Top-Down Finger Tracking on Relief Surfaces for Interactive Content Exploration

Evaluation of Existing Technology

Access to museums for blind and visually impaired people through 3D technology

Erasmus+

Funded by the European Union

This project (2014-1-AT01-KA204-001014) has been funded with support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Content

1 Introduction

Reliefs are a widely accepted method to help conveying complex images and paintings to blind and visually impaired people [45]. A common misconception is, that a touch relief should stand for itself. The same is true for all kinds of tactile material. In practice, they are only intended to be an *additional* help. The main part is still the professional description of the image or object. The tactile relief or object may help in getting the composition, and spatial relationships that are hard to explain with words alone. Further, a touchable object allows further own investigation. But even there, most people like a dialog, where they can ask about things they are touching. Especially, color and visual properties are much better explained, since vivid words stimulate the imagination much better, than the hard surface of a touch relief is capable of.

1.1 Vision

The idea of this work is to investigate whether such information can be given automatically during haptic interaction with the tactile materials. Specifically we think about position specific audio comments, triggered by a digital touch interface.

There are currently two approaches:

- In Digital Touch Replicas (e.g. [39]), touch sensors are included in a haptic exhibit. This approach has proven useful and content and parameters can be updated, but physical sensors need to be integrated in the replica.
- In purely virtual haptic interaction devices (e.g. Probos [40]), interaction regions can be easily changed since they are only defined by software, but the haptic qualities are limited since the virtual objects cannot be touched with the whole hand.

In this work, we intend to bridge this gap by developing a new digital touch interface for physical objects, which does not require any modification of the object.

Our idea is based on the recently emerging tracking sensors that allow precise three-dimensional body, hand and finger detection, specifically targeted at human computer interaction with the whole body. Examples include Microsoft's game controller Kinect, the structure.io, and LEAP Motion especially for finger-tracking. With these devices, it should be possible to track the visitor's hands and fingers on the exhibit, and to detect touch-events as well as special gestures. Based on this input, and the precise location on the exhibit, specially designed audio comments and different narratives can be interactively explored. Special gestures can be incorporated to e.g. skip or replay comments, or maybe to change interaction modes. Different exploration modes are possible, from overview descriptions, to detailed stories about different parts, measurement functions using multiple fingers, or quizzes where the visitor has to find and touch a specific part. Modes especially for visually impaired people can be easily integrated, e.g. to name the touched part, to get color or material information from under the finger or to guide the visitor over the touched parts.

These scanners are typically low-cost since they were developed for the consumer market, and are easy to operate. One or more scanners are placed above the object, and a special software will detect the touch events and generate the feedback.

The software holds a 3D scan of the object on which the interaction regions and their feedback are stored. Therefore, it is easily possible to change the interaction regions and the narrative, without any modifications on the object. The same scanner can even be used on different objects.

1.2 Related work

1.2.1 tiptoi

An interesting tool intended for sighted people is the toy “tiptoi“, marketed by Ravensburger [43]. It is a pen-shaped devices with a small camera at the tip. When placed over specially printed objects, tiny, nearly invisible dot patterns allow to recover the precise location of the pen, and may trigger audio events played over an in-built speaker. Several different products are available, like books, but also 3D models, board games or a globe. The control and audio files for each product are downloaded via internet, and allow a large variety of interactive experiences. Most products feature different information layers or modes, for example: discover, knowledge, narration and games. Depending on the activated mode, touching the same spots may trigger different audio. Game modes allow even more interactivity like finding objects, sorting things, keeping track of scores, and so on.

The tiptoi is not directly suitable for 3D exhibits, since the coded print can only applied to mostly planar faces, and the toy itself is probably not sturdy enough and might be easily stolen. But the large variety of interactive possibilities is definitely a big inspiration for this project.

1.2.2 Probos

The Probos interactive [40] goes one step further and allows touching of virtual three-dimensional objects. Based on a haptic force-feedback human-computer input device (e.g. [46]), the user may touch, feel, and even interact with virtual objects. These haptic input devices are like a small robotic arm with rotation sensors and motors in all joints. At the end of the arm is typically a stylus or knob that can be grasped and moved around. With the measurements of the rotation sensors, the 3D position of the tip of the stylus can be computed. If this position touches a virtual object, the motors build a force against the user's motion, creating the sensation of touching an actual object.

Either scanned real-world objects or completely computer generated, any object can be touched, as long as a digital 3D file of the surface is available. Since all touch events are virtual and happen in the computer, these events can be easily used to trigger pop-up information or audio comments, depending on the touched location. In the real world, this is unfortunately more difficult.

1.2.3 Embedded Touch Sensors

Touch Graphics Inc. [41] developed several tactile exhibits or tactile interactive maps amongst others, all built around interactive tactile technology. By integrating touch sensors into the objects, sounds or visuals can be triggered, enriching the object, and giving different narratives.

Similar is the interactive stele of Hesysunebef [39], developed by Loughborough University and exhibited at Manchester Museum. A 3D-printed hand-painted one-to-one replica of an Egyptian stone plate is equipped with 22 touch sensors. Touching one of the figures, a short back story is given in audio and as graphics and text on a monitor next to it.

This technology allows direct interaction on 3D objects, but special sensors need to be placed inside. While the narrative can be easily adapted, the placement of the sensors is not, and has to be planned beforehand. Such sensors may not be implemented on original artworks. And the sensors may not be reused on other exhibits.

1.2.4 Talking Tactile Tablet

Touch Graphics developed the Talking Tactile Tablet [42], a multi-purpose touch sensitive tablet on which raised-line graphics may be placed, to get interactive narratives, depending on where the graphics are touched. A simple authoring tool allows to designate areas and assign audio files. This technology however only works on nearly flat objects, and not on high relief surfaces or fully three-dimensional objects.

1.2.5 3D-Finger

3D-Finger [44] was a feasibility study for adding narratives on 3D objects without the need for internal sensors. A number of video cameras were placed around the object and could track a specially colored dot on the user's fingertip, and thus, allowed touch detection.

Our intent is similar to this study, but we intend to avoid the need of colored dots, but use novel finger tracking technology based on depth camera sensors.

2 Requirements on the Tracking Cameras

While it is difficult to give exact figures or features that will be required by the tracking cameras for our application, at least a general direction and some key figures can be estimated.

The optimum would be an out-of-the-box direct solution for articulated finger tracking, i.e., a technology that reconstructs the position and orientation of all parts of the complete skeleton of each detected hand. With such a technology at hand, it should be straight forward to detect touch events, and even other gestures, and we can invest more time into different interaction possibilities.

Since we intend to use the tracking on tactile reliefs, the target size needs to be at least the size of the reliefs, which, taken from past events, is around DIN A3 (420×297mm). The users will mostly interact near the surface of the relief, but may also perform some gestures above the relief. The track-able interaction space should therefore have a minimum height of at least 25cm.

According to our vision, the camera should be mounted above, or diagonally behind the relief, so that it oversees the whole interaction volume, but still does not disturb touch and vision (for people with rest of sight, and for seeing people, who might also want to use the installation). Observation from below is not possible, since our surface is not planar but a relief, and which cannot be made see-through. In this configuration, touch events cannot be measured directly, since we may only scan the backsides of the hands. From the backside of the fingertips, and a guess of the finger diameter, we should be able to distinguish touch events from hovering over the surface.

A key requirement is therefore that the sensors need to be able to reliably distinguish the fingers from the relief surface, even if the relief is touched. The sensors need to have a small enough error in the depth measurements, and low enough noise to allow reliable separation of the fingers from the background in order to detect touch events. In addition, the spatial resolution of the sensors need to be high enough to be able to distinguish the different, often small areas in a tactile relief.

Further, it is beneficial, when the relief itself can be scanned or tracked as well, in order to account for drifts in the setup, and to get a reference of the exact height at the different relief locations.

3 Evaluation

Due to budgetary and time constraints not all available depth cameras could be bought and tested. We had to rely on testing devices we had access to, and on online resources for the other devices. In the following, we discuss available depth cameras and software projects specialized on finger and hand tracking.

3.1 Leap Motion

Announced in 2012, Leap Motion [2] announced a device that allows fully articulated hand and finger tracking, as a novel method for human computer interaction, based on gesture detection and precisely detecting where the user is pointing with her fingers, or even with pointy objects like pencils. They advertised a much higher resolution than the then known Kinect sensor (see Section 3.2) with sub-millimeter precision. Videos were showing its superior quality and precise control, with users waving their hands over the 3×1×0.5 inch (7.6×2.5×1.3cm) small sensor. Full skeleton tracking was demonstrated, showing a very stable detection of the position and orientation of all 19 finger bones and the palm of the hand in all possible poses. Special gestures can be detected like waving in the air or touching virtual planes.

The device seemed to be perfect for our purpose as it directly computes the position of all fingertips in the three-dimensional space, even if not fully visible, which is needed to detect touch events. In fact, the announcement and videos inspired our project, and led to the proposal for the now-granted part of the project AMBAVis.

There were also videos showing a sort of point cloud, raising the hope in many people, that the device was in fact a depth camera.

Technology

After the device came to the market, reports of people disassembling the device were published (e.g. [3]), showing that the device is not a depth sensor as many had hoped, but a comparatively simpler stereo camera. The tiny housing holds three infrared LEDs, lighting the space above the device with bright, but for humans invisible infrared light, and two infrared cameras equipped with fisheye lenses to provide a wide field of view and a large operating volume at close distances. This was required, since the device is designed to lie on the desk, with the hands hovering above it. It also became obvious, that the black and white camera images are directly transmitted to the computer over USB, and no processing is performed on the device. All hand-detection and skeleton fitting is performed on the computer using “complex math”. Nevertheless, the videos showed good finger tracking in many hand positions, but also often tracking problems.

Evaluation

We could borrow a Leap Motion from a colleague before purchase. Installation was easy, and the developer SDK seemed to be well written and well documented. However, tests did not work as expected.

In our office environment, hand tracking in normal bottom-up mode worked, but had often severe problems, giving strongly twisted finger positions, and registering hands the wrong way around (e.g. left hand as right hand, and 180° turned). When testing the device upside-down, mounted above the desk, and pointing on the hands below, it turned out, that the device was simply not built for this task. Although there is a VR mode, normally used when mounting the Leap Motion on a VR-helmet on the forehead, which seems to detect hands better when viewed from their back side, it did not work in our setup, with the hand close to or even touching a surface. It seems the device is especially built for mid-air operation, where the infrared LEDs only light the nearby hands, but leave the whole distant background dark, presumably to aid in easier foreground background segmentation.

We could not get any reasonable results for our use case. It seems, that we need to rely on real depth camera sensors that can better separate the hands from the nearby background.

3.2 Microsoft Kinect for XBOX 360

Introduced in 2010, this device was the first depth camera targeted at the mass market, with a price of around 100 USD. Unfortunately, all official web sites seem to be no longer online, but there is e.g. an article on Wikipedia [4], especially in the Section History. The device was developed as an additional input device for Microsoft’s Xbox 360 game console, with the purpose of whole body tracking allowing to play games with the whole body.

Despite its original intent, the device was largely adapted for scientific and enthusiast’s applications, since it allowed low-cost access to the technology of depth cameras, and the protocol was quickly reverse engineered making it usable outside the XBOX game console. In any case, it set off a trend towards this novel technology, and a lot of similar devices and software using it appeared over the years.

Technology

Its core is a novel depth sensor, developed by PrimeSense, which projects an invisible infrared pseudo-random pattern of bright and dark spots [1] onto the scene, using a LASER projector. This pattern is filmed with an infrared camera that is offset 7.5cm from the projector, equipped with a narrow bandpass filter tuned to the wavelength of the LASER. The offset between camera and projector causes a horizontal offset of the projected points, depending on the distance of the surface hit by the projector. From these offsets, the depth of a number of sample points relative to the camera is computed directly on the device, and transmitted via a USB 2.0 video stream to the computer for further processing. Since only about 1/9-th of the projected spots are bright, the effective resolution of the probably around 210×160 pixels [1], but that is hidden in that the interface gives a depth image stream of 320×240 with up to 2048 discretized depth values. This image is already filtered and processed in an undocumented way.

In addition, it features an RGB color camera with a resolution of 640×480 pixel with 30 fps, and a microphone array that allows spatial localization of voice commands of up to 4 players as an additional input.

The housing sits on a tilt mechanism that may be used to automatically tilt the cameras up and down to adjust for optimum scene coverage. The device requires a quite high amount of power, for that it requires an additional power supply and cannot be run on the power provided by the USB 2.0 standard alone.

Tests

Since we already owned a Kinect 360 sensor from past projects, we could quickly do tests related to finger/hand tracking. Since the Kinect was built for whole body tracking, its nominal operation range is limited to distances of 0.8-3.5 meter. At longer distances, the laser dots become too dim to be detected, at shorter distances the laser dots become too bright and their images fuse together and cannot be detected any more.

With a horizontal field of view of 57.5° and a vertical of 43.5°, the area covered at minimum distance is 88×64cm, which is already more than double the size we need, effectively halving the already low resolution. With our requirement for a working volume with a minimum height of 25cm, the area increases at a distance of 1.05 meters to 115×84cm, leaving a resolution of over 3.5mm per pixel assuming a sensor resolution of 320×240, or even over 5mm with the 210×160 resolution derived from the projected pattern, which is quite low for finger detection. In addition, there is some amount of noise in the depth measurements, depending on the distance, the color of the objects, and the lighting situation, with up to 5mm noise.

We conclude, that the Kinect 360 may work, but is definitely not optimal for our purpose.

3.3 Kinect Variants

The Kinect for Xbox 360 was never intended for usage outside Xbox applications, and was largely driven by the open source movement. These led to clever developments by the scientific community and enthusiasts, showing the large potential of this new technology. Several companies reacted on this and developed variations of the original Kinect.

3.3.1 Re-Packaged Variants

The original inventor of the technology, PrimeSense, offered other companies the technology.

Asus created the product line ASUS xtion [5], with probably the same hardware, but in a smaller and lighter case. The most notable improvement was that it would run on the power of a USB port alone, removing the need for a separate power supply. There are variants with and without color camera.

Funded by a Kickstarter campaign [6], Occipital developed the Structure sensor [7], a variant in a very small case intended to be used on handheld tablet computers, also operated on USB power alone.

Microsoft launched the Kinect for Windows, basically the same as the Kinect for Xbox 360, but with a different license, officially allowing use outside the Xbox 360 environment.

PrimeSense also launched its own product line called Carmine [8], with the Carmine 1.08 very similar to the ASUS xtion.

These variants are very similar, and share probably the same core technology, but some differences have been documented [9] [13].

Together with these variants came official software and drivers. PrimeSense offered their NiTE Middleware, and Microsoft their Kinect for Windows SDK, both offering diverse post processing capability, like body tracking, but no finger and hand tracking.

The most notable functional extension was the introduction of a near mode, which in theory offers tracking up to 40 cm in a special mode. This slightly increases the resolution to around 2.5mm (with 320×240) or 3.3mm (210×160) per pixel with a working distance of 65 cm. Also noise was reduced a bit, making the Kinect more suitable, although not perfect. Test was performed with a Kinect for Windows, bought in former project. Testing all variants is probably not necessary, since all are built around the same hardware.

3.3.2 Carmine 1.09

All above variants are only re-packaged versions with presumably the same, or very similar hardware. The only exception is PrimeSense's product line [8]. The Carmine 1.09 was specifically designed to operate in near mode (0.35-1.5m). According to its appearance, at least the projector design was changed. Further products were planned.

Unfortunately, PrimeSense was bought by Apple Inc. in 2013 [10], and these variants are no longer officially available. There are some eBay and amazon shops, where remaining stock is sold for a multiple of the original costs. The only official source found was the face tracking software company faceshift, who obviously branded their own variant of the Carmine 1.09 [9]. However, they only sell their cameras bundled with their software, which would have exceeded our available budget.

3.3.3 Lens Modifications

An interesting modification kit is offered by Nyko [11]: The snap on device features 3 lenses in front of the projector, IR and RGB cameras, and is reported to reduce the working range by 40%. This modification was also reported for the Carmine 1.09, using off-the-shelf +2.5 diopter reading glasses mounted in front of the projector and camera [12].

However, we did not try this, since a) the near mode of the Kinect allowed us already to get as close as necessary without introducing lens distortions; and b) since we could not get a Carmine 1.09.

3.3.4 Orbbec 3D

The newest announcement is the Orbbec 3D depth sensor range, which seems to be based on a similar technology. Little is yet known, but this device may be interesting, as they claim to have improved upon the Kinect technology in terms of range and quality.

3.4 Time of Flight Depth Sensors

The Kinect and its variants are all based on a static projected pattern used for depth estimation across the image plane. Time of flight measurement is another technology developed for this purpose. It is based on the measurement of time that light emitted from the sensor needs to travel in order to reach an object in the scene, being reflected back and, finally, reach back to the sensor where it is registered. From the (nearly) constant speed of light and the time needed, the distance can be measured. In practice, the tiny time difference is not measured directly, but is typically detected as a phase shift of a pulsing light, averaged over a longer period.

Recently, this technology became affordable, and even low-cost devices are now available, entering the mass market. Current examples are the *SoftKinetic DS325* [14] and rebranded as *Creative Senz3D* [15], the *Kinect for Xbox One* [16] and rebranded as *Kinect for Windows v2,* and various sensors of pmd Technologies, like the CamBoard pico [17] that was to be bundled in the nimble UX [18] project (see Section 3.6).

Although each sensor has its advantages and disadvantages, there are some general observations. One advantage of this technology is that there is a measurement for each pixel of the sensor at a large frame rate. Often additional measurements are available as a byproduct of the measurement process, e.g., the scene fully lit by the sensor's light source and without this light, or a confidence value of the depth measurement. The Kinect for Xbox One is able to measure a person's pulse from small infrared fluctuations on her skin. However, the sensor range is often limited (which should not be a limitation in our case), the sensor resolution is typically rather small, since the electronics for each pixel are more complicated. But worst for our application, the depth measurements are typically very noisy with around 1 cm even in the near field (e.g. SoftKinetic < 1.4cm at 1m), often depending on the color (absorption of infrared).

The SoftKinetic DS325 has a bundled hand-tracker software, but judging from online videos [19] fingertip detection is very noisy. There is a research project, demonstrating hand tracking using the Microsoft Kinect for Xbox One [20], but no code or executable was published, yet.

We only did a test with the Kinect for Xbox One, which was available at our lab from a previous project. The noise level was indeed significantly higher than with the Kinect 360. We had no access to the other sensors. The Creative Senz3D was previously part of Intel's Perceptual Computing SDK, which is now replaced by the RealSense technology.

3.5 Intel RealSense

RealSense [21] is Intel's answer to novel imaging techniques to be embedded into handheld and laptop devices, and a collection of software and middleware to effectively use this technology. It currently consists of

- a light-field camera RealSense Snapshot;
- a larger distance depth camera RealSense R200 meant to be mounted rear-facing in order to take pictures of farther away objects, e.g. to scan an entire room; and
- a short range depth camera RealSense F200 [23], to be mounted front-facing used for self-portraits or videos for chat applications, or for finger-tracking as a natural interface.

These are already available inside some laptop and tablet computers [22].

The device with the largest potential for our application is the front-facing short-range RealSense F200. Designed by Creative [24], it is based on a different technology using an active (i.e. changing) projected pattern. According to [25] a time-multiplexed Grey-coded stripe pattern [26] is projected in consecutive frames. The given photo shows several Grey-patterns stacked on top of each other, which we believe is a side effect of a rolling shutter of the used camera, as typically a pattern with stripes along the whole height is used, one Grey-code bit plane at a time. This is underpinned by a teardown of the device [27] showing the projector composed of a laser widened into a vertical line by a line lens, and then deflected by a vibrating MEMS mirror. Thus, if the laser diode is pulsed accordingly, the typical Gray stripe patterns can be produced. A high-speed infrared camera registers these between 7 and 9 (user selectable) coded frames, and computes a depth estimate from these measurements.

We could borrow such a device before purchase. Advantages are a full 640×480 depth resolution with a depth range of 0.2 – 1.2 meter, a field of view of 90°, and high quality depth measurements with low noise. Noise level seems to be the lowest for such low-cost devices to date, which can probably be attributed to the technology, using a number of consecutive frames per measurements. This technology is used in a number of professional scanners as well (e.g. [47]), but these don't offer real-time speed as available here. The device is also equipped with a 1080p RGB camera, and can transmit a lot of high-quality data streams via its USB 3 interface, over which it is also powered, without the need for an external power source.

One disadvantage of using consecutive frames is that fast moving objects may produce measurement errors, although this is of minor importance for our application, since the hands will not move very fast during tactile exploration. In addition, the device offers a lot of hardware settings to be adjusted, allowing to minimize such effects: e.g. the number of frames used per measurement (tradeoff between capture time and depth resolution), laser power, integration time, amount of processing/filtering, and so on.

The accompanying SDK features several post-processing algorithms including articulated finger tracking. The latter turned out to be less suitable for our purpose, having troubles tracking hands captured from the back side. But it features a good depth resolution allowing to distinguish fingers from touched objects. The working range is sufficient for reliefs of size A3. For larger objects, a larger distance is required, quickly degrading the scan quality, probably due to too little laser power. Therefore, we judge, that it is probably very usable for our intended application.

As a side note, the SDK requires Microsoft® Windows® 8.1 and above, which might not yet be available on all computers. Fortunately, enthusiasts found out, that the protocol is based on the USB Video standard, and it could already be made to work on Linux [28]. If necessary, we believe we can adapt it to previous Windows versions as well.

3.6 Nimble VR

The former company 3Gear Systems [29] seems to be the only company solely focusing on articulated finger tracking based on depth camera measurements. End of 2014 they teamed up with the hardware developer pmd Technologies GmbH [18] and formed Nimble VR [30]. Through a Kickstarter campaign [31], interested people quickly pledged more than twice of what they requested, but it was cancelled two days before the end, since the 3Gear part was acquired by Oculus (which was bought by Facebook before).

Although the start page now only informs about the acquisition, it is still possible to find information about the previous Nimble UX beta phase [32] and also about the Nimble SDK [33]. It is still possible to download the software, but it is no longer possible to get the needed license files.

According to video demos, the nimble/3Gear software is perfectly capable of tracking hand gestures in mid-air, but even on a surface [34]. According to the video, even touch events of all 10 fingers can be tracked in a high quality. It seems very stable, with online adaption to the hand size of a user and has a low latency. The software was compatible with a wide variety of depth cameras, and was even demonstrated with two Kinect depth cameras concurrently, although this is no longer recommended [35].

This software would probably be a perfect basis for our project.

3.7 Other Software

Noteworthy other software projects for articulated finger tracking with depth cameras are AnyTouch [36], Touchless Touch [37] and Sigmanil [38]. While they are mostly concerned with mid-air virtual touch event detection, gesture detection and other applications, they could be interesting for our problem. Touchless Touch is capable of registering touch events on surfaces, but mainly with the whole hand. Sigmanil has an additional module, developed for articulates hand detection.

4 Conclusion

From the devices that we could test with our limited budget and time, and the web research we conclude, that there may not be the optimal device for our purpose on the market, yet. Several products were looking promising, but turned out to be not suitable for our purpose. In addition, several developers were bought by large companies, and ceased to deliver products or development kits. This however suggests, that there is a large interest in this kind of devices, and that a number of suitable depth cameras will probably become available in the near future.

A very good sensor comparison of available sensors is available on the website of faceshift [9], specifically targeted at face scans, which is, at least to a certain extent, applicable to our problem. The operating range is similar, and the detail and amount of noise can be judged from the scans of faces as well. Only, when scanning hands, the gaps between fingers may pose additional challenges, and the proximity to the background surface can only be compared to the protruding nose from the face. According to their ranking, the Time of Flight sensors “Microsoft Xbox One Kinect” and “Creative Senz3D” are not acceptable, “Kinect for Xbox 360” and variants are good, and their Carmine 1.09 variant and Intel RealSense F200 are the best. Judging from the screenshots, the RealSense may even get a better and cleaner scan. This is in line with our own observations.

For now, the Intel RealSense F200 seems to be the most suitable device for our purpose on the current market, although with no working finger tracking solution.

A similar situation is for software that offers finger and hand tracking based on depth cameras. Either the developers were bought by large companies, or the products already available are not satisfactory. All tested products only work with the hands in mid-air, but not when operating on a surface. Online resources suggest, that there are already working solutions, but are no longer available for the public.

For the project at hand this means, that we cannot rely on a hand tracking product and concentrate on the interaction as anticipated. Instead we need to develop the tracking software ourselves. As this is a major endeavor, we will have to limit ourselves to a minimum tracking possibility, in order to at least demonstrate and test the basic concept and feasibility of our approach. With the limited resources at hand, we will try to implement at least single finger interaction.

On the other hand, if it may become possible to use one of the working finger-tracking solutions (especially Nimble VR [33] or the scientific prototype based on the Kinect for Xbox One [20]), this would be a huge advantage for our project, as we could proceed as originally anticipated.

References

[1] Andreas Reichinger, Kinect Pattern Uncovered, https://azttm.wordpress.com/2011/04/03/kinect-pattern-uncovered/, accessed February 2015.

[2] Leap Motion website, http://www.leapmotion.com, accessed April 2015.

[3] J. Phung, Leap Motion Teardown in Macro HD, https://www.youtube.com/watch?v=uF0NSUmxFYA, accessed April 2015.

[4] Kinect, Wikipedia, https://en.wikipedia.org/wiki/Kinect#History, accessed April 2015.

[5] Asus, xtion 3D sensors, https://www.asus.com/3D-Sensor/, accessed April 2015.

[6] Kickstarter, Structure Sensor, https://www.kickstarter.com/projects/occipital/structure-sensor-capture-the-world-in-3d, accessed April 2015.

[7] Structure sensor website, http://structure.io/, accessed April 2015.

[8] PrimeSense, Sensors, http://web.archive.org/web/20130807161100/http://www.primesense.com/solutions/sensor/, accessed April 2015.

[9] Faceshift, Comparison of Depth Cameras, http://www.faceshift.com/cameras, accessed April 2015.

[10] M. Isaac, J. Paczkowski, Apple Confirms Acquisition of 3-D Sensor Startup PrimeSense, http://allthingsd.com/20131124/apple-confirms-acquisition-of-3d-sensor-startup-primesense/, accessed April 2015.

[11] Nyko, Zoom for Xbox 360® Kinect®, http://nyko.com/collections/xbox-360/products/xbox-360-zoom-for-kinect, accessed April 2015.

[12] Christoph Heindl, Zoom Lense Results, http://reconstructme.net/2012/03/18/gallery-update-2/, accessed April 2015.

[13] iPiSoft, Depth Sensor Comparison, http://wiki.ipisoft.com/Depth_Sensors_Comparison, accessed April 2015.

[14] SoftKinetic, DepthSense cameras, http://www.softkinetic.com/Products/DepthSenseCameras, accessed April 2015.

[15] Creative, Senz3D depth camera, http://en.europe.creative.com/p/web-cameras/creative-senz3d, accessed April 2015.

[16] Microsoft, Kinect for Xbox One, http://www.xbox.com/en-US/xbox-one/accessories/kinect-for-xbox-one, accessed April 2015.

[17] pmd Technologies GmbH, CamBoard pico, http://www.pmdtec.com/products_services/reference_design_pico_pico_s.php, accessed April 2015.

[18] pmd Technologies GmbH, nimble UX, http://pmdtec.com/nimbleux/, accessed April 2015.

[19] SoftKineticUSA, SoftKinetic DS325 vs. Kinect Video, https://www.youtube.com/watch?v=s5M09Lq7FzU, accessed April 2015.

[20] Toby Sharp, Cem Keskin, Duncan Robertson, Jonathan Taylor, Jamie Shotton, David Kim, Christoph Rhemann, Ido Leichter, Alon Vinnikov, Yichen Wei, Daniel Freedman, Pushmeet Kohli, Eyal Krupka, Andrew Fitzgibbon, and Shahram Izadi, Accurate, Robust, and Flexible Real-time Hand Tracking, CHI, April 2015.

[21] Intel, RealSense Overview, http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html, accessed April 2015.

[22] Intel, Devices with RealSense Technology, http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-devices.html, accessed April 2015.

[23] Intel, RealSense F200 Camera, https://software.intel.com/en-us/RealSense/F200Camera, accessed April 2015.

[24] Creative Technology Ltd., Intel RelaSense F200, http://support.creative.com/kb/ShowArticle.aspx?sid=124661, accessed April 2015.

[25] Intel RealSense Forum, Entry about how the Intel RealSense F200 works, https://software.intel.com/en-us/forums/topic/537872, accessed April 2015.

[26] J. Salvi, J. Pagès, J. Batlle, Pattern codification strategies in structured light systems, Pattern Recognition, Volume 37, Issue 4, April 2004, Pages 827–849.

[27] **S. Dixon-Warren**, Inside the Intel RealSense Gesture Camera, http://www.chipworks.com/about-chipworks/overview/blog/inside-the-intel-realsense-gesture-camera, accessed April 2015.

[28] D. P. Johnson, Intel Real Sense Camera on Linux, http://solsticlipse.com/2015/02/10/intel-real-sense-on-linux-part-2-3d-camera-controls.html, accessed April 2015.

[29] 3Gear Systems, website, http://threegear.com/, accessed April 2015.

[30] Nimble VR, website, http://nimblevr.com/, accessed April 2015.

[31] Kickstarter, Nimble Sense: Bring Your Hands into Virtual Reality & Beyond, https://www.kickstarter.com/projects/nimblevr/nimble-sense-bring-your-hands-into-virtual-reality/, accessed April 2015.

[32] Nimble VR, Nimble UX beta, http://nimblevr.com/nimbleux/, accessed April 2015.

[33] Nimble VR, Nimble SDK, http://nimblevr.com/download.html, accessed April 2015.

[34] 3GearSystems, Tracking arbitrary motion of all 10 fingers with the 3Gear DevKit, https://www.youtube.com/watch?v=exZ6wukQCpk, accessed April 2015.

[35] Nimble VR, Notes on supported depth cameras, http://nimblevr.com/latest/doc/cameraNotes.html, accessed April 2015.

[36] Ayotle, AnyTouch Technology, http://www.ayotle.com/, accessed April 2015.

[37] Touchless Touch, website, http://www.touchlesstouch.com/, accessed April 2015.

[38] Sigmanil, website, http://www.sigmanil.com/, accessed April 2015.

[39] Judy Wing, Ancient hieroglyphics meet cutting-edge technology at Loughborough University, PR 12/197, 9 November 2012. (http://www.lboro.ac.uk/service/publicity/news-releases/2012/197_Manchester-Museum.html, accessed April 2015)

[40] Touch & Discover Systems, the Probos interactive, http://www.touch-and-discover-systems.co.uk/probos-2/, accessed April 2015.

[41] Touch Graphics Inc., Portfolio, http://touchgraphics.com/portfolio-2/, accessed April 2015.

[42] Touch Graphics Inc., Talking Tactile Tablet, http://touchgraphics.com/OnlineStore/talking-tactile-tablet-2-ttt.html, accessed April 2015.

[43] Ravensburger, tiptoi, https://www.tiptoi.com, accessed April 2015.

[44] 3D Finger, AAT – Zentrum für angewandte assistierende Technologiedn, Technische Universität Wien, http://www.aat.tuwien.ac.at/3dfinger/, accessed April 2015.

[45] A. Reichinger, S. Maierhofer, and W. Purgathofer. High-Quality Tactile Paintings. ACM J. Comput. Cult. Herit. 4, 2, Article 5 (November 2011), 13 pages.

[46] Sensable, Phantom Omni Haptic Device, http://www.dentsable.com/haptic-phantom-omni.htm, accessed April 2015.

[47] Aicon 3D, Breuckmann Scanner, http://aicon3d.com/products/breuckmann-scanner.html, accessed April 2015.

Acknowledgements

Thanks to Michael May (VRVis), for borrowing his Leap Motion, and Martin Humenberger (Austrian Institute of Technology) for borrowing one Intel RealSense F200 before purchase to safe our budget.

Author: Andreas Reichinger | VRVis Zentrum für Virtual Reality und Visualisierung Forschungs GmbH
(August 2015)

AMBAVis – Access to Museums for Blind and Visually Impaired People through 3D-Technology

Erasmus+ project - cofounded by the European Union

www.ambavis.eu
office@ambavis.eu

Project Partners:

Economica Institute of Economic Research
VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH
Österreichische Galerie Belvedere
Austrian Federation of the Blind and Partially Sighted (BSVÖ)
Trnka n.o.
German Association for Blind and Visually Impaired People
Manchester Museum