EUROGRAPHICS Workshop on Graphics and Cultural Heritage (2017)
R. Sablatnig and B. Štular (Editors)

Short Paper

Multisensory Virtual Experience of Tanning in Medieval Coventry

Y. Dong[1,4] M. Webb[2] C. Harvey[3] K. Debattista[4] A. Chalmers[4]

[1]College of Mechanical & Power Engineering, China Three Gorges University, China
[2]Department of Archaeology and Ancient History, University of Leicester, UK
[3]Department of Creative Technologies, University of Bournemouth, UK
[4]WMG, University of Warwick, UK

Abstract

In the medieval period, Coventry, in the English Midlands was a major centre for tanning as well as for its better-known cloth industry. Heavily damaged during the Second World War, and unsympathetically rebuilt in the 1950s and 1960s, there is little left in modern Coventry to remind visitors of this important period in Coventry's history. The tanning of cattle hides was a labour intensive and smelly process. After the hooves and horns had been chopped off the animal skins from butchers, the skins were immersed in pits filled with dog dung or lime, scraped and then sprinkled with urine and left to rot in a warm environment. Finally the hides were soaked for up to twelve months in a tanning liquid to achieve the desired quality of leather. To fully appreciate the medieval tanning process, any virtual recreation needs to be multisensory: the sites of the skins in various stages of processing, the shouts of the workers, and the pungent smells. This paper presents the process of recreating this important multisensory experience and discusses how this can provide visitors with a richer experience of Coventry's illustrious past.

CCS Concepts

•***Computing methodologies*** → ***Virtual reality;*** *Ray tracing; Perception;* •***Applied computing*** → *Architecture (buildings);*

1. Introduction

Virtual Reality (VR) offers the possibility of recreating environments which no longer exist. This is particularly valuable as VR can allow people to experience cultural heritage sites which have long since disappeared. The real world is multisensory with sights, sounds, smells and feelings, including motion, temperature, etc, combining to provide richness to any real experience. Traditional VR systems, however, typically only include a few senses, such as visuals or visuals and audio. If we are to truly experience the past as our ancestors might have, then VR has to include more senses. This is especially important as the impact of one sense on another, so-called cross-modalities, can have a significant effect on how a site is perceived [CSS04]. However, if we are to recreate past multisensory experiences, then particular attention has to be given to the authenticity of any sense we are simulating. Inaccurate sensory simulation could lead to the serious problem of "misrepresenting the past".

The city of Coventry is in the West Midlands of England. Of Saxon origin, by the 15th century Coventry had grown to one of the largest and most important cities in England and was well known throughout Europe for its wool and cloth exports. Unfortunately many of Coventry's medieval buildings were destroyed by enemy action in 1940-41. But in one street, Spon Street, a significant number of medieval houses survived, only to have several of them demolished in the early 1960s when an inner ring road was cut through the area. Fortunately in 1967, the City Council decided to launch a scheme to preserve the remaining houses and indeed relocate surviving medieval houses from elsewhere in Coventry to join the *i*n situ houses.

In the medieval period, Spon Street was a major thoroughfare into and out of Coventry and also a centre for tanning and dyeing. The road is located just outside of the city walls, as 15th century urban regulations, as detailed in the Coventry Leet Book [DH08], required these trades with their noxious odours to be on the outskirts of town. The buildings are near a source of running water which was required for the tanning process; in this case the river Sherbourne which ran to the south of Spon Street.

The surviving medieval buildings, as can be seen in Figure 1(a), are used as shops and restaurants, and not for their original purpose. Virtual Reality offers the possibility of recreating Spon Street as it would have appeared in the Middle Ages and allowing people today to experience what it may have been like to explore Spon Street in the late 15th century. The goal of this project is for visitors to Coventry's Herbert Art Gallery and Museum to be able to use VR to wander down modern Spon Street and then to be virtually transported back to the 15th century and visit one of the tanning

DOI: 10.2312/gch.20171297

Figure 1: *Medieval houses in Spon Street today. (a) Current businesses (b) Tudor house*

works that was there then. The project is focussing on recreating authentic multiple senses (visuals, audio, smell and temperature) to enable the visitors to experience what the past may have been like.

2. Related work

The desire to recreate cultural heritage on a computer to gain better insights of the past has come a long way since the early 3D models of Reilly from IBM [RS89]. Museums have always had a interest in using VR technology to present information to visitors, for example [BFS00, GCR01], but up until a few years ago the technology was expensive and limited what museums could afford to provide [CB10]. With the advent of inexpensive VR technology, such as the Oculus Rift, HTC Vive etc., museums are now able to showcase whole cultural heritage environments in VR for visitors to enjoy, including entire ancient cities, for example [FAG*08, WOFK13]. However, as with any recreation, care needs to be taken, otherwise there is a real chance of presenting misleading representations of the past [Mar01, MR94]. The real world is multisensory. Our perception of any environment is based on what we see, hear, smell and feel (touch, temperature, etc). In addition, the interaction of the senses, or cross-modalities, can significantly affect how anything is perceived, for example [HT66, ZS04]. To authentically recreate a site, the multiple senses need to be physically simulated and delivered to the user of the VR system in a natural manner [CHM09, MRBB16]. Furthermore, any simulation needs to take into account the materials that would have been used in the past [HMD*10].

3. The leather industry

The medieval leather industry is an under-studied area. Domestic in nature, it was never as wealthy as the cloth industry upon which England's export trade was based. But leather was used in a variety of key ways including clothing, shoes and boots, tents, saddles, gloves, purses and for carrying wine. As the industry grew in importance and the cloth industry in some towns and cities declined by the end of the 15th century, the status of those controlling the leather industries such as tanning, rose to such a degree that they often became part of the ruling urban oligarchy, including becoming mayor [Swa89].

By its nature, the leather trades worked closely with the butchers, who sold-on the hides to tanners complete with head and hoofs. Archaeological evidence for tanning and its associated industries often includes large deposits of cattle horn cores at particular locations, together with rectangular or circular pits. Cattle hides were associated with tanning, which was essentially a "wet" process that required access to running water. After washing the hides in running water, they were immersed in warm bird or dog droppings to loosen the hair and fat, which was then scraped off with a two-handled knife. The hides were treated with urine and stored in shallow vats filled with water. The treated hides were dried, treated with crushed oak bark, and then put in deep vats filled with a vegetable tanning material made of oak bark. This last stage in the process could take up to 12 months. The hide was subsequently dried and stored in a dark shed with louvered panels. At this stage in the process, the hide was known as "red" leather.

The red leather was then handed to the currier who worked the leather to make it more flexible and turn it into "black" leather. The finished product was then sealed up and sold to cordwainers, girdlers, cardmakers and saddlers. Other skins (calf, goat, cat, dog and horse) were treated by tawyers to create a "white" leather. Tawyering was a "dry" process that involved rubbing alum into the hide with salt to create a white colour. The hide was then stretched and softened with oil and egg yolks and put into large tubs. A further stretching process involved softening the leather with a blunt blade and then the leather would be dyed if needed [Swa89,Che01].

It is notoriously difficult to precisely match documentary evidence with existing standing buildings, but on the evidence we have, it would seem likely that surviving building no 14 Spon Street, was the site of a tanning workshop, known to have existed on a large plot on the south side of the street, near the gate, and with access to the River Sherborne. The current building dates from the 15th century, and was upgraded in the 16h century with gables, and is now known as 'Tudor House', Figure 1(b) [Jon58].

4. Creating the multisensory environment

A human continually experiences a range of sensory stimuli in the real world, including visuals, audio, smell, temperature, motion, taste, pain, etc. Our brains process these separate sensory inputs into a common perception of a scene [CSS04]. Exactly how the brain processes these sensory stimuli is still an active area of research. Fortunately, to simulate a multisensory experience and deliver this in a VR system, it is only necessary to deliver the "perceptually equivalent" level of sensory stimulation to the user for each sense [CHM09]. Perceptual equivalence is the level of accuracy needed in a simulation for the user to *believe* that what they are experiencing is the real world. It is not necessary to integrate these senses in any way during delivery, as the brain will do this naturally [CF08]. To recreate medieval Coventry we simulate four senses: visuals, audio, smell and temperature. As mentioned previously, the goal of this experience is to have the visitor to start in a virtual model of modern Spon Street and then be virtually transported back to the 15th century.

4.1. Capture and modelling

A total of 14 scans using a Focus3D X 330 Faro Scanner were taken of Spon Street from the ring road to John the Baptist church, which stands on the site of a former gate into Coventry. Each point cloud scan comprised a resolution $40,000 \times 20,000$ points. Neighbouring scans where controlled in such a fashion that reference targets (calibrated spheres) for aligning subsequent scans were visible

Figure 2: *The point cloud of modern Spon Street*

and close enough that they would receive enough point information upon their surface, > 60 points, that they could be used as control points for limiting degrees of freedom between scans and stitching the scans together. This process was achieved using correspondence matching. The stitching was accomplished using Scene 6.0 software, Figure 2.

The point cloud was homogenised to maintain point density across the scan. This was achieved by culling points using statistical inference of their likelihood to be noisy and their local point density R, calculated through Equation 1:

$$R = \frac{2\pi}{hd} \tag{1}$$

where R is point density in metres, h is the horizontal resolution of the point cloud scan and d is the distance of the recorded point from the scanner position. For example, given a scan resolution of $8,000 \times 4,000$ points, the point density at 10 metres from the scanner is 7.85 mm and 15.7 mm at 20 metres from the scanner. A point was deemed to be noisy if, by evaluating its local neighbourhood, it could be adjudged to be an outlier to the point set by some variable radius. This happened if, for example, the laser scanner hit a dust particle.

Once the point cloud was processed the information was a large set of points P: $\forall p \in P : \exists\{x,y,z,r,g,b\}$ where x,y,z is the 3D point location in Cartesian coordinates and r,g,b is the colour triplet of the point p.

In addition, High Dynamic Range (HDR) panoramas were taken using a Spheron PanoCam to ensure the full range of light in the scene was captured [BADC11]. These HDR panoramas were subsequently used to accurately relight both the model of modern and medieval Spon Street.

4.2. Re-enacting tanning

A major challenge when including virtual humans in any reconstruction is the "uncanny valley" effect [Mor12]. Despite the very latest human modelling and rendering techniques, it is still not possible to achieve virtual humans that look as realistic as real humans. It was thus decided to film historical re-enactors undertaking tanning in the medieval manner and using chroma key methods to composite the actors back into the virtual model of the tanning environment. To avoid users trying to interact with the virtual tanning process or view it from a direction that had not been filmed, they are only able to view it from a first floor window of the Tudor House.

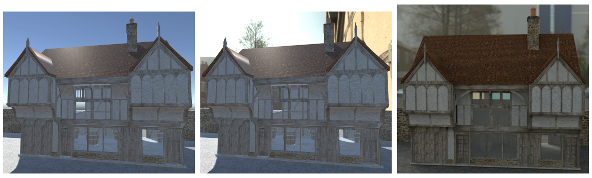

Figure 3: *Different quality rendering of the Tudor house (a) low: Unity "Fastest" (b) medium: Unity "Fantastic" (c) high: Path traced*

The re-enactment was filmed in Stereo using two GoPro Hero 4 cameras so that the result looked more like "looking out of a window" than "watching a video".

The process of incorporating the stereo footage in Unity for delivery to the HMD, is as follows:

1. The tanning activities are shot from two separated cameras (left and right) in the green screen environment, with a suitable focal length, filming angle and shooting distance.
2. The frames of two videos are synchronised and then synthesized into 3D side-by-side format.
3. Two same sized planes are placed in the exact position as the screens; one for the left eye (left screen), and another for the right eye (right screen).
4. Each frame of the side-by-side video is read, cut into two textures, left and right, and then these textures are send to the left screen and right screen respectively.
5. Finally a Shader makes the green part of the texture on each screen transparent, allowing the virtual environment to be seen.

4.3. Rendering

The quality of the rendering of any virtual environment can play a key role in its perceived realism. When rendering there is, of course, a trade-off between the amount of time required to render any image and the quality required. Rendering a high quality, physically based image using, for example path tracing [Kaj86], can take many minutes or even hours for a highly complicated scene, while non-physically based, rasterized rendering can achieve several hundreds of frames per second.

In this project, three quality levels were considered: a low quality rendering using the standard Unity renderer, a medium quality using the best quality settings for rendering that Unity provides, and a high quality image rendered using a physically based path tracer with 2,048 samples per pixel Figure3. For the two higher-quality rendering, a HDR environment map captured at Spon Street was used. The low quality level used Unity's standard ambient lighting model. The two Unity renderings (low and medium) used the quality settings "Fastest" and "Fantastic". Details of the actual parameter settings for each of these can be found at [uni].

4.4. 3D audio

Hearing is a vital human sense, enabling us to locate and identify objects spatially and react to events even if we don't see them. For an authentic experience in a virtual environment, the sound needs to be simulated and delivered in 3D [HHD*12]. To recreate a medieval tanning environment, the sounds that are required include: the workers shouting to each other, the sound of the work in progress, and other ambient sounds, such as water sloshing, dogs barking and pigs rummaging in the yard.

A ray-tracing approach was adopted for the acoustic rendering [STKS]. We rendered a B-format (Ambisonics) Room Impulse Response (RIR) at 352800 Hz sampling frequency as the highest temporal resolution for sound. Such high sampling rates are used in Digital extreme Definition (DxD) applications for editing high-resolution audio tracks. Every RIR is convolved with a two-channel head related impulse response (HRIR) signal using the CIPIC HRTF database [ADT01]. The HRIRs capture all source localisation cues and they encode elevation and azimuth effects of the incoming sound signal to a listener's ears. These cues are affected by human anatomical characteristics (head size, shoulder, pinna shape, etc.) and the HRIR content is used to capture this interaction.

4.5. Smells and Temperature

Real world smells can be captured by placing an Automated Thermal Desorption tube close to the smell source - and sucking air across the tube. There is a fine granular material in the tube which traps the smell molecules. The tube is then placed in a gas-liquid chromatography device which separates the mixture of odourants into consistent molecules which are then passed through a mass-spectrometer which records a histogram of which molecules are present.

The dominant smell in the tanning process is that of urine. Urine consists of more than 95% water in which is dissolved a number of other materials, including urea, chloride and potassium. The pungency of the smell from urine is from the urea, $CO(NH_2)_2$, which is a non-toxic form of ammonia.

The smells of tanning were captured using an Automated Thermal Desorption placed 12cm [ISO] from the source. The tube was then given to a perfumer who analysed the smell and provided a reconstructed smell which is perceptually close to the real smell. This was validated by the perfumer by comparing samples of the real smell (captured in bags at the site of the tanning) and the virtual smell. The intensity of the smell that needed to be delivered to the user was calculated based on the distance from the window from where the user will be viewing the tanning and assuming little or no wind. If desired, wind can be simulated through fans, and the intensity of the smell determined by Computational Fluid Dynamics (CFD). As a human's perception of smell is not very precise, only a simplified version of CFD is required to compute the level of smell intensity at the user [RBC10]. The virtual smell was delivered to the user in the virtual environment in small controlled amounts, at the appropriate intensity, via Polytetrafluoroethylene (PTFE) tubes situated under the nose [HDBRC11].

In addition to wind, the intensity of the smell may also be affected by temperature. Currently there are three temperature settings in our multisensory reconstruction: "cool", "ambient" and "hot". The cool and hot settings are provided by lightly blowing cool or hot air at the user via a standard fan heater. As an approximation of the relationship between smell intensity and temperature, if the cool setting is chosen, the smell intensity determined by the CFD is halved, while it is doubled for the hot setting. Future work will investigate this relationship in detail and provide a more accurate "modification factor" if necessary.

5. Conclusions and Future Work

While the manufacture of cloth and caps made Coventry wealthy in the 15th century, leather-making, of which tanning was an essential process, was also an important industry in the city. Medieval tanning was a noisy, smelly process. For modern visitors to a museum to fully appreciate what the past may have been like, any virtual reconstruction of tanning needs to be multisensory. Furthermore, to avoid misleading visitors, or simply delivering "pretty pictures", the multiple sensory stimuli in the environment need to be accurately recreated and their distribution within the environment simulated by means of physics. For visuals this is path tracing, audio - beam tracing, and smell - computational fluid dynamics.

One current problem is the time to compute the path traced images. The quality of the path traced images, Figure 3(c) is significantly higher than even the "Fantastic" Unity rendering, Figure 3(b), and thus provides an enhanced user experience. However, each path traced image currently takes approximately 2,438 seconds to compute on a PC (CPU i7-3770K @ 3.50Ghz, RAM 16GB, GPU NVidia 750Ti). On the same computer, a "Fantastic" quality Unity image takes 9.8ms (102fps), and the "Fastest" setting can compute images at 6.9ms (145fps). This substantial time per frame for path tracing precludes its use in any interactive environment. For a predetermined "fly through" of medieval Spon Street, each frame can be precomputed, however if interaction by the user is desired in the environment, then the visuals will need to be computed "on the fly". Current work is investigating optimisation algorithms for path tracing that will enable the user to have a lower quality image as they interact in the environment, perhaps with fewer samples and only with direct lighting, and this quality will increase automatically whenever they keep still.

Another current limitation is the need to capture the HDR environment maps on the pavement, rather than in the middle of a busy street. This can lead to "scaling and alignment issues" in the background of the images rendered with the environment map, Figure 3(b) and (c). Future work will investigate automatically scaling and aligning any background information in an environment map (eg. trees, building etc) to better match the positioning of the virtual geometry.

Future work will also consider reinforcing the learning that visitors have been able to achieve through the multisensory virtual experience, by directly linking the virtual experience with the archaeological finds displayed in traditional cases in the museum, as well as more tangible exhibits. This will enable the visitors to experience through directly being able to handle samples from all stages, how the leather has progressed from a raw hide to "black leather" and finally the completed leather products.

Acknowledgements

We would like to thank the China Scholarship Council for supporting Yuanfa Dong's visit to the UK.

References

[ADT01] ALGAZI V., DUDA R., THOMPSON D.: The Cipic hrtf database. In *IEEE workshop on Applications of Signal Processing to Audio and Electroacoustics* (2001), pp. 99–102. 4

[BADC11] BANTERLE F., ARTUSI A., DEBATTISTA K., CHALMERS A.: *Advanced High Dynamic Range imaging*. CRC Press, 2011. 3

[BFS00] BARCELO J. A., FORTE M., SANDERS D. H. (Eds.): *Virtual Reality in Archaeology*. ArchaeoPress, 2000. 2

[CB10] CARROZZINO M., BERGAMASCO M.: Beyond virtual museums: Experiencing immersive virtual reality in real museums. *Journal of Cultural Heritage 11* (2010), 452–458. 2

[CF08] CHALMERS A., FERKO A.: Levels of Realism: from virtual reality to real virtuality. In *SCCG08. Spring Conference in Computer Graphics* (Slovakia, April 2008), ACM Press, pp. 27–33. 2

[Che01] CHERRY J.: Leather. In *English Medieval Industries* (2001), Hambledon, London, pp. 295–318. 2

[CHM09] CHALMERS A., HOWARD D., MOIR C.: Real Virtuality: A step change from Virtual Reality. In *SCCG09. Spring Conference in Computer Graphics* (Slovakia, April 2009), ACM Press. 2

[CSS04] CALVERT G., SPENCE C., STEIN B.: *The multisensory handbook*. MIT Press, 2004. 1, 2

[DH08] DORMER HARRIS M.: *The Coventry Leet Book or Mayors Register Parts I & II*. Early English Text Society, 1908. 1

[FAG*08] FRISCHER B., ABERNATHY D., GUIDI G., MYERS J., THIBODEAU C., SALVEMINI A., MÜLLER P., HOFSTEE P., MINOR B.: Rome reborn. *ACM SIGGRAPH new tech demos* (2008). 2

[GCR01] GAITATZES A., CHRISTOPOULOS D., ROUSSOU M.: Reviving the Past: Cultural Heritage Meets Virtual Reality. In *Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage* (New York, NY, USA, 2001), VAST '01, ACM, pp. 103–110. 2

[HDBRC11] HARVEY C., DEBATTISTA K., BASHFORD-ROGERS T., CHALMERS A.: Visual saliency for smell impulses and application to selective rendering. In *EG UK* (2011), Eurographics. 4

[HHD*12] HULUSIC̀ V., HARVEY C., DEBATTISTA K., TSINGOS N., WALKER S., HOWARD D., CHALMERS A.: Acoustic Rendering and Auditory-Visual Cross-Modal Perception and Interaction. *Computer Graphics Forum 31*, 1 (February 2012), 102–131. 4

[HMD*10] HAPPA J., MUDGE M., DEBATTISTA K., ARTUSI A., GONÇALVES A., CHALMERS A.: Illuminating the past: State of the art. *Virtual Real. 14*, 3 (September 2010), 155–182. 2

[HT66] HOWARD I., TEMPLETON W.: *Human spatial orientation*. John Wiley and Sons Ltd., 1966. 2

[ISO] ISO STANDARD: *Workplace air quality: sampling and analysis of volatile organic compounds by solvent desorption/gas chromatography*. BS ISO16200-1. 4

[Jon58] JONES S.: *Survey of Coventry's timber framed buildings for Victoria County History*. unpublished, 1958. 2

[Kaj86] KAJIA J.: The rendering equation. *ACM SIGRAPH Computer Graphics 20*, 4 (1986), 143–150. 3

[Mar01] MARTINEZ P.: Digital realities and archaeology: a difficult relationship or a fruitful marriage? *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage* (2001), 9–16. 2

[Mor12] MORI M.: The uncanny valley. *(Translated by K.F. MacDormanand and N. Kageki), IEEE Robotics and Automation 19* (2012), 98–100. 3

[MR94] MILLER P., RICHARDS J.: The good, the bad and the downright misleading: archaeological adoption of computer visualization. *Computer Applications and Quantitative Methods in Archaeology* (1994), 19–22. 2

[MRBB16] MELO M., ROCHA T., BARBOSA L., BESSA M.: The impact of body position on the usability of multisensory virtual environments: Case study of a virtual bicycle. In *Proceedings of the 7th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion* (New York, NY, USA, 2016), DSAI 2016, ACM, pp. 20–24. 2

[RBC10] RAMIC-BRKIC B., CHALMERS A.: Virtual smell: Authentic smell diffusion in virtual environments. In *Proceedings of the 7th International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa* (New York, NY, USA, 2010), AFRIGRAPH '10, ACM, pp. 45–52. 4

[RS89] REILLY P., SHENNAN S.: *Applying Solid Modelling and Animated Three-dimensional Graphics to Archaeological Problems*. IBM UK Scientific Centre, 1989. 2

[STKS] SILTANEN S., TAPIO L., KIMINKI A., SAVIOJA L.: The room acoustic rendering equation. *Acoustical Society of America 122*, 3, 1624–1632), year = 2007. 4

[Swa89] SWANSON H.: *Medieval Artizans: an urban class in late medieval England*. Blackwell, Oxford, 1989. 2

[uni] https://docs.unity3d.com/manual/class-qualitysettings.html. 3

[WOFK13] WEBEL S., OLBRICH M., FRANKE T., KEIL J.: Immersive experience of current and ancient reconstructed cultural attractions. In *Digital Heritage International Congress* (October 2013), IEEE. 2

[ZS04] ZAMPINI M., SPENCE C.: The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of Sensory Studies 15* (2004), 347–363. 2