand that each time he creates/selects/edits something and a window appears he needs to click *"OK"* to

submit the changes or *"CANCEL"* to close, otherwise he can't do any other action. He didn't know where is the infinity of the Z axis, because he couldn't locate the Z axis. When the model is downloaded in step 4&5 it is in the glb format, but it's not supported by Fusion 360, so I had to provide the obj format. After going into mesh edition mode he didn't know that he should confirm the window, so he couldn't do anything. the participant didn't know how to scale/move/rotate the building, so I had to show it. When scaling he had to pull the arrow, but he didn't see it. He moved the building around the place where the border is, but he didn't rotate or move it in the Z axis. the participant didn't know he must click finish mesh to exit the editing mode. When creating the offset plane he couldn't locate the right floor since the camera was turned in a way, from where it was impossible to see. Then he realized that the building was not in place, he didn't seem to care. The participant pointed out that there is no point to put a picture for the second floor in the second floor plane. When selecting the high at which second floor was located he was not using the 3D model, he just measured it randomly. When creating the sketch of the building border he missed a few spots, so in the lines were not connected, because of this he was unable to extrude the walls. When he wanted to exit the manual and he got back to step 4&5, he thought it's a new step and he was confused. He said that manual should be another step rather than to be separated from the pipeline.

Step 6 Front-Side: The participant was not sure if he should upload all pictures that I have given him.

Step 7 completed: The participant thought he needs data from the previous step, so he used the download and upload button. He didn't realize that he doesn't need to do anything to proceed. He wasn't sure that the panorama shown in the step 7 completed is the same as he uploaded earlier.

Step 8 completed: Same comments, like in the previous step.

Result: The participant thought that he must press the Home button to proceed.

Use case #2

Since the participant already knew the interface, he did the same actions as in the previous use-case.

Use case #3

The participant got to the step 2 completed page, opened the gallery and said that it was not clear to him what these things are and how they were created. He didn't associate those with the pictures that he uploaded a step ago. He assumed those were some kind of 3D models.

Post-test questionnaire

The participant did not like the application, the experience wasn't that good. He felt lost, he lacked info, the info wasn't clear. He compared my app to a buying a ticket application: "In the ticket buying app every step is defined, there's only one thing you can do at one step, you know what's happening in every step, and that's what I lack here".

He wanted to be provided with an explanation to what is happening, doesn't want to be given an option to upload his own data for each step. He wants to have just one path, so he doesn't get lost. "The buttons are below so it's confusing - why are they there? " - asked the participant. He thinks that steps that are bonded together don't make sense and bring confusion. He thinks I should remove the possibility to go directly to all steps except 1, because it doesn't make sense since it's a pipeline, so to remove it from the header. He thinks an explanation of each step is needed. Manual button should be more like a button and less like a text box. *"HOME"* button on the result page should be deleted. *"BACK"* button should be *"Previous Page"*. Manual should be more precise and more pictures with red signs that point to the used controls in Fusion 360. In the step where the new 3D model with straight walls is created, it is not clear why to the participant. "Explain to the user that the reconstructed building doesn't have straight walls, " - said the participant. He said he would not recommend this program to his colleagues in the state that it is now. But if it every gets better, then the idea is very interesting and could be used in many fields.

Summary

The participant was mostly lost in the process. In each step where he saw the upload and download button he thought he needed to use it. He also did not know what is going on in each step and what is the result. The participant did not associate the figure that says "result" with the result. He had no idea where he was located in the process and what is going to happen next. The hardest part was the manual: it did not provide enough information for a user to navigate and to be able to easily perform each step. After this testing session a several improvements have been concluded about the web pages:

1. Specify what to do in each step, explain what it does and explain the data that are needed for that step.
2. *"BACK"* button on each page should be *"Previous Page"*.
3. Maybe consider removing downloading and uploading in each step.
4. In step where a new 3D model is created explain why.
5. Header menu should be removed, instead in each step there should be some kind of explanation where you are in the pipeline.
6. If the upload button stays, the indication for the amount of files downloaded should be more visible.
7. Manual link should be in a button.

8. When downloading the 3D model in step 4&5 it should be in obj format.
9. Manual should be its own step rather than a branch outside of the pipeline.
10. Create a new step 7 completed page for front-side view panorama, so it doesn't confuse the users.

The manual should be improved in the following way:

1. The *"Pro Tips"* in the manual should appear when the user needs them instead of the beginning.
2. Specify using the images with red indicators and text the following:
    a. Axis, planes - where to find the names for each of them when user hovers
    b. Each time user has to use any kind of menu show it how to get there
    c. If user needs to edit canvas/object/plane/sketch show how to do it
    d. User is in edit menu - let them know how to exit the menu
    e. Every time new object/canvas is added to the scene and a small menu appears - show the user that they need to click *"OK"*
    f. If something is hidden in the left-side menu it should be shown
    g. When scaling/rotating/moving make sure even the small controls are highlighted
3. Each small step should be written on its own instead of writing everything in one step
4. Each small step should be depicted with red indicators and text
5. Less words and more pictures
6. Provide more information on how to scale/rotate/move the 3D model when it's uploaded
7. Remove manual for inserting the second and third floor plan - it's pointless
8. Point out that all lines should be connected in the sketch of the building border

7.2.12 Overall Usability Testing Summary

After the testing sessions many key issues have been found regarding the user interface of the high-fidelity prototype and also about the manual. To summarize, the following changes should take place regarding the user interface:

1. Upload and download buttons should be indicated to be non-obligatory.
2. Indication of download and upload action should be more clear. There should be a bar that shows the progress of upload.

3. Names of the steps should not be confusing a common user, e.g. *"Computer feature matching Remove redundancies"* should be renamed to *"Remove background from the pictures"*.
4. Each step should have a pipeline schematic illustration.
5. *"BACK"* button on each page should be *"Previous Page"*.
6. Download of 3D model should be in the manual.
7. Manual should be a part of the pipeline, instead of a fork.
8. There should be a radio form that lets the user chose if they want to upload their own data or use the result produced by the program. If the user chooses their own data, then an upload form appears.
9. Steps should be removed from the header menu.
10. The picture gallery should have a grid and a button to enter.
11. Steps should be renamed in such a way that they do not confuse the user, e.g. *"4&5"* can be renamed to just *"4"*.
12. Step 4&5's name can be changed from *"straight walls"* to *"simplified"*.

Then, there are a few things that can be improved about the manual:

1. Steps in the manual can be changed like this: text in the left and a picture in the right.
2. Rename manual's name from *"straight walls"* to *"simplified"*.
3. More images, less text.
4. One step - one action.
5. Images should have highlighted controls.
6. Each modification confirmation of canvas, mesh or sketch should have its own step in the manual.
7. Showing/hiding canvases, meshes and sketches should be highlighted in the manual.
8. Each rotation of the camera should be shown in the manual.
9. Each scaling, rotation and other transformation should be shown in the manual in details.
10. The first step of creating the model should be unhiding the origins, since by the default they are hidden.
11. Clicking *"OK"* should be it's own step and should have a picture with an illustration.
12. Left-side menu should be called a browser.
13. *Z-axis* and *XZ-plane* should be highlighted in the manual.
14. It should be explained with pictures how to use lines when creating a sketch.
15. Add a description on how to remove snapping when moving objects.
16. When user's supposed to be right-clicking it should be specified.

I have created an illustration in vector format in a pdf file, so it can be used for the future work as an illustration if the pipeline located in each step of the process. Find it in Figure 7.5 and a pdf in the attachments. It has changed names of the steps and descriptions.

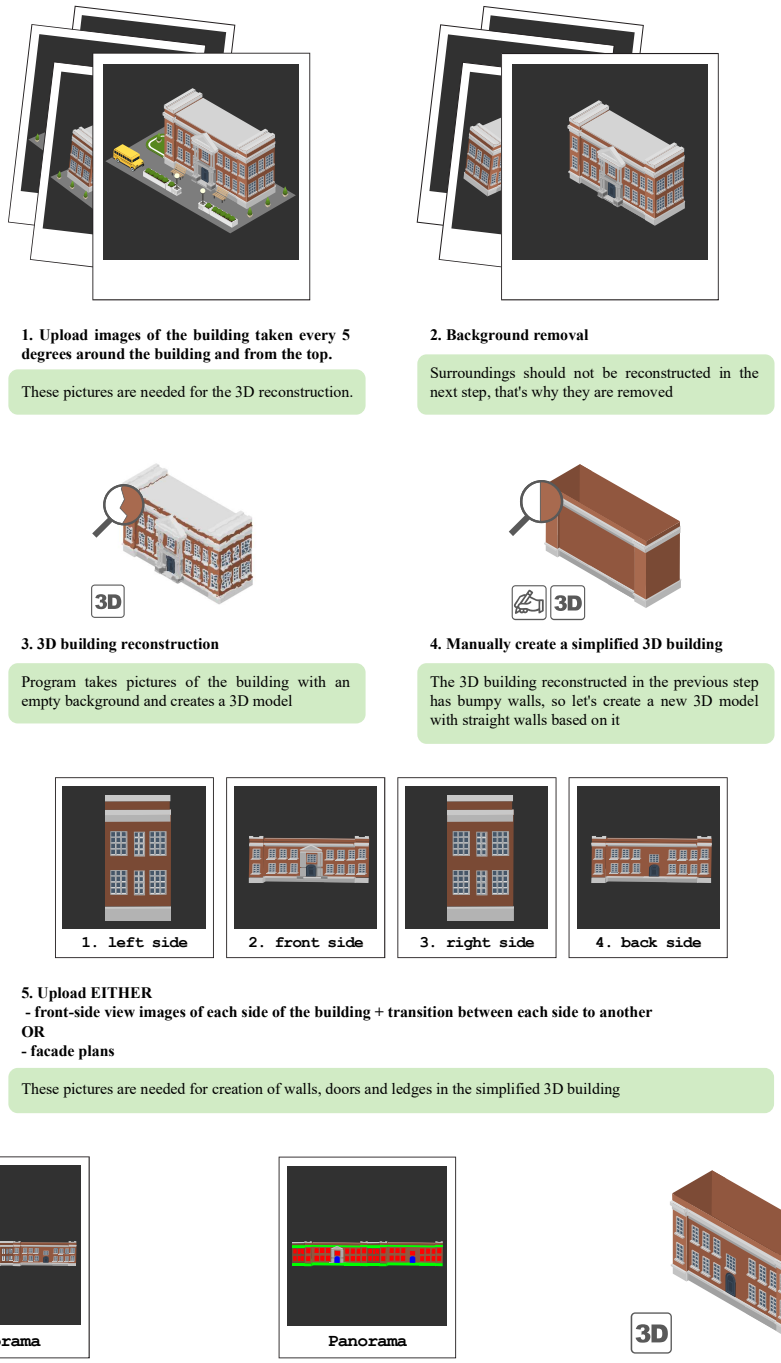


Figure 7.5: Pipeline schematic illustration

7.3 3D Model

The printed 3D model was evaluated by a visually impaired expert Lukáš Treml. Unfortunately, testing with actual Palata clients was not possible due to the Covid-19 situation. The session took place in the same lab, where the model was printed, and lasted for 52 minutes.

7.3.1 Front-side Part

The front-side part of the model should not block for drawers opening. The tiles should be glued firmly to the wooden part. The edges of the tiles that stick out, e.g. the ones on the corners of the front-side should be smoothed. If windows with an option to turn the light off and on will be implemented, then the drawer desk will block it. The front-side floors should have, e.g. not more than *1mm*, distance between each other, so that they do not lean on each other and allow for the drawers to open. For the visually impaired a small distance will not be a problem. The tiles should be a bit thicker, e.g. *1mm*, so that it is not bouncing when is touched. The expert really liked the front-side central upper opening for the sign.

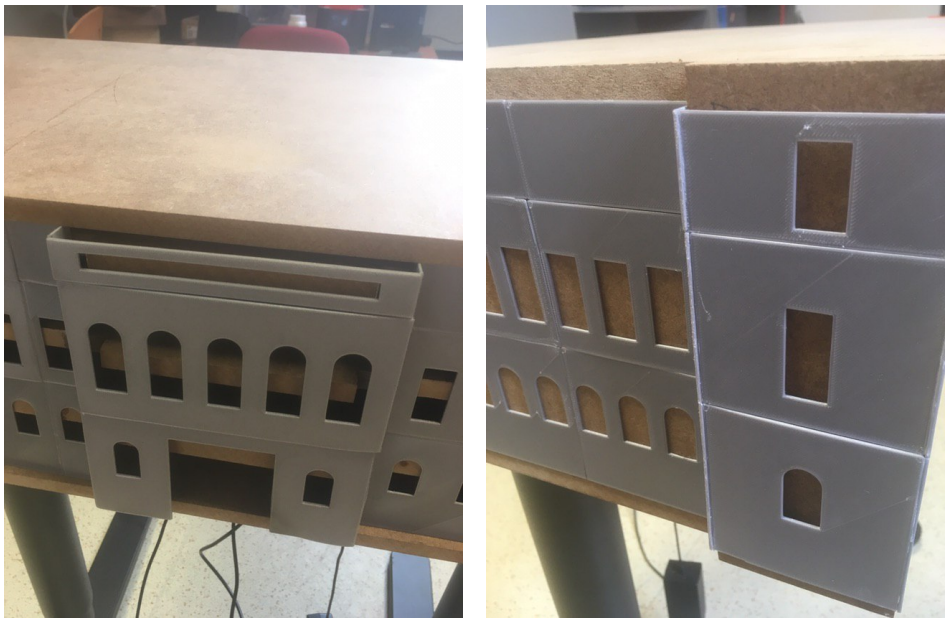


Figure 7.6: First picture: The floors are not on the same level due to drawer desks; Second picture: Sharp corners that exceed the wall

7.3.2 Drawers

When the drawers are opened they do not stay at the same level, which does not allow for the front-side part floors to be on the same level (Figure 7.6). This might confuse the users. When drawers are opened fully, they are not stable. It is a problem, because when users explore the floor they touch the drawer and can slightly move it.

Figure 7.7: First picture: A hole between the tiles. The crossing of four tiles.; Second picture: A "good" corner

7.3.3 Corners

The expert liked some corner tiles, where two perpendicular walls were printed at once, they make smooth transition and the corner feels nice. (Figure 7.7) "This is good, because users will not be afraid to touch the model if all corners were like this, " - said the expert. The other corners that exceeded the wall and were sharp were not advised to use (Figure 7.6). All corners should be printed as a separate tile in the "L" form to provide the smooth transition between to two walls. This will prevent also for 4 tiles crossing, that create a very unpleasant feeling, plus, it is hard to grind off.

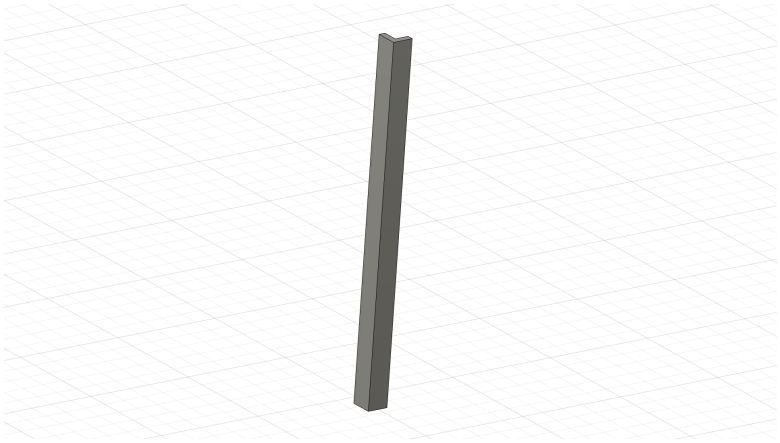


Figure 7.8: An example of how all corners should be printed

7.3.4 Texture

All tiles have different texture. It is not suitable for the visually impaired, because it creates a non-uniform feeling and confuses them. Due to small holes between the tiles (Figure 7.7) the perception of the wall is non-uniform, so it again needs some grind off and coating. Any wall should give an impression of one indivisible piece. "I really like the temperature of the material. For example, I know some models that were made from metal, and my hands always got very cold, " - added the expert. Some tiles were printed with a small bug, e.g. a small line of filament that was not supposed to be there, so that is another reason why it should be smoothed and coated with a color. The color should be as close as possible to a real Palata's facade color.

7.3.5 Summary

After this valuable testing with Lukáš Treml I have highlighted the following:

1. Holes
   Holes between the tiles should be covered with, e.g. Epoxy Putty. The same holes should be grind off with a coarse-grit sandpaper to make them smoother.
2. Tile Texture
   Tiles should be smoothed using a coarse-grit sandpaper.
3. Corners
   Corners should be printed separately in an "L" form (Figure 7.8) and can't be sharp.
4. Tile thickness
   Front-side tiles should be thicker.

5. Distance between tiles in the front
   Front-side floors should have 1mm distance.
6. Tiles glued firmly
   All tiles should be glued firmly to the wooden model.
7. Color
   The color of tiles should be similar to Palata's facade color.
8. Drawers
   Drawers should be stable and when are closed should remain on the same level.

Tiles should be colored and smoothed. In order to achieve that the following should be done [17]:

1. Each tile should be smoothed using a coarse-grit sandpaper.
2. Then it is necessary to apply a layer of liquid metal polish to fill in any gaps.
3. After that, plastic-compatible primer is used to cover up small imperfections in the exterior of the print.
4. Next, the paint is applied. Either acrylic or enamel paint can be used.
5. To protect the painted layers it is necessary to apply a layer of clear coat on it.

Chapter 8

Conclusion

The aim of this thesis was to design and implement an exterior part of Palata residential care institution for visually impaired, while proposing a method for reconstruction of any other building from images to satisfy the requirements of the patent [33] for creation of a tactile 3D map for the visually impaired older adults. The realization of the goals defined for this thesis is described below.

1. **Analysis of the outcomes of current research on the spatial cognition of visually impaired with a particular focus on older adults**

   I have analyzed the statistics on visual impairment in the Section 2.1 Chapter 2. Then, I have studied means of perception if the visually impaired: their orientation and navigation abilities in Subsection 2.2.1, their auditory abilities in Subsection 2.2.2 and their tactile senses in Subsection 2.2.3.

   In the same chapter I have analyzed the tactile maps as means for acquisition of spacial cognition in Section 2.3. In this section I discussed about the maps in general in Subsection 2.3.1, then I studied different researches on tactile maps in Subsection 2.3.2, subsequently I have reviewed new methods for tactile map production: 3D printed maps (Subsection 2.3.3) and interactive tactile maps (Subsection 2.3.4).

   Since my thesis is focused on creation of an exterior part of Palata, I have also analyzed the current research in that area in the Section 2.4: what is Palata (Subsection 2.4.1), how do clients of Palata navigate and orientate inside (Subsection 2.4.2). In the same section I have also investigated on the current research and inventions of my colleague in Subsection 2.6 and 2.6.1. Finally, I have studied why exterior part is important in the 3D tactile model of Palata in Subsection 2.6.2.

2. **Study of a method for 3D model reconstruction from images**

   In the Chapter 3 I have studied a method for reconstruction of 3D models from images. It consists of 5 sections, which talk about image association (Section

3.1), multi-view geometry (Section 3.2), homography (Section 3.3), non-linear estimation algorithms (Section 3.4) and, finally, the 3D model reconstruction pipeline itself and the steps it consists of (Section 3).

3. **Analysis of a technique for segmentation of building facades**

   In the Chapter 4 I have focused on one concrete method of facade segmentation. The method is based on procedural models like shape grammar, that is explained in Section 4.1. Learning of facade elements is made using randomized forests (Section 4.2). This method can be optimized using random walk algorithm (Section 4.4).

4. **Definition of a pipeline for creation of a 3D building model from images**

   I have defined a pipeline for reconstruction of a 3D building from images in the Chapter 5. The pipeline consists of several steps, all of them are described in each section of that chapter.

5. **Implementation of the low and high-fidelity prototypes of the pipeline using the UCD method**

   Following the UCD I have designed and implemented two generations of prototypes of a pipeline for reconstruction of the exterior of 3D buildings based on images (Chapter 6). The low-fidelity prototype was designed as a pdf file and allowed almost no interaction (Section 6.1). The high-fidelity prototype was created in a form of a website with several interactive features (Section 6.2).

6. **Design and implementation of a 3D printable exterior part of Palata**
   In the same Chapter 6 I have analyzed the requirements for the 3D model (Subsection 6.3.1). Subsequently, I designed (Subsection 6.3.2), 3D printed and glued to the wooden model (Subsection 6.3.3) an exterior building part of Palata based the patent [33] requirements (Section 6.3).

7. **Evaluation of the prototypes and the 3D model with the target user group**

   Following the UCD method the low-fidelity prototype was analyzed using the HTA (Subsection 7.1.1) and evaluated using the Cognitive Walkthrough method in three rounds total (Subsection 7.1.2).

   It was tested using Usability Testing technique on a target user group (2 females, 4 males, average age 28.8, professions: IT, architecture, film industry) (Section 7.2). The testing showed that there is a lot to be changed in the user interface and the manual before the creation of an actual product. Thus, I have proposed some changes based on the testing results and have implemented a schematic illustration of a pipeline that can be used when creating an actual product (Section 7.2.12).

   The exterior building part was tested with a visually impaired expert (Section 7.3), that has given me a lot of valuable insights. His advises on improvements are connected to safety, stability, endurance, tactile readability and universalisability of the 3D model. The expert commented on the texture, material and color of the model, the corners of the buildings and the drawers.

8.1 Future Work

For the future development of my project it would be necessary to improve the pipeline following the requirements proposed based on the result of usability testing. One of the participants who studies computer vision and also focuses on 3D model reconstruction from images has highlighted the fact that a web application would be very handy. Thus, the final product could also be in a form of web application. Another thought is that the application could be in a form of checklist, rather than a pipeline, where users can decide what they have and what they still need to do.

There should also be made a few improvements regarding the 3D exterior part of the building that I have created, i.e. the ones that were highlighted by the expert. It is also necessary to design and implement facade elements, like doors, windows and ledges, in order to complete the user experience. Another important part is the roof of the building, which has not yet been tackled.

Bibliography

[1] Google earth. Available at `https://www.google.com/earth/versions/`.

[2] Sidepad. Available at `https://sitepad.com/`.

[3] Balsamiq wireframes, 2008. Available at `https://balsamiq.com/`.

[4] Haptorender, 2009. Available at `http://wiki.openstreetmap.org/wiki/HaptoRender`.

[5] E. Baumgartner, C. B. Wiebel, and K. R. Gegenfurtner. A comparison of haptic material perception in blind and sighted individuals. *Vision research*, 115:238–245, 2015.

[6] I. BELOVOVA DOLEJSI. Speciálně andragogická intervence u zrakově postižených seniorů v rezidenčním prostředí [online], 2009.

[7] A. Bílková, P. Kocurová, and R. Zdařilová. Orientation of blind and visually impaired people-case study. 2019.

[8] A. M. Brock. Touch the map! designing interactive maps for visually impaired people. *ACM SIGACCESS Accessibility and Computing*, (105):9–14, 2013.

[9] C. Büchel, C. Price, R. Frackowiak, and K. Friston. Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain: a journal of neurology*, 121(3):409–419, 1998.

[10] A. Caspo, G. Wersényi, and M. Jeon. A survey on hardware and software solutions for multimodal wearable assistive devices targeting the visually impaired. *Acta Polytechnica Hungarica*, 13(5):39, 2016.

[11] J. Corbin and A. Strauss. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications, 2014.

[12] B. Crandall, G. Klein, G. A. Klein, R. R. Hoffman, et al. *Working minds: A practitioner's guide to cognitive task analysis*. Mit Press, 2006.

[13] J. C. G. T. J. K. M. R. A. V. M. W. . K. B. D A L Maberley, H Hollands. The prevalence of low vision and blindness in canada, 2006.

[14] L. Dandona and R. Dandona. Revision of visual impairment definitions in the international statistical classification of diseases. *BMC medicine*, 4(1):7, 2006.

[15] P. Dominika. Interaktivní hmatové plány pro seniory se zrakovou vadou. 2019.

[16] E. Dubrofsky. Homography estimation. *Diplomová práce. Vancouver: Univerzita Britské Kolumbie*, 5, 2009.

[17] J. Flynt. How to paint pla prints – easy to follow guide, feb 2019.

[18] I. O. for Standardization. *Ergonomics of Human-system Interaction: Part 210: Human-centred Design for Interactive Systems*. ISO, 2010.

[19] E. E. Freeman, B. Munoz, G. Rubin, and S. K. West. Visual field loss increases the risk of falls in older adults: the salisbury eye evaluation. *Investigative ophthalmology & visual science*, 48(10):4445–4450, 2007.

[20] D. Goldreich and I. M. Kanics. Tactile acuity is enhanced in blindness. *Journal of Neuroscience*, 23(8):3439–3445, 2003.

[21] T. Götzelmann. Interactive tactile maps for blind people using smartphones? integrated cameras. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, pages 381–385, 2014.

[22] T. Götzelmann and A. Pavkovic. Towards automatically generated tactile detail maps by 3d printers for blind persons. In *International Conference on Computers for Handicapped Persons*, pages 1–7. Springer, 2014.

[23] R. K. Harrison Chau. Robust panoramic image stitching.

[24] T. K. Ho. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, volume 1, pages 278–282. IEEE, 1995.

[25] L. Holloway, K. Marriott, and M. Butler. Accessible maps for the blind: Comparing 3d printed models with tactile graphics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2018.

[26] K. Hötting and B. Röder. Auditory and auditory-tactile processing in congenitally blind humans. *Hearing research*, 258(1-2):165–174, 2009.

[27] T. Kapic. Indoor navigation for visually impaired. *A project realized in collaboration with NCCR-MICS*, pages 1–10, 2003.

[28] E. Karami, S. Prasad, and M. Shehata. Image matching using sift, surf, brief and orb: performance comparison for distorted images. *arXiv preprint arXiv:1710.02726*, 2017.

[29] Kärkkäinen. Touch mapper.

[30] A. Lobben. Tactile maps and mapping. 2015.

[31] M. Macik, V. Gintner, D. Palivcova, and I. Maly. Tactile symbols for visually impaired older adults. 08 2018.

[32] M. Macik, I. Maly, J. Balata, and Z. Mikovec. How can ict help the visually impaired older adults in residential care institutions: The everyday needs survey. In *2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, pages 000157–000164. IEEE, 2017.

[33] M. Macik, M. Treml, Z. Mikovec, J. Balata, and I. Maly. Trojrozměrný model budovy určený především pro zrakově postižené uživatele, 08 2019.

[34] J. Markel. The sift algorithm for fundamental frequency estimation. *IEEE Transactions on Audio and Electroacoustics*, 20(5):367–377, 1972.

[35] F. W. Mast, N. Preuss, M. Hartmann, and L. Grabherr. Spatial cognition, body representation and affective processes: the role of vestibular information beyond ocular reflexes and control of posture. *Frontiers in integrative neuroscience*, 8:44, 2014.

[36] D. J. McGrenere. Fundamental in designing interactive computational technology for people (hci), 2019.

[37] C. Muchnik, M. Efrati, E. Nemeth, M. Malin, and M. Hildesheimer. Central auditory skills in blind and sighted subjects. *Scandinavian audiology*, 20(1):19–23, 1991.

[38] J. Nielsen. Usability 101: Introduction to usability. *www.nngroup.com*, jan 2012. How to define usability? How, when, and where to improve it? Why should you care? Overview defines key usability concepts and answers basic questions.

[39] M. L. Noordzij, S. Zuidhoek, and A. Postma. The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. *Cognition*, 100(2):321–342, 2006.

[40] D. Pascolini and S. P. Mariotti. Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, 96(5):614–618, 2012.

[41] I. Patel, K. A. Turano, A. T. Broman, K. Bandeen-Roche, B. Munoz, and S. K. West. Measures of visual function and percentage of preferred walking speed in older adults: the salisbury eye evaluation project. *Investigative ophthalmology & visual science*, 47(1):65–71, 2006.

[42] B. Röder, W. Teder-SaÈlejaÈrvi, A. Sterr, F. Rösler, S. A. Hillyard, and H. J. Neville. Improved auditory spatial tuning in blind humans. *Nature*, 400(6740):162–166, 1999.

[43] A. Rosebrock. Image stitching with opencv and python, dec 2018.

[44] L. Rosenblum. *See what I'm Saying: The Extraordinary Powers of Our Five Senses*. W. W. Norton, 2010.

[45] J. Rowell and S. Ungar. The world of touch: results of an international survey of tactile maps and symbols. *The Cartographic Journal*, 40(3):259–263, 2003.

[46] N. Sadato, A. Pascual-Leone, J. Grafman, M.-P. Deiber, V. Ibanez, and M. Hallett. Neural networks for braille reading by the blind. *Brain: a journal of neurology*, 121(7):1213–1229, 1998.

[47] N. Sadato, A. Pascual-Leone, J. Grafman, V. Ibañez, M.-P. Deiber, G. Dold, and M. Hallett. Activation of the primary visual cortex by braille reading in blind subjects. *Nature*, 380(6574):526–528, 1996.

[48] J. L. Schoenberger. Colmap.

[49] J. L. Schönberger. *Robust Methods for Accurate and Efficient 3D Modeling from Unstructured Imagery*. PhD thesis, ETH Zurich, Zurich, 2018.

[50] A. SINGH. A detailed guide to the powerful sift technique for image matching (with python code), oct 2019.

[51] J. C. Stevens, E. Foulke, and M. Q. Patterson. Tactile acuity, aging, and braille reading in long-term blindness. *Journal of experimental psychology: applied*, 2(2):91, 1996.

[52] S. Swaminathan, T. Roumen, R. Kovacs, D. Stangl, S. Mueller, and P. Baudisch. Linespace: A sensemaking platform for the blind. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2175–2185, 2016.

[53] B. K. Swenor, B. Munoz, and S. K. West. Does visual impairment affect mobility over time? the salisbury eye evaluation study. *Investigative ophthalmology & visual science*, 54(12):7683–7690, 2013.

[54] B. Taylor, A. Dey, D. Siewiorek, and A. Smailagic. Customizable 3d printed tactile maps as interactive overlays. In *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 71–79, 2016.

[55] O. Teboul, L. Simon, P. Koutsourakis, and N. Paragios. Segmentation of building facades using procedural shape priors. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 3105–3112. IEEE, 2010.

[56] D. Tyagi. https://medium.com/data-breach/introduction-to-surf-speeded-up-robust-features-c7396d6e7c4e, mar 1019.

[57] S. Ungar, A. Espinosa, M. Blades, E. Ochaíta, and C. Spencer. Blind and visually impaired people using tactile maps. *Cartographic Perspectives*, 28:4–12, 1998.

[58] R. W. Van Boven, R. H. Hamilton, T. Kauffman, J. P. Keenan, and A. Pascual-Leone. Tactile spatial resolution in blind braille readers. *Neurology*, 54(12):2230–2236, 2000.

[59] V. Voženílek, M. Kozáková, Z. Štávová, L. Ludíková, V. Ržičková, and D. Finková. 3d printing technology in tactile maps compiling. In *Proceedings of 24th International Cartographic Conference, International Cartographic Association*, 2009.

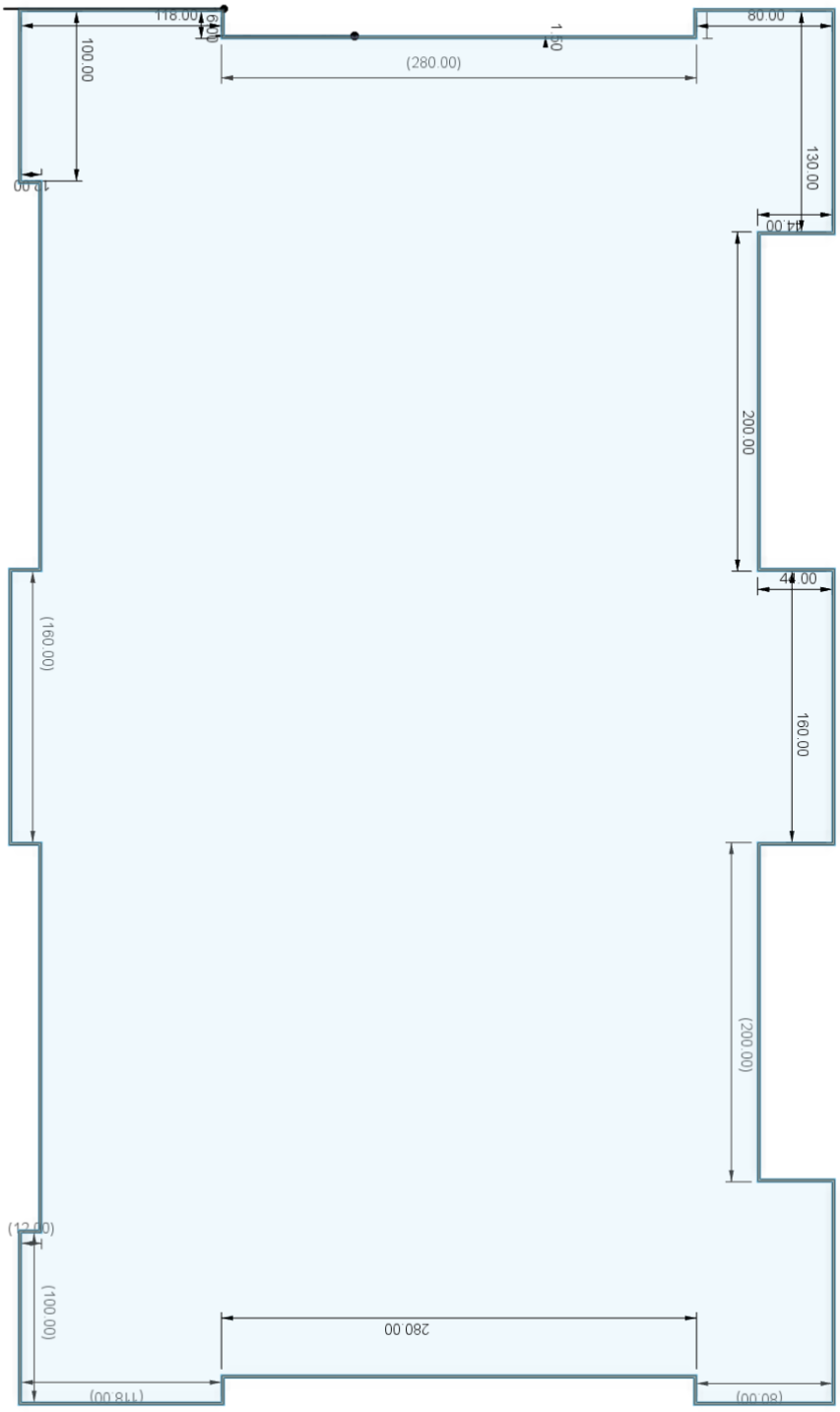
[60] T. WEEVERS. A cry for looking to other methods for user centered design, apr 2012.

[61] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000.

[62] Český Statistický Úřad. Výběrové šetření zdravotně postižených osob - 2013, apr 2014.

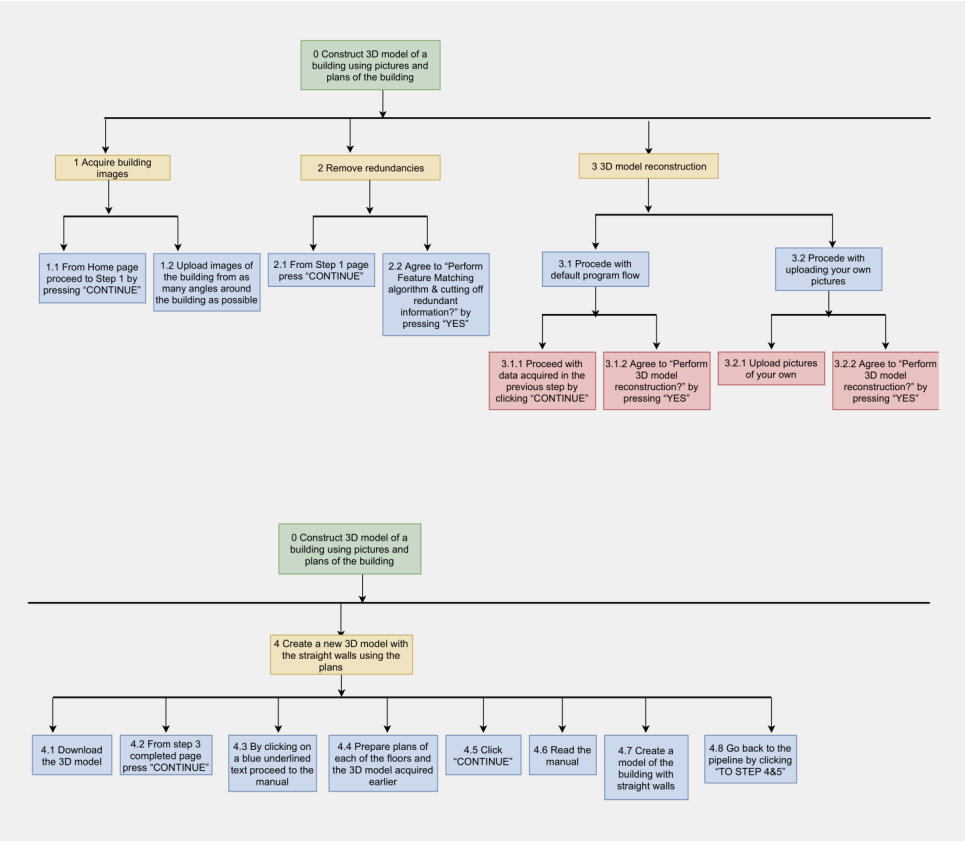
Appendix A

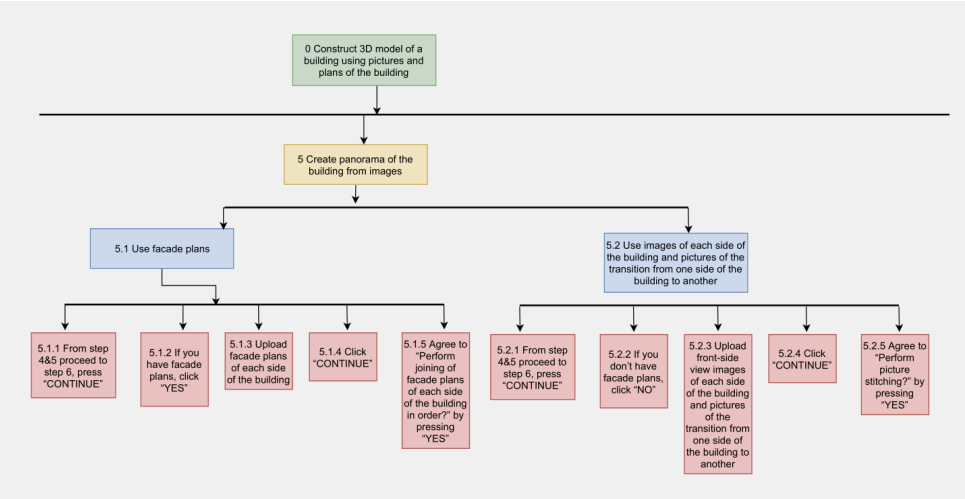
Measurements Of The 3D Model

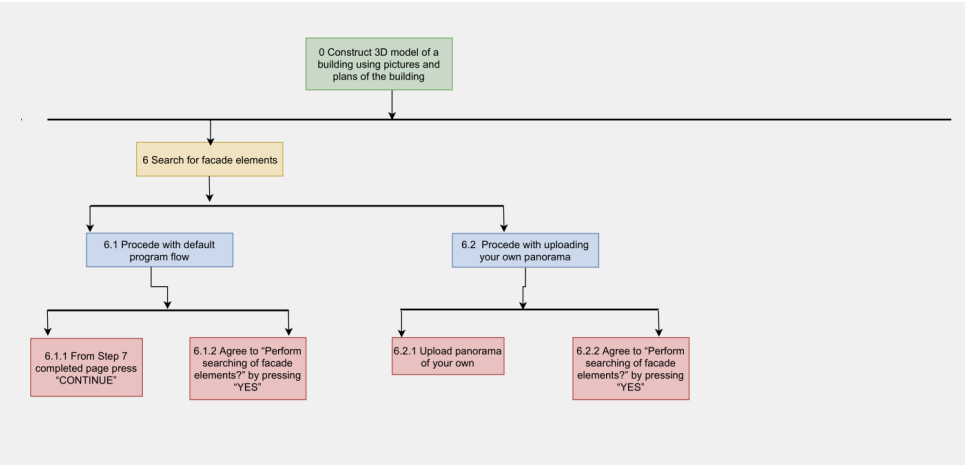

118.00
100.00
(280.00)
80.00
130.00
44.00
200.00
44.00
(160.00)
160.00
(200.00)
(12.00)
(100.00)
280.00
(118.00)
(80.00)

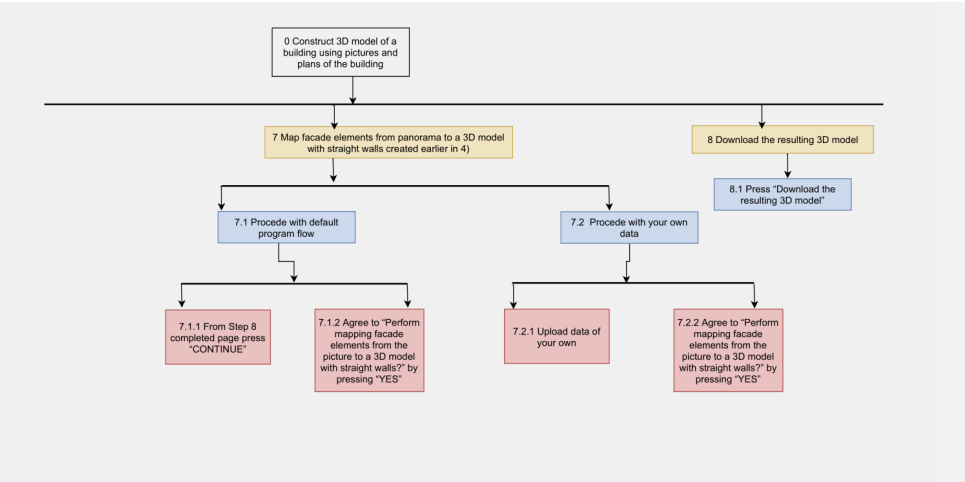
Appendix B

HTA diagram




0 Construct 3D model of a building using pictures and plans of the building
5 Create panorama of the building from images
5.1 Use facade plans
5.2 Use images of each side of the building and pictures of the transition from one side of the building to another
5.1.1 From step 4&5 proceed to step 6, press "CONTINUE"
5.1.2 If you have facade plans, click "YES"
5.1.3 Upload facade plans of each side of the building
5.1.4 Click "CONTINUE"
5.1.5 Agree to "Perform joining of facade plans of each side of the building in order?" by pressing "YES"
5.2.1 From step 4&5 proceed to step 6, press "CONTINUE"
5.2.2 If you don't have facade plans, click "NO"
5.2.3 Upload front-side view images of each side of the building and pictures of the transition from one side of the building to another
5.2.4 Click "CONTINUE"
5.2.5 Agree to "Perform picture stitching?" by pressing "YES"


0 Construct 3D model of a building using pictures and plans of the building
6 Search for facade elements
6.1 Procede with default program flow
6.2 Procede with uploading your own panorama
6.1.1 From Step 7 completed page press "CONTINUE"
6.1.2 Agree to "Perform searching of facade elements?" by pressing "YES"
6.2.1 Upload panorama of your own
6.2.2 Agree to "Perform searching of facade elements?" by pressing "YES"


0 Construct 3D model of a building using pictures and plans of the building
7 Map facade elements from panorama to a 3D model with straight walls created earlier in 4)
8 Download the resulting 3D model
8.1 Press "Download the resulting 3D model"
7.1 Procede with default program flow
7.2 Procede with your own data
7.1.1 From Step 8 completed page press "CONTINUE"
7.1.2 Agree to "Perform mapping facade elements from the picture to a 3D model with straight walls?" by pressing "YES"
7.2.1 Upload data of your own
7.2.2 Agree to "Perform mapping facade elements from the picture to a 3D model with straight walls?" by pressing "YES"

Appendix C

Cognitive Walkthrough

	1. Will users understand how to start the task?	2. Are the controls conspicuous?	3. Will users know the control is the correct one?	4. Was there feedback to indicate the task was complete or incomplete?
1. Upload images of the building from as many angles around as possible (a)	yes	no	yes	no
2. Perform feature matching cutting off redundant information		no	yes	yes
3. Perform 3D model reconstruction		yes	no	yes
4. Upload facade plans (b)		yes	yes	no
5. Perform joining of facade plans of each side of the building in order		no	yes	yes
6. Perform searching of facade elements in the panorama created in the previous step		no	yes	no
7. Perform mapping facade elements from the picture to a 3D model with straight walls		no	yes	no
8. Download the final result		no	yes	yes

Table C.1: Cognitive Walkthrough Round 1 Task 1

	1. Will users understand how to start the task?	2. Are the controls conspicuous?	3. Will users know the control is the correct one?	4. Was there feedback to indicate the task was complete or incomplete?
1. Upload images of the building from as many angles around as possible (a)	yes	no	yes	no
2. Perform feature matching cutting off redundant information		no	yes	yes
3. Perform 3D model reconstruction		no	yes	yes
4. Go back and reupload pictures of the building taken from different angles around it without redundant information (a)		no	yes	no
5. Perform 3D model reconstruction again		no	yes	yes

Table C.2: Cognitive Walkthrough Round 1 Task 2

	1. Will users understand how to start the task?	2. Are the controls conspicuous?	3. Will users know the control is the correct one?	4. Was there feedback to indicate the task was complete or incomplete?
1. Upload images of the building from as many angles around as possible (a)	yes	no	yes	no
2. Perform feature matching cutting off redundant information		no	yes	yes
3. Perform 3D model reconstruction		no	yes	yes
4. Upload facade plans (b)		yes	yes	no
5. Perform joining of facade plans of each side of the building in order		no	yes	yes
6. Perform searching of facade elements in the panorama created in the previous step		no	yes	yes
7. Perform mapping facade elements from the picture to a 3D model with straight walls		no	yes	yes
8. Download the final result		no	yes	yes

Table C.3: Cognitive Walkthrough Round 2 Task 1

	1. Will users understand how to start the task?	2. Are the controls conspicuous?	3. Will users know the control is the correct one?	4. Was there feedback to indicate the task was complete or incomplete?
1. Upload images of the building from as many angles around as possible (a)	yes	no	yes	no
2. Perform feature matching cutting off redundant information		no	yes	yes
3. Perform 3D model reconstruction		no	yes	yes
4. Go back and reupload pictures of the building taken from different angles around it without redundant information (a)		no	yes	no
5. Perform 3D model reconstruction again		no	yes	yes

Table C.4: Cognitive Walkthrough Round 2 Task 2

	1. Will users understand how to start the task?	2. Are the controls conspicuous?	3. Will users know the control is the correct one?	4. Was there feedback to indicate the task was complete or incomplete?
1. Upload images of the building taken from every 5 degrees around it (a)	yes	no	yes	yes
2. Perform feature matching cutting off redundant information		no	yes	yes
3. Perform 3D model reconstruction		no	yes	yes
4. Upload 3D model with straight walls to be used in the next steps (A)		no	yes	yes
5. Upload facade plans (b)		no	yes	yes
6. Perform joining of facade plans of each side of the building in order		no	yes	yes
7. Perform searching of facade elements in the panorama created in the previous step		no	yes	yes
8. Perform mapping facade elements from the picture to a 3D model with straight walls		no	yes	yes
9. Download the final result		no	yes	no

Table C.5: Cognitive Walkthrough Round 3 Task 1

	1. Will users understand how to start the task?	2. Are the controls conspicuous?	3. Will users know the control is the correct one?	4. Was there feedback to indicate the task was complete or incomplete?
1. Upload images of the building from as many angles around as possible (a)	yes	no	yes	yes
2. Perform feature matching cutting off redundant information		no	yes	yes
3. Perform 3D model reconstruction		no	yes	yes
4. Go back and reupload pictures of the building taken from different angles around it without redundant information (a)		no	yes	no
5. Perform 3D model reconstruction again		no	yes	yes

Table C.6: Cognitive Walkthrough Round 3 Task 2

Appendix D

Screener For Usability Testing

D.1 Questions

1. What is your age?
   a. 15-25
   b. 26-36
   c. 37-47
   d. 47-60
2. How would you define your computer skills?
   a. Level 1 - Fundamental Skills (Typing, Mouse)
   b. Level 2 - Basic Computing and Applications
   c. Level 3 - Intermediate Computing and Applications
   d. Level 4 - Advanced Computing and Applications
   e. Level 5 - Proficient Computing, Applications, and Programming
3. How often do you use a computer?
   a. Every day
   b. Very often
   c. Sometimes
   d. Never
4. What is the highest degree or level of school you have completed? If currently enrolled, the highest degree received.
   a. High school graduate, diploma or the equivalent

b. Some college credit, no degree
c. Trade/technical/vocational training
d. Associate degree
e. Bachelor's degree
f. Master's degree
g. Doctorate degree

5. What is your English language knowledge level?
   a. A1 (Beginner)
   b. A2 (Elementary English)
   c. B1 (Intermediate English)
   d. B2 (Upper-Intermediate English)
   e. C1 (Advanced English)
   f. C2 (Proficiency English)
6. What is your profession/field of study?
   a. Programmer
   b. 3D model designer
   c. Architect
   d. Civil Engineer
   e. Other

   Please specify the field _

7. What operation system do you use on your computer?
   a. Windows
   b. Linux
   c. macOS
   d. I don't know

e. Other, specify _

8. How often do you create 3D models in a 3D modelling environment, like Autodesk Fusion 360, AutoCAD, Blender, Maya and so on?
   a. Every day
   b. Very often
   c. Sometimes
   d. Never
   e. Long time ago
9. Have you ever encountered/heard of an application for 3D model reconstruction from images?
   a. Yes
   b. No

D.2 Answers

D.2.1 Pilot Test Participant

1. 2
2. 5
3. 1
4. 5
5. 5
6. 1 front-end development
7. 1
8. 5
9. 1

D.2.2 Participant 1

1. 2
2. 3
3. 1
4. 6
5. 3
6. 3
7. 1
8. 3
9. 1

D.2.3 Participant 2

1. 1
2. 5
3. 1

4. 5
5. 5
6. 1 front-end development
7. 3
8. 4
9. 2

D.2.4 Participant 3

1. 2
2. 5
3. 1
4. 5
5. 4
6. 1 computer vision
7. 1
8. 4
9. 1

D.2.5 Participant 4

1. 2
2. 5
3. 1
4. 5
5. 5
6. 1 data science & machine learning
7. 3
8. 5
9. 2

D.2.6 Participant 5

1. 3
2. 5
3. 1
4. 3
5. 5
6. 5 movie producer
7. 1
8. 1
9. 2

Appendix E

Usability Testing Sessions

E.1 Test 1

Participant's profile

The participant is a female, 26 years old. She's born in Russia and lived in Prague for 8 years. The participant has a B1 level of English. She works with computers every day and she uses the Windows operating system. She has a Master Degree in Architecture and works as an architect, so she had experience with modeling of 3D buildings. She has heard of an application that produces 3D models out of images.

Testing environment

The testing session took place at a study cafe in Prague on Sunday, 18.07.2020 at 12:30. There were almost no people in the cafe, so it was very quiet. The participant allowed me to record her voice and the screen. The whole session lasted for 1 hour and 26 minutes.

Pre-test questionnaire

The participant imagines the application of 3D model reconstruction from images as an application where a user submits a map of a city and the program returns a 3D map, where each building is extruded to a cube or a rectangle. She says that she might need to use it in her work. "Though, this kind of a program for a 3D building reconstruction from images would be too primitive. In my job we need

detailed buildings," - she said.

Session

Use case 1

Step 1: The participant thought that after she uploaded the pictures, the gallery that was shown as an example was containing those pictures that she downloaded. After the upload, she did not see the indication that the pictures were uploaded immediately.

Step 2 completed: She immediately recognized that the resulting pictures are the buildings with no redundant information by commenting it out loud.

Step 3 completed: The participant commented that scanning of buildings works the same.

Step 4&5: The participant saved the 3D model right away. She mixed the download button with the upload button and missed the link to the manual. After I helped her find it, she said it's better if there is a button for it.

Manual-1: The participant commented that it's better if the download of the 3D model is located directly in the manual.

Manual-2: The participant read the "pro tips" and said she is used to working with AutoCAD, and this program is similar. The origin was hidden in the beginning, it was confusing for the participant in terms of finding the *XZ* plane. After some time she has found a way to unhide the origins and has successfully uploaded the first floor plan. She was unable to edit the canvas, instead she opened the "Delete entities" popup window and could not have exited it. After a while she was able to find the right control and move the picture to the origin. She was able to upload the mesh into the scene, but was confused on how to move, scale and rotate the figure. She said she would expect the manual to be more concrete with more detailed pictures. The most problematic deformation of the 3D model was scale, since it had its own control in the modelling environment. After she found them, she had a problem with scaling itself. She did not see that there's a box where she can put the scaling number or that there's a control that she can pull. She was complaining that even with her knowledge of modelling she still has problems with finding the controls, so she advises to make the tutorial more detailed, where each step is explained up to the smallest substeps. After all controls were found she did not have any problem with placing the building into the right position and place. When performing the step with lowering of the opacity, she did not see the information in the manual where she was supposed to click "OK", so she could not have moved to the default environment, even though this information was present in the manual. The participant said the problem is that several steps of the manual are written in one, so when she read

the concrete step, she had already forgotten the whole thing by the time when she finished some part of that step. She was again failing to locate the XZ plane, since it was invisible when I provided Autodesk to her, so I had to help her locating the origins menu in the left corner. In the next step, where the building base is created she forgot to hide the mesh body. The participant complained that she would prefer more pictures and less text in the manual. When she had to extrude the plane there was a major problem: she missed a manual point for creation of the offset on the sketch, so extrude did not work. After I pointed it out and she tried extruding again she faced another problem: while clicking on everything in the left-side menu the participant accidentally hidden the sketch of the base, so it was impossible to extrude it. In the end she successfully finished the manual. When she was done, she clicked the "Continue to step 4&5" button and was returned to the step 4&5, it confused her - she thought it was the next step.

Step 4&5 result: When uploading the 3D model with straight walls she did not at first see that the file was uploaded.

Step 6 front-side: The participant was not sure if she needed to upload all files that she was given.

Step 7 completed: The participant stated out loud that she doesn't understand why there are download and upload buttons and what she needs to with them. She complained that these buttons are the same color and she mixed them. The participant said it's not understandable why facade elements need to be found and more explanation could be helpful. The participant says she doesn't understand where she is now in the pipeline and what and when is the final outcome, so she would be happy if there was a graphic scheme of the pipeline in each page. She says she doesn't understand why users would need to upload and download anything in between the steps. The participant says it is confusing that in the header the steps are joined together, e.g. "4&5" or "6&7".

Step 8 completed: Again the participant was confused why there are the upload and download buttons.

Result: The participant successfully downloaded the resulting 3D model, but she said it would be better if the indication of the final result was more visible.

Use case 2

There was no problem in completing this task since the participant got familiar with the pipeline and the user interface in the previous task.

Use case 3

The participant successfully completed the task. The pictures in the gallery seemed confusing to the participant. She said it would be better if there's a grid, so it's more visible that those are separate pictures.

Post-test questionnaire

Since she works in architecture, she had her own professional view on the subject of the topic of this high-fidelity prototype. She pointed out that this kind of program could be very successful in architectural prototypes if certain thing were fixed. The participant 1 liked that this application is made in a pipeline way. "If certain things are changed in a good way then it would be hard to get lost in this application, " - she said. The only thing that she knew which she could compare this high-fidelity prototype to is an application she used at her university for creating a 3D map out of a 2D. The problem with that application is that it is not detailed at all. This gives an advantage to the application that we were testing.

The participant said that the user interface of the testing application could be improved. For example, upload and download buttons could be different color, or removed at all. If they will not be removed, then the indication of an upload and download action should be more visible. The gallery could have a grid, so it is more clear that pictures are separate. A link to the manual could be more visible or it could be a button instead. Another alternative that she came up with is having a manual as part of the pipeline. "It is confusing that after the manual I return back to the same step where I have started, " - said the participant. Another thing that seemed wrong to the participant 1 is that steps are joined together in the header menu. The participant said, that placing the download of the reconstructed 3D building into the manual itself could give more understanding to a user. In the last page the indication that it is the final result should be more clear, as well as the button to download it. The manual was another thing that was confusing to the participant. She said, it would be great if it was more detailed and if each step in the manual consisted of just one action. "More pictures, less text, " - she said. She would recommend this program to her colleagues if it was fixed in a more user friendly way.

Summary

The participant indicated several problems along the pipeline. She was mostly confused in each step's completion page, where she had an option to download the result and upload her own data: she thought it was an obligatory step. The participant 1 was lost in the pipeline, she wasn't sure where she is and when does it end. Another problem is the manual - it was hard to navigate in an unknown program to create a 3D model. The idea that manual is a side page, rather than a part of a pipeline, was unclear to the participant.

After this testing session a few things have been detected for improvement:

1. Each step should have a graphical scheme of the pipeline, where the current step is highlighted.
2. Upload and download buttons should be different colors.
3. Indication of the upload and download action should be more visible.
4. The picture gallery should have a grid.
5. Steps should be renamed in such a way that they do not confuse the user, e.g. "4&5" can be renamed to just "4".
6. The default program flow should be more clear to the user. "Continue" button could be renamed to "Continue with the computed result". Before the "Download" and "Upload" buttons it should be clearly stated that it is not obligatory to use them.
7. Manual should be part of the pipeline, rather than a side step.

Manual could be improved in the following way:

1. More images, less text.
2. One step - one action.
3. Images should have highlighted controls.
4. Each modification confirmation of canvas, mesh or sketch should have its own step in the manual.
5. Showing/hiding canvases, meshes and sketches should be highlighted in the manual.
6. Each rotation of the camera should be shown in the manual.
7. Each scaling, rotation and other transformation should be shown in the manual in details.
8. The first step of creating the model should be unhiding the origins, since by the default they are hidden.

E.2 Test 2

Participant's profile

Participant is a male and is 24 years old. He has C1 English and a Bachelor Degree in programming. He works as a front-end developer and uses the macOS operating system on his computer. He has never had an experience with 3D modelling and has never heard of an application that transforms images into a 3D model.

Testing environment The testing session took place in a National Technical Library in Prague on Sunday, 19.07.2020 at 18:00. The library was almost empty and we had a quiet place where we could speak without any problem. The participant allowed me to record his voice and the screen. The whole session lasted for 1 hour and 32 minutes.

Pre-test questionnaire

The participant said that he did not know anything about the 3D model reconstruction from images and he has never encountered such an application. Nevertheless, he said that he heard of 3D scanners. "You put an object on the table, make sure that colours of the table and the object differ are bright. You take a lot of pictures around, then put into a program and run it. You get a 3D model, " - explained the participant. He said he expects it to be somewhat like the scanner and that he might use it at work if it's ever needed. He never had an experience with 3D modeling.

Session

Use case 1

Intro: The participant didn't read anything in the intro page. He wasn't sure if he should click something in the header menu or in the page to get to step 1.

Step 1: Using drag and drop he has selected files in the files from the folder that he was provided. He wasn't sure if files were uploaded, it took him some time to see the sign below. He said he was expecting some loading bar and more indication.

Step 2 completed: The participant read the sign "Building images without redundant information are ready" for a few times. He said the sign is too long to understand it from the first time. The participant is confused whether he should download and upload something. He tried downloading the data and uploading them for the next step.

Step 3 completed: The participant is happy with how the 3D model is presented - he rotated it a few times.

Step 45: Participant wasn't sure what he had to do in this step. He did not download the 3D model, but he has found a clickable link and clicked it.

Manual-1: In the manual he saw that he needs the 3D model, so he went back and downloaded it. He said the downloading of the 3D model should take place directly in the manual.

Manual-2: The participant read all “Pro tips”. The origins were hidden in the beginning, so I had to interfere and unhide them. When inserting the canvas, he wasn't sure which origin is *XZ*, so it took him some time to guess it. After he successfully placed the first floor picture into the scene he didn't see the rotation controls, instead he tried using the “Edit Canvas” popup menu. He did not see the controls for moving the canvas, so he used the popup windows instead, which took him some more time than it should have. After rotating the image he forgot to click “OK”, so he had problems continuing with the manual. When trying to move the picture he again had troubles finding the controls. After successfully inserting the 3D model he was not sure how to scale, rotate and move the object, as it was not specified in the manual, so I showed him the controls instead. The participant complained that this kind of manipulation in an unknown program especially for someone who does it for the first time in his life ought to be explained in detail. After the transformation was performed, he again forgot to click “OK” to exit the editing mode, so he got stuck for a while. It was hard to the participant to locate the second and the third floor since he did not rotate the camera in the right angle. Before creating the building base sketch, the participant did not hide everything that he was suggested in the manual, so it was a bit problematic for him to perform the action. After I pointed it out, he said, that the step in the manual was written instead of shown in the picture, so he forgot to hide some elements. The participant created the building base sketch, but the lines were not connected. I pointed it out to him. After he finished the sketch, he did not click “Finish Sketch”, so he could not find the extrusion control, so it took him some time to realize that. After he did, he complained about the manual steps being too full of information and pointed out that they should be as short as possible with more pictures. He successfully extruded the walls, but the level at which he extruded did not match the building levels, because the planes the indicate the floor were hidden. When the participant was done, he returned to step 45, but he was confused, because he thought that it's the next step. He said the he would expect manual to be a part of the pipeline.

Step 45 result: The participant did not notice that indication of successful upload after he uploaded his 3D model.

Step 6: He wasn't sure if he had plans or front-side view pictures, so he checked in the folder that he was provided with.

Step 6 Front-Side: The participant 2 did not notice right away after uploading pictures, that they were uploaded.

Step 7 completed: It is again unclear for the participant whether he should download or upload something on this page.

Step 8 completed: The participant has the same confusion like in the previous step.

Use case 2

This task was completed successfully since the user was familiar with the user interface from the previous task.

Use case 3

The participant successfully located the gallery and commented that there should be a button to enter the gallery, instead of clicking on the pictures.

Post-test questionnaire The participant liked the idea of the application. He said that his team at work might need something similar in their future projects. Though, he did not like several things.

Firstly, the indication of the file being uploaded and downloaded were not visible to him and he would expect a progress bar. Secondly, the upload and download buttons in some steps are confusing, he doesn't understand why are they necessary. He proposes to remove them at all. Then, the names of the steps should be more clear to a common user, they should not be long and complicated to read. Another proposition was to put the downloadable 3D model in the tutorial itself, rather than keeping it in the "4&5 step". Also, he would make manual steps explanation more brief, since he think that a person cannot concentrate on so much information at once. The participants is also proposing to put more pictures in the manual. Manual should be a part of the pipeline, rather than a fork. The header should not include steps at all according to him. "It's a pipeline, so why would a user want to go to a certain step immediately?" - said he.

The participant does not know anything similar to compare this application to. If he could change something, he would add a depiction of the whole pipeline to each page while highlighting the part that the user is at now. Also, he would change everything that he explained before. The participant would recommend this program to his colleagues when it is fixed.

Summary

The participant was successful at completing the task. Though, he got confused several times. Here are a few things that can be changed about the interface:

1. Make user understand, that "Upload" and "Download" buttons are not obligatory.
2. Rename the steps so that they are more clear to a user and shorter. Add some additional explanatory text under each step.
3. Add 3D model to the tutorial, instead having it in step 45.
4. Manual should be a part of the pipeline.

5. Each page should include an illustration of the whole pipeline highlighting the current step.
6. Remove the steps from the header.

The manual can be edited in the following way:

1. More pictures. Less text.
2. Tell the user to unhide the origins in the beginning.
3. Clicking "OK" should be it's own step and should have a picture with an illustration.
4. Add pictures where controls for moving/editing canvas/object/sketch are highlighted.
5. Before creation of the border sketch the elements that are supposed to be hidden should be highlighted in the picture individually.
6. Add more specifications and pictures when elements should be hidden.

E.3 Test 3

Participant's profile

The participant is male and is 31 years old. He's Russian and loves in Prague for 2 years. The participant's a student studying master's in computer vision. His job is also connected to computer vision, so using a computer is his everyday matter. He speaks English with the level B2. The participant uses Windows operating system. He has never had an experience with modelling of 3D objects, but he used an application for reconstruction of 3D objects from images.

Testing environment

The testing session took place in Karlovo Náměstí faculty building of the Czech Technical University on August 3rd, 2020 at 14:15. It ended at 15:47. The environment was quiet and without any people around. The participant allowed me to record his voice and the screen of the computer that we were testing on.

Pre-test questionnaire

The participant had an experience with using Colmap. When he first used Colmap, he was not sure what the pipeline consisted of, but this application had a manual,

which he read and then it was clear. The pipeline consists of the following steps: feature reconstruction, matching and 3D reconstruction with a “dense” option. He liked everything about that application. “It’s an easy, good interface, a lot of buttons you can press to see some more details. It can import into txt or ply format. A lot of settings, including default. You can use this program from the console, so if you need to reconstruct a lot of models, you don’t need to click a lot, ” - said the participant. He would expect an application similar to that to have a web interface and a bit more information concerning the formats, or statistics on some file reconstruction. The participant’s job focus is now connected to 3D model reconstruction from images, so he will definitely use such an application. Though, he never modeled anything in any 3D environment.

Session

Use case 1

Step 1: When he uploaded files, he did not see the indication of the file being uploaded. He doesn’t understand what will happen in the next step and why, because the words are too difficult.

Step 2 completed: He doesn’t understand why there are two buttons. “If I use these buttons, do I have to press “CONTINUE”? Then why is it above those buttons?” - said the participant. But he did understand what these buttons are made for.

Step 3 completed: He doesn’t understand why there’s a need to create a building with straight walls.

Step 4&5: He did not see the clickable link for the manual, but he did download the 3D model.

Manual-1: Did not read anything on this page.

Manual-2: The participant noted that when the tutorial mentions a menu, he looks on the upper menu, not the menu below it. “ Also the left-side menu should be called a browser, ” - he said. He was not sure where the infinity of the Z axis is located. Instead of using the controls, he tried using the edit menu, because he didn’t see it. In the end he found the controls and noted that it should be depicted in the menu. He did not click “OK”, since he did not read the whole step of the manual. When he entered the mesh edit mode, he did not notice that he was in the edit mode, so he tried editing the mesh again and he noted that this information should be depicted in the manual. Then, the participant could not find the “modify” menu button. When he tried moving the mesh, he did not realize that he needed to click on the mesh first. When drawing the sketch he was not sure how to use the lines. When he finished modelling the building and pressed “CONTINUE to 4&5”, he did realize that he went back to step 45, unlike the other participants.

Step 7 completed: the participant noted that even though he understands why there are upload and download buttons, the context is still missing. Also, he said that those buttons don't seem to be connected to the continue button.

After that, the participant did not encounter any problems.

Use case 2

The participant did not have any problem with successful task completion. But again, he noted the fact that upload and download buttons should be connected to the "CONTINUE" button in another way.

Use case 3

The participant did not know what is "removed redundancies". He got confused when completing the task, he though he needed to upload the pictures without the background into the first step and noted that the task should have been formulated in another way. He thought that he already had pictures with no background. In the end, he opened the pictures of the building from step 1. He augmented it by saying that in the step 2 completed page he did not notice that there is some connection to the pictures he uploaded in step 1, plus, he did not see that they were clickable. He would rather have a button next to them.

Post-test questionnaire

The participant would like to do as less manual work as possible, e.g. without modelling in Fusion 360. He also liked the buttons design: non-obligatory buttons are grey and obligatory are green. Each step in the manual should be depicted as following: text in the left and a picture in the right, so when scrolling down, the user can see what they did in the previous steps and what they will do in the next steps. Upload and download options should be marked as non-obligatory and should be located before the "CONTINUE" button. He would expect each step of the pipeline to have a depiction of the whole process and where the user is located now in that process. Plus, in each step where users can upload their own data, it should also be visible in that scheme. Another thing is that "straight walls" he would rename into something, like "uniform" or "simplified", and he would also add explanation why the 3D model is being simplified.

The participant really liked that the application is online, so he wished that Colmap had the same option. Compared to Colmap, the application that we were testing had some manual steps, which he did not like. The participant advised to use some kind of triangle reduction algorithm to make the manual step automatic. The participant in his work does not need to have a perfectly reconstructed exterior of the building with minimized number of triangles, so he said he will probably not use the program that I have presented to him.

Summary

Overall, the participant did a great job in completing all three tasks compared to other participants. I assume it has to do with the fact of him knowing the process of 3D model reconstruction from images. Though, there are a few things that came up to be unusable.

Regarding the user interface the following changes can be done:

1. Upload and download options should be marked as non-obligatory and should be located before the "CONTINUE" button.
2. Add illustration of the pipeline to each step and mark the current step.
3. Add information to the illustration about the user being able to upload his own data.
4. Rename manual's and step 4&5's name from "straight walls" to "simplified".
5. Make steps' names less difficult & add explanation, e.g. "removed redundancies" is confusing.
6. Explain why there's a need to create a new simplified 3D model.
7. Manual should be a part of the pipeline.
8. Gallery should be opened using a special button, rather than a picture be clickable.

The manual can be changed in the following way:

1. Left-side menu should be called a browser.
2. Highlight the controls, so users know they can use them.
3. Clicking "OK" should be its own step and depicted.
4. Each step should be explained more detailed with pictures.
5. Steps in the manual can be changed like this: text in the left and a picture in the right.

E.4 Test 4

Participant's profile

The participant is a female and is 26 years old. She's Russian and lives in Prague for 9 years. She is studying for a master's degree in machine learning and data science. The participant's a programmer and has a C1 level of English. The participant uses the macOS operating system. She had an experience with modelling of 3D objects a few years ago and never heard of an application for 3D model reconstruction from images. At the moment she is pregnant, 8th month. She asked me to note it, because she says she feels very unfocused in the last months.

Testing environment

The testing session took place on 7.08.2020 at 19:00 in the participant's house, it was quiet there. It lasted for 2 hours and 32 minutes, she had to take a few breaks. The participant allowed me to record her voice and the screen of my computer, where we were testing.

Pre-test questionnaire

She has never heard of an application for 3D model reconstruction, but she would expect an intuitive design and some kind of a pipeline method. The participant says she might hardly ever need to use this kind of application, since it's not connected with the field of her study. She had an experience with modelling 3D objects in Blender a few years ago, but she did not like it: the design of those applications was complicated and required a lot of knowledge.

Session

Use case 1

Step 1: The participant didn't understand what "5 degrees around the building" meant. She did not see that the pictures were uploaded for the first time.

Step 2 completed: The reason for upload and download buttons are not clear to her. But she did understand that the background was cut off from the initial pictures.

Step 3 completed: She was not sure if the provided 3D model is the result of her previous actions.

Step 4&5: She followed the instructions and downloaded the model right away, but she didn't understand what building plans were and she thought she didn't have them.

Manual-1: She read all the text.

Manual-2: After successfully uploading the first floor plan she was unable to rotate it and "infinity in Z axis" seemed very confusing to her. She supposed it's the one that points up. Then after she finally found the rotating option she did not see the controls, instead she was trying to rotate the camera. She spent a lot of time rotating the camera instead of the canvas itself. The participant inserted the mesh and tried to edit it, but then another problem occurred: she forgot that she edited it and that's why she got stuck on trying to click "Edit" on the mesh. After she realized it, she proceeded to scaling, but again using a wrong approach. Instead of looking at a tutorial, she tried to do it herself and started clicking on the mesh to find a scale option. I pointed it out to her. When performing rotation she encountered another problem: she did not notice the rotation controls. When trying to construct a new plane of the second and the third floor the participant was not sure which one is the XZ plane, she mixed it up with the XY plane. Another problem that appeared is that she did not know how to use the line for the sketch. She thought that if she chose the line in the menu it would appear just as one line on the sketch and she would have to repeat the choosing. In the end, the sketch was created. Now that she had to create an offset, e.g. create a wall, she made it too thick and in the inside part of the building. After finishing the manual she clicked "CONTINUE TO STEP 4&5", and was returned back, which confused her. She downloaded the model again and thought it was the object that she just created.

Step 4&5 result: The participant realized that she was wrong assuming that the download button in step 4&5 was for downloading the model she just created in Fusion 360. After she uploaded the model she just created in the high-fidelity prototype, she again was confused about the upload status, as it was not clear to her.

Step 7 completed: She was not sure what she needed to do with the download and upload buttons again.

Step 8 completed: The same issue as step 7 completed.

Use case 2

The participant did not have any problems with this task.

Use case 3

Instead of opening the "pictures of the building with no redundant information" in step 2 completed page, she opened a gallery in step 1. She said she didn't understand the task.

Post-test questionnaire

The participant liked the idea of the application she was presented with. Also, she liked the design and layout. She can't compare it to anything else, because she doesn't know anything similar.

She would change the program's feedback, e.g. "Images have been uploaded" should be more explicit. Upload and download buttons should be marked as optional according to the participant. She pointed out a few things about the manual: the "infinity in the Z plane" should be rewritten, controls should be highlighted in the manual pictures, it should be explained how to use lines in the sketch, manual itself should be part of the pipeline.

The participant said that she would recommend this program to her friends and colleagues.

Summary

The hardest part of the pipeline for the participant became the manual modelling. She spends a lot of time on it compared to other participants. A few things can be learnt from this testing session regarding the UX:

1. Upload and download indication should be more explicit.
2. Upload and download buttons should be marked as optional.
3. Manual should be part of the pipeline.
4. Pipeline should be illustrated in each step, so the user knows what's going on and where they are in the pipeline.

The manual can be changed in the following way:

1. When talking about the Z or XZ plane they should be highlighted in the manual
2. Controls should be highlighted as well.
3. Each edit mode should be depicted and highlighted in the manual.
4. It should be explained with pictures how to use lines when creating a sketch.

E.5 Test 5

Participant's profile

Participant is a male, 39 years old and is Italian. He speaks English with the level C1 and is a proficient computer user - he uses a computer every day. The participant is a movie producer, but used to be a front-end programmer and a graphical designer. He has professional diplomas in advertising and graphic design and visual and web design. The participant had an experience with creating 3D models a long time ago. He has never heard of an application for 3D model reconstruction from images.

Testing environment

The testing took place at the participant's home on Friday 8.08.2020 at 18:00. He lives alone, so it was a quiet environment. Testing took 1 hour and 28 minutes to complete. The participant agreed to his speech being recorded with a voice recorder and for the screen of the computer to be recorded while testing.

Pre-test questionnaire

The participant has never heard of an application that transforms images to a 3D model, so the only thing he expected was for application to be as automatic as possible. He would probably like to use this kind of an application in future if his film producing job ever demands it, e.g. for recreating real-life objects in 3D.

Session

Use case #1

Step 1: The participant uploaded the files, but he's not sure if they were downloaded. After some time he noticed a sign. Before jumping to the next step he read the name of the next step and he said he isn't sure what it means and why it's important.

Step 2 completed: The participant did realize that the background was removed, but did not get why. In the step he was confused by the upload and download buttons, but then he guessed that it was optional. Nevertheless, he tried to download and upload the data. "It should be specified what these buttons are for and if they are obligatory they should be above the "CONTINUE" button, " - said the participant.

Step 3 completed: The participant understood what happened and why he needed pictures with no redundant information.

Step 45: He downloaded the 3D model, but did not see the manual link. I had to point it out to him.

Manual-2: The participant suggested that the 3D model download should be located in the manual and that there should be a link to Fusion 360 (he did not notice it). The participant resized the window with the manual into a smaller one to the side, this was not letting him see the manual pictures, so he was just reading the text. Throughout the whole manual part the participant was complaining about the software, specifically that it snaps when something is moving, so the user cannot define their own distances. He advised me to write a tip on how to remove it. When trying to edit the canvas the manual said to "click on the canvas in the menu" which confused the participant, because he was left-clicking, and it should have been right-clicking. The participant did not know how to rotate - the manual did not specify how. Then, he had a problem with moving the canvas - the manual did not specify the controls. Then, he found the number input menu and was struggling with choosing the right distance. Moving and scaling of the building was not a problem. The participant did not understand how he's supposed to measure the next floor by eye. When creating a sketch he had a problem with using the line, it was jumping all around the screen. After that all went well. After he successfully created a building, he was redirected back to step 45, which confused him. "I'd imagine to proceed with the pipeline rather than going backwards, " - commented the participant.

Step 6 Front-side: He uploaded the files, but again did not notice the indication. Also, the participant commented on the sign "Next step is..." advising to remove the dots.

Step 7 completed: The participant again commented on the upload and download buttons that he doesn't understand if those are obligatory.

Result: The participant said that he likes the result.

Use case #2

The participant had no problem with this task since he already did it in the previous task.

Use case #3

The participant advised to have a button next to the pictures and to remove the text "click to open gallery" because it doesn't look professional.

Post-test questionnaire

The participant liked the application, though, he would change a lot. He liked that it was a pipeline and that the buttons that lead the user through the process are green, when the ones that are not important are grey. He also liked the result. He did not like that upload and download buttons are not explained well, that the

names of the steps are complicated and there's no explanation why they are there. Also, he pointed out that the manual should not return back, but rather lead forward. "The manual itself should give users an option to remove the snapping. It was very annoying , " - said the participant. He also said that he would play with the user interface a bit to make it look more "professional". The participant said that he might need to use this kind of software in the future, e.g. in movies. But he would use/advise to colleagues the software that we just tested only if it was fixed.

Summary

The participant #5 was the quickest among all to complete all the tasks. He was also quicker in completing the manual. I assume it is due to his knowledge of image editing software. Among with the common issues that I have found from this session the participant offered me a few new interesting improvements.

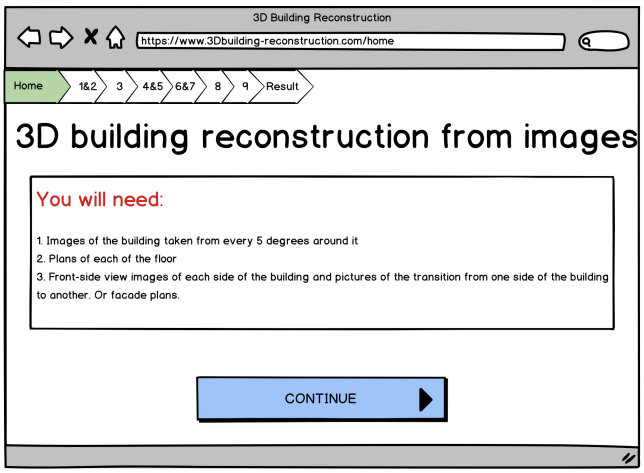
1. Upload and download buttons should be explained more.
2. Rename steps in a simpler manner.
3. Add description to the steps.
4. Manual should be a part of the pipeline.
5. From "Next step is..." dots should be removed.
6. Gallery should have a button instead of clicking on pictures.

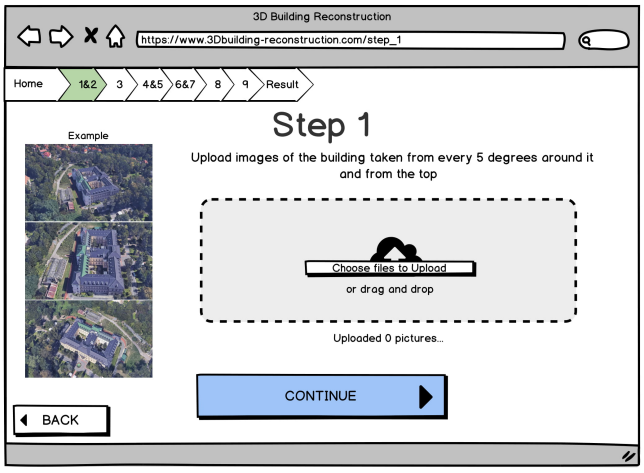
The manual can be changed like this:

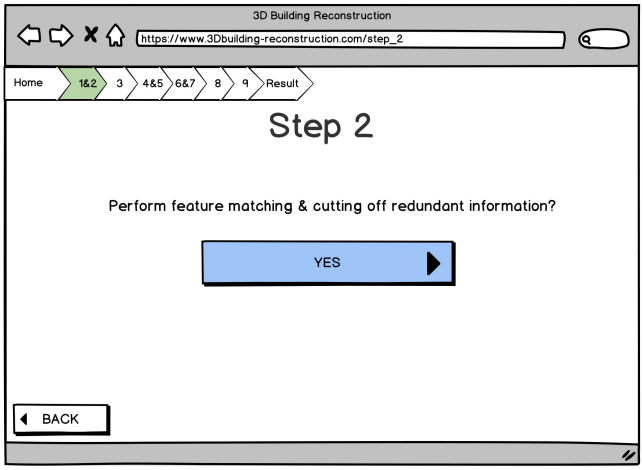
1. Add a description on how to remove snapping.
2. When user's supposed to be right-clicking it should be specified.
3. Controls should be described in the manual.

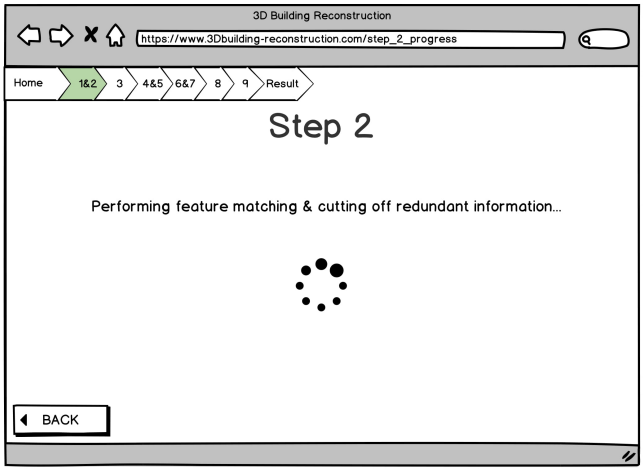
Appendix F

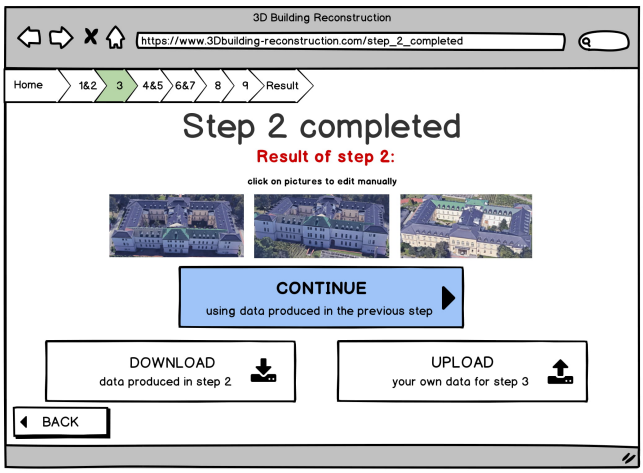
Low-Fidelity Prototype



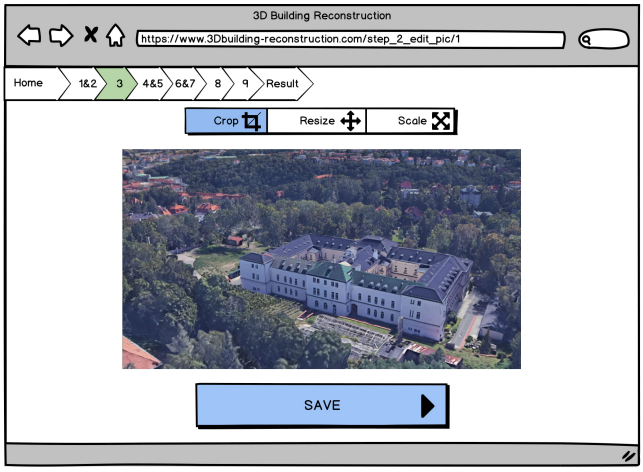



3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_2
Home
1&2
3
4&5
6&7
8
9
Result
Step 2
Perform feature matching & cutting off redundant information?
YES
BACK

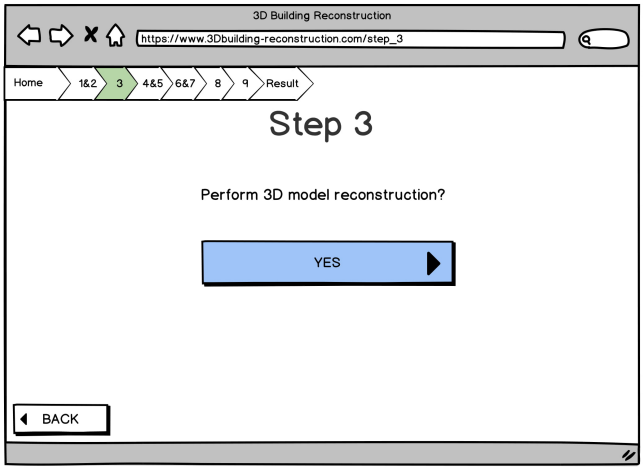

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_2_progress
Home
1&2
3
4&5
6&7
8
9
Result
Step 2
Performing feature matching & cutting off redundant information...
BACK

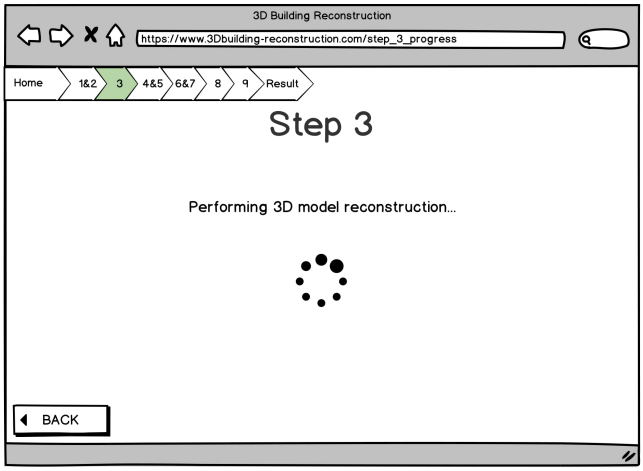

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_2_completed
Home
1&2
3
4&5
6&7
8
9
Result
Step 2 completed
Result of step 2:
click on pictures to edit manually
CONTINUE
using data produced in the previous step
DOWNLOAD
data produced in step 2
UPLOAD
your own data for step 3
BACK

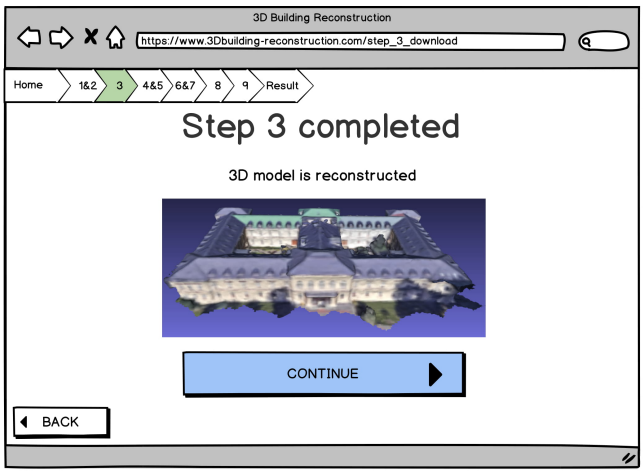
3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_2_edit_pics
Home
1&2
3
4&5
6&7
8
9
Result
Edit Pictures
Choose a picture to edit
DONE EDITTING

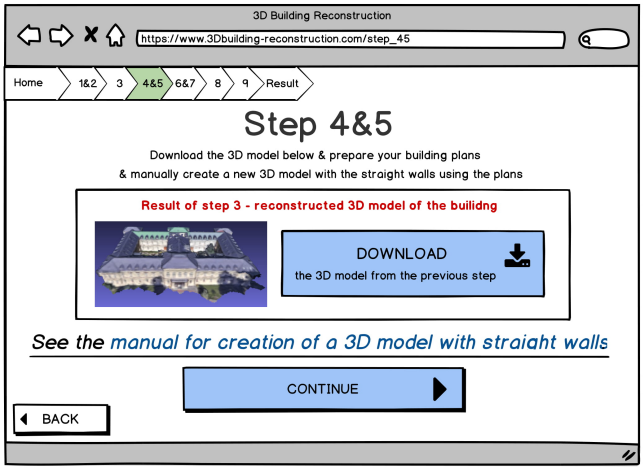

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_2_edit_pic/1
Home
1&2
3
4&5
6&7
8
9
Result
Crop
Resize
Scale
SAVE

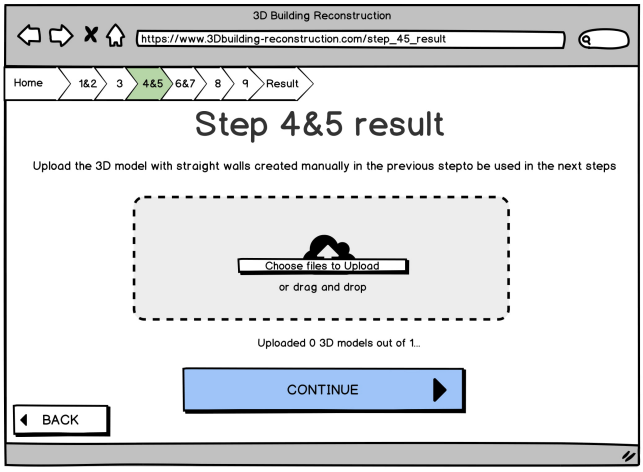
3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_2_completed_after_manual_edi
Home
1&2
3
4&5
6&7
8
9
Result
Step 2 completed
Edited pictures:
CONTINUE
using edited pictures
DOWNLOAD
data produced in step 2
UPLOAD
your own data for step 3
BACK


3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_3
Home
1&2
3
4&5
6&7
8
9
Result
Step 3
Perform 3D model reconstruction?
YES
BACK

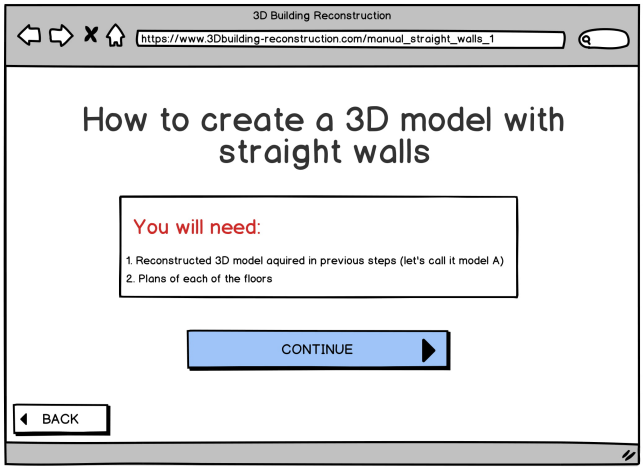

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_3_progress
Home
1&2
3
4&5
6&7
8
9
Result
Step 3
Performing 3D model reconstruction...
BACK

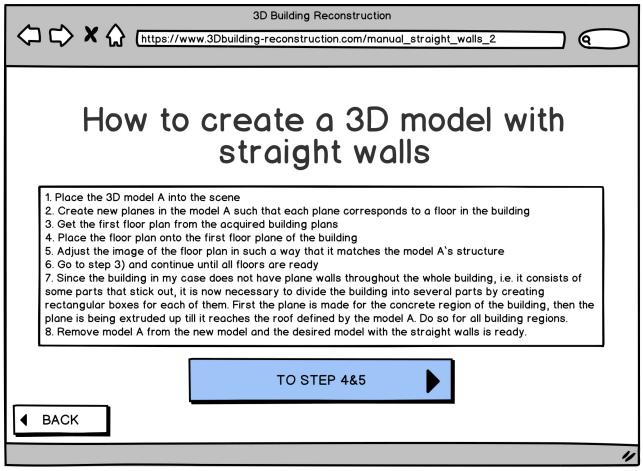

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_3_download
Home
1&2
3
4&5
6&7
8
9
Result
Step 3 completed
3D model is reconstructed
CONTINUE
BACK

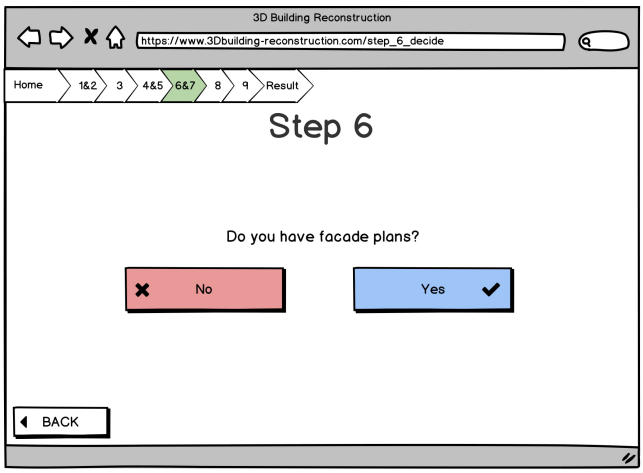

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_45
Home
1&2
3
4&5
6&7
8
9
Result
Step 4&5
Download the 3D model below & prepare your building plans
& manually create a new 3D model with the straight walls using the plans
Result of step 3 - reconstructed 3D model of the building
DOWNLOAD
the 3D model from the previous step
See the manual for creation of a 3D model with straight walls
CONTINUE
BACK

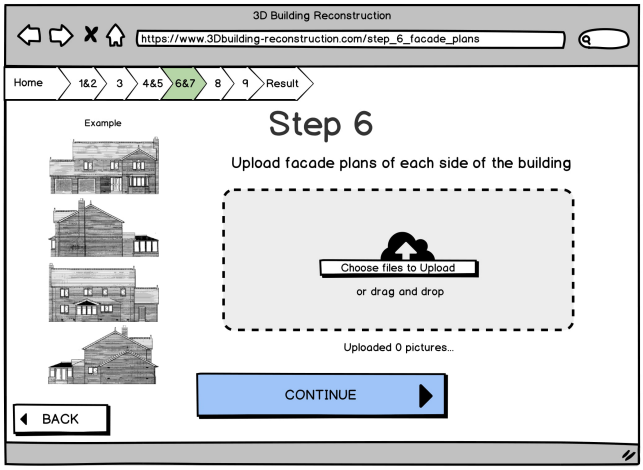

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_45_result
Home
1&2
3
4&5
6&7
8
9
Result
Step 4&5 result
Upload the 3D model with straight walls created manually in the previous stepto be used in the next steps
Choose files to Upload
or drag and drop
Uploaded 0 3D models out of 1...
CONTINUE
BACK

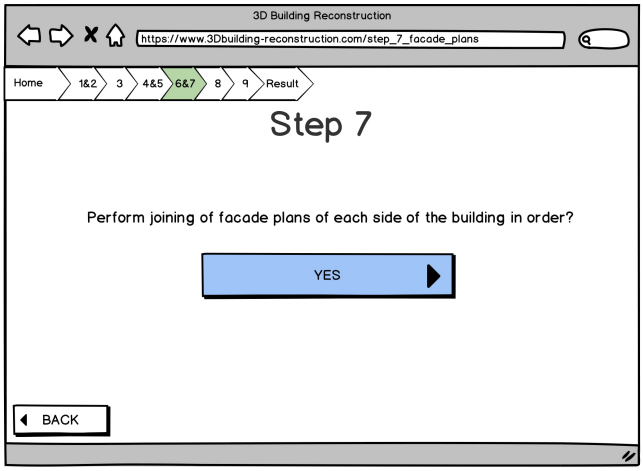
3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_45_result_visualization
Home
1&2
3
4&5
6&7
8
9
Result
Step 4&5 result
3D model uploaded
CONTINUE
BACK

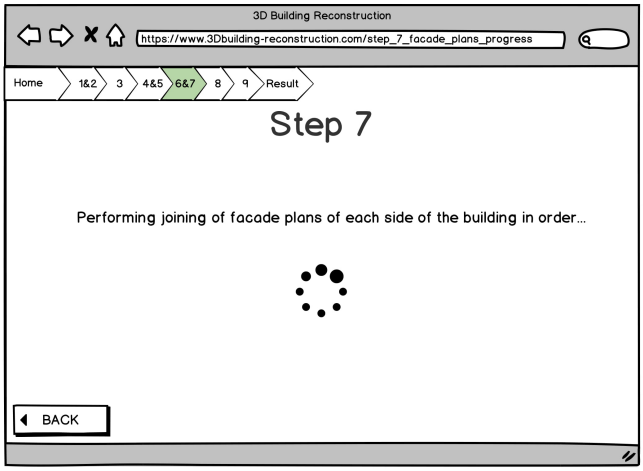

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/manual_straight_walls_1
How to create a 3D model with straight walls
You will need:
1. Reconstructed 3D model aquired in previous steps (let's call it model A)
2. Plans of each of the floors
CONTINUE
BACK

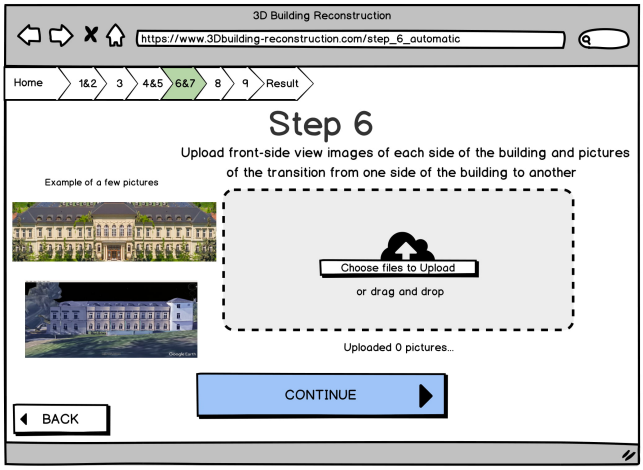

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/manual_straight_walls_2
How to create a 3D model with straight walls
1. Place the 3D model A into the scene
2. Create new planes in the model A such that each plane corresponds to a floor in the building
3. Get the first floor plan from the acquired building plans
4. Place the floor plan onto the first floor plane of the building
5. Adjust the image of the floor plan in such a way that it matches the model A's structure
6. Go to step 3) and continue until all floors are ready
7. Since the building in my case does not have plane walls throughout the whole building, i.e. it consists of some parts that stick out, it is now necessary to divide the building into several parts by creating rectangular boxes for each of them. First the plane is made for the concrete region of the building, then the plane is being extruded up till it reaches the roof defined by the model A. Do so for all building regions.
8. Remove model A from the new model and the desired model with the straight walls is ready.
TO STEP 4&5
BACK

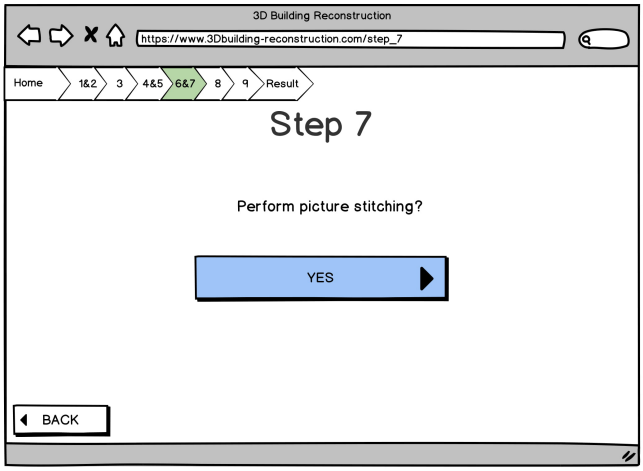

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_6_decide
Home
1&2
3
4&5
6&7
8
9
Result
Step 6
Do you have facade plans?
No
Yes
BACK

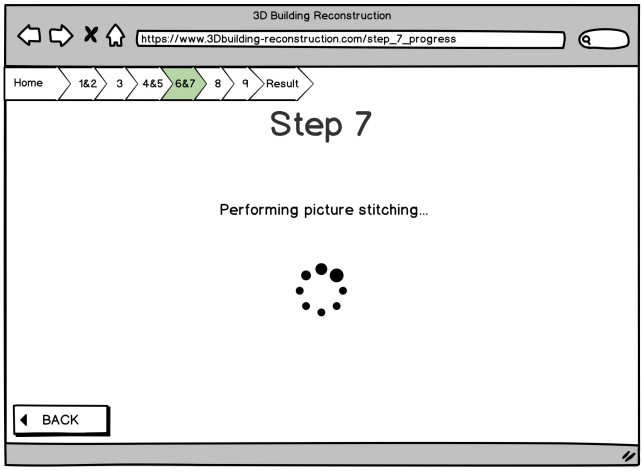

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_6_facade_plans
Home
1&2
3
4&5
6&7
8
9
Result
Example
Step 6
Upload facade plans of each side of the building
Choose files to Upload
or drag and drop
Uploaded 0 pictures...
CONTINUE
BACK

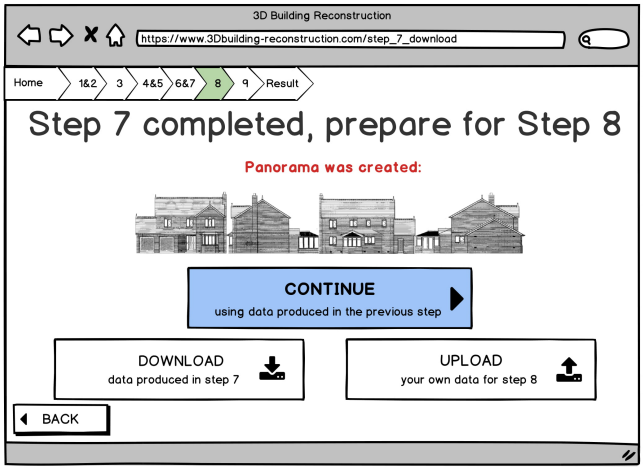

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_7_facade_plans
Home
1&2
3
4&5
6&7
8
9
Result
Step 7
Perform joining of facade plans of each side of the building in order?
YES
BACK

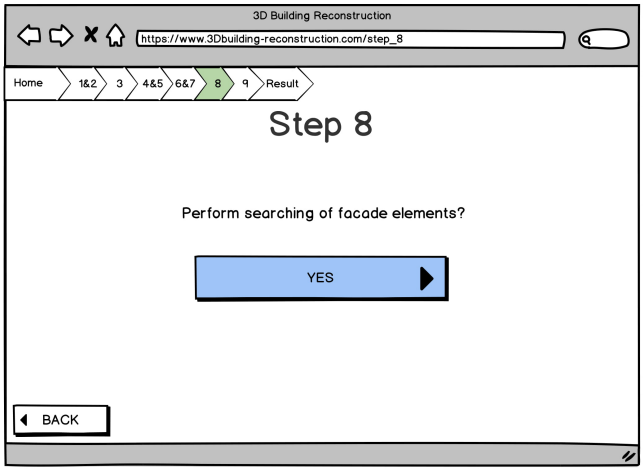

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_7_facade_plans_progress
Home
1&2
3
4&5
6&7
8
9
Result
Step 7
Performing joining of facade plans of each side of the building in order...
BACK

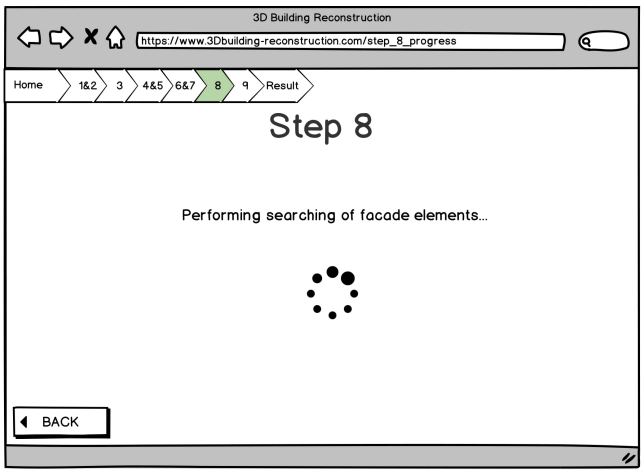

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_6_automatic
Home
1&2
3
4&5
6&7
8
9
Result
Step 6
Upload front-side view images of each side of the building and pictures of the transition from one side of the building to another
Example of a few pictures
Choose files to Upload
or drag and drop
Uploaded 0 pictures...
CONTINUE
BACK

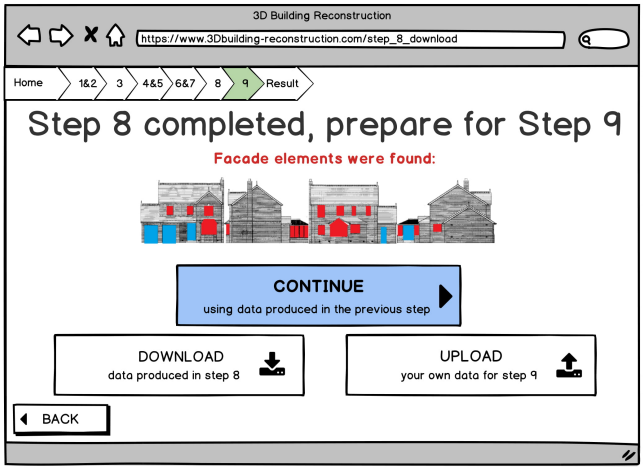

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_7
Home
1&2
3
4&5
6&7
8
9
Result
Step 7
Perform picture stitching?
YES
BACK

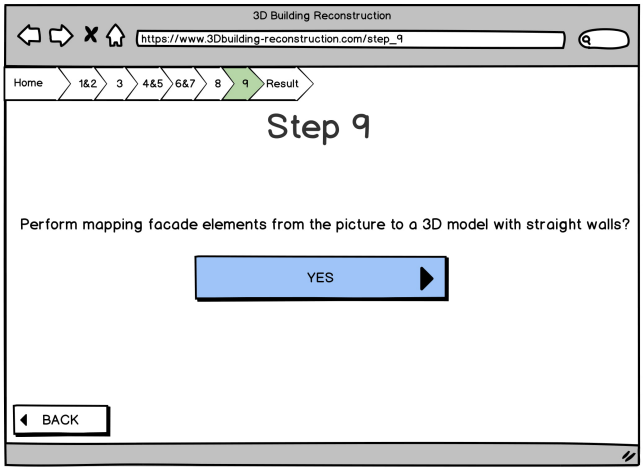

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_7_progress
Home
1&2
3
4&5
6&7
8
9
Result
Step 7
Performing picture stitching...
BACK

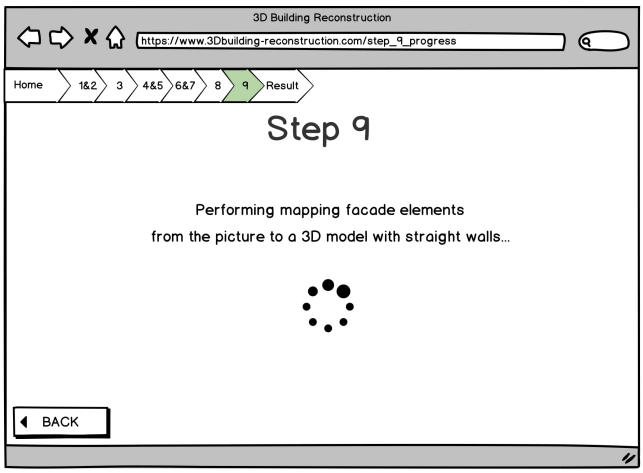

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_7_download
Home
1&2
3
4&5
6&7
8
9
Result
Step 7 completed, prepare for Step 8
Panorama was created:
CONTINUE
using data produced in the previous step
DOWNLOAD
data produced in step 7
UPLOAD
your own data for step 8
BACK

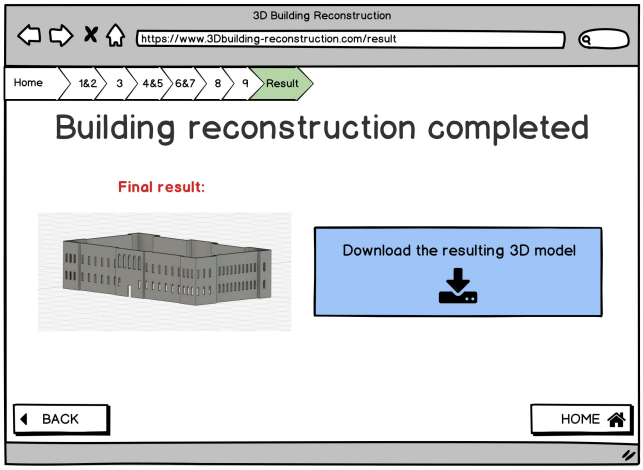

3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_8
Home
1&2
3
4&5
6&7
8
9
Result
Step 8
Perform searching of facade elements?
YES
BACK


3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_8_progress
Home
1&2
3
4&5
6&7
8
9
Result
Step 8
Performing searching of facade elements...
BACK


3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_8_download
Home
1&2
3
4&5
6&7
8
9
Result
Step 8 completed, prepare for Step 9
Facade elements were found:
CONTINUE
using data produced in the previous step
DOWNLOAD
data produced in step 8
UPLOAD
your own data for step 9
BACK


3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_9
Home
1&2
3
4&5
6&7
8
9
Result
Step 9
Perform mapping facade elements from the picture to a 3D model with straight walls?
YES
BACK


3D Building Reconstruction
https://www.3Dbuilding-reconstruction.com/step_9_progress
Home
1&2
3
4&5
6&7
8
9
Result
Step 9
Performing mapping facade elements
from the picture to a 3D model with straight walls...
BACK



Abbreviations

HCI Human-Computer Interaction

HTA Hierarchical Task Analysis

UCD User-Centered Design