PROCEEDINGS of the 2nd Symposium: The Acoustics of Ancient Theatres
6-8 July 2022 Verona, Italy

Analysis Of Auralisation Techniques For 3d Models Of Music Theatres

Sophie Schauer[1]; Natalya Shakhovska[2]; Jürgen Sieck[3]

[1] HTW Berlin - University of Applied Sciences, Germany, sophie.schauer@htw-berlin.de
[2] LPNU Lviv - Lviv Polytechnic National University, Ukraine, nataliya.b.shakhovska@lpnu.ua
[3] Htw Berlin - University Of Applied Sciences, Germany, Juergen.Sieck@ htw-berlin.de

ABSTRACT
The work displayed in this paper was achieved during the AURA project, a Creative-Europe-funded project with mul- tiple international partners both with technical and cultural expertise. The paper describes current auralisation techniques and recites the past course of the project. It draws comparisons to similar projects before developing an own concept for an auralised application. This will include one application using a hand-modelled 3D object of the Konzerthaus Berlin, which opened in 1821, and one which will use a laser-scanned model. These applications, using the different types of models, will be further analysed regarding audio and video quality as well as overall performance. Furthermore, the de- velopment of a neural network for instrument positioning relative to the listener will be described. A conclusion on the suitability of the two model types for auralised applications will be drawn and a further outlook on the ongoing work in the AURA project will be given.
Keywords: Auralisation, Virtual Reality, 3D modelling

1. INTRODUCTION
State-of-the-art technology usage in the cultural sector has risen immensely in the past years. Especially since the Covid-19 pandemic, more and more cultural institutions had to switch to an online concept [1], making modern technology an integral part of how we experience art, music and culture in general. With digital tour guides omnipresent in almost all museums and exhibitions and musical plays and concerts being live- streamed, technology is indispensable in our cultural lives nowadays.

The rise of technology usage in music venues has also increased the popularity of auralisation techniques. By using auralisation a realistically sounding 3D model ofany musical theatre around the world can be virtually visited and experienced. The use cases for auralisation are widespread and can range from experiencing the soundscape of ancient heritage sites or hearing thesame piece of music in multiple different musical venues without the need for travelling.

The AURA project, a Creative-Europe-funded project with multiple international partners from Germany, Italy and Ukraine, focuses on exactly this kind of technology usage in music venues.

The paper will give detailed insights into state-of- the-art technology regarding auralisation and present a concept for an auralised application using two different models from the Konzerthaus Berlin, one of them being hand-modelled, the other one produced using the laser-scanning technique. These will be the foundation for an analysis of both model types. The audio, video quality and overall performance will be analysed, before then giving insight into an analysis of listener and instrument location using a neural network. A conclusion will be drawn on which 3D modelling is more suitable for auralisation and how instrument placementand listener location can influence the visitor's experience.

Figure 1 – Music venues from the AURA projectin Germany, Italy and Ukraine

2. STATE OF THE ART
Auralisation is the virtual reconstruction of sound fields [2]. Taking into consideration the material parameters of all objects located in the room, how theobjects react with sound, as well as the room geometryin general, a realistically sounding audio experiencecan be created.

Over the course of the AURA project, three music venues have been auralised. First, a prototype wasmade using a hand-modelled 3D object from the GreatHall of the Konzerthaus Berlin. After laser scans and the 3D modelling process of the Teatro del Maggio in Florence were finished, the completed 3D object was auralised as well, using the same technique as in the previously developed prototype. The next step was thelaser scan of the Konzerthaus Berlin, which resulted in a second 3D model from the Great Hall. The outcome was two

10.58874/SAAT.2022.164

comparable auralised models of the GreatHall in the Konzerthaus Berlin which will be analysedin the 4th chapter.

The auralisation was done using the Steam Audio plugin for Unity [3]. Steam Audio adds functionality to create spatial audio in applications developed in the Unity editor. Features like material parameters and geometry specifications can be added to all object components. A fundamental part of spatialisation is the use of head-related transfer functions, short HRTFs. HRTF describes how sound from a specific point within the environment will travel to the listener's ear [4]. Steam Audio gives the developer the option to substitute the default HRTFs with their own SOFA file and thus the ability to create sound experiences specific- ally customised for each user.

During the APOLLO project[1], carried out by the INKA research group, a digital 3D model of Konzerthaus Berlin (KHB) and all its halls had been created by hand. The model was used for several projects including a guided tour in virtual reality. This model was reused for the first auralised prototype for the AURA project.

Figure 2 – Small Hall from the Konzerthaus Berlin, de-veloped during the APOLLO project

The laser scan of the Konzerthaus Berlin was produced by the DiDa lab, the department of architecturefrom the University of Florence, one partner of the AURA project. Many previous digital surveys have been conducted. One of them was the analysis of the St. George Church in Girne, Cyprus. During a digital survey campaign, the volumes and architectural features of the church were studied and morphological data, drawn up through the use of laser scanners, was integrated with the material information acquired through Structure from Motion (SfM) methodologies. This resulted in a complete textured 3D model [5]. The same approach was taken for the digitalisation of the Konzerthaus Berlin and the production of the 3D model of the Great Hall.

3. APPLICATION CONCEPT AND IMPLEMENTATION FOR ANALYSIS

Both KHB models were imported into the Unity Editor. The concept was based on the previously developed auralised models using both Unity and the plugin Steam Audio [3]. The models vary in the number of polygons and therefore also in size. The material parameters for the auralisation were the same. The following table gives an overview of the models' geometry parameters.

Figure 3 – Laser scan data during the digital survey of the Great Hall in the Konzerthaus Berlin

Table 1 – parameters of both models

	Hand-modelled	Laser-scanned
File size	37,6 MB	191,8 MB
Objects	63	1 116
Vertices	1 470 066	2 356 068
Edges	2 439 618	3 525 955
Faces	1 171 600	1 549 111
Triangles	1 171 600	1 551 095

Each model was then placed in the prepared scenes. To further analyse the two versions of the Great Hall the target platform was changed from virtual reality glasses to Windows/macOS. Therefore, the interaction system was changed as well. Position switches can bedone by pressing the numbers 1 to 8 on the keyboardof the used laptop or PC. Musicians can be enabled and disabled by pressing the keys F1 to F10. This helped in making the testing simpler and more efficient.

4. ANALYSIS

After preparing the scene setup for both models ofthe KHB Great Hall, test runs were done analysing a handful of graphic parameters. All tests were done in the Unity Editor on a MacBook Pro 2016 with a 2,6 GHz quad-core Intel Core i7 CPU. Recording over the course of 1 minute from the same position in both models the FPS, meaning the number of frames Unitydraws per second, was captured. Even though the laser-scanned model has a more complex geometry, the FPS were between 104 to 114, averaging at around 108 FPS. The number of FPS for the hand-modelled version was more consistent however a lot lower in com- parison, being between 54 and 58.

Next, the CPU usage was measured. This took into consideration the total amount of time taken to process one frame as well as the time taken to render one frame. Again for 1 minute, both scenes were measured.For the scene of the laser-scanned model, the CPU took between 8.1 and 9.6 milliseconds to process one frame and between 1.6 and 2.2 milliseconds to render.In comparison, the hand-modelled Great Hall model was averaging between

[1] https://inka.htw-berlin.de/project/apollo/

17 and 19 milliseconds and 0.7to 1.1 milliseconds to render. The number of batches drawn is for the laser-scanned model at 544 a lot high-er than 166 for the hand-modelled one. This explains the difference in time.

The optical analysis revealed big differences in di-mensions, lighting and textures. The laser-scanned model is accurate to the centimetre whereas the hand-modelled one was done using approximate measurements. This led to very different sizes of the stage, and overall dimensions of the hall. The laser-scanned mod el seems a lot more spacious and realistic. The lighting and textures vary as well. Even with the same lighting settings, the laser-scanned model is darker and has more muted tones. The hand-modelled version is brighter and has more white tones.

Figure 4 – Optical comparison from same position (left hand-modelled, right laser-scanned)

The auralisation is based on the model's geometry. As already mentioned, the laser-scanned model is ac- curate to the centimetre resulting in a more realistic auralisation. However, these differences are very minor and did not make for a noticeable audible dis- tinction between the two models. The overall performance of both scenes on the testing device was good. The higher number of FPS made camera movement in the laser-scanned model significantly smoother. Even with far fewer vertices and objects in general, the hand-modelled version was slower and had more anti-aliasing. This can solely be attributed to the visualisation and not the auralisation part since all settings and parameters remained the same for both models.

The analysis confirmed the suitability of accurate laser scan models for the AURA use case. The graphical rendering is more performant, making the visualisation overall superior. Especially for virtual reality applications the higher number of FPS makes for a better experience and can possibly prevent motion sickness. Generally speaking, frame rates below 90 FPS can lead to disorientation and nausea for the user[6]. Even though there was no distinct difference in sound between both models, for larger scaled auralisations the difference in audio can be more significant.

5. NEURAL NETWORK DEVELOPMENT

There are significant differences in the placement of musical instruments on stage or in the orchestra pit. The selection, arrangement and choosing of the number of musical instruments are related to the achievement of creative tasks. However, from an acousticpoint of view, the number of musical instruments is associated with sound balance problems. Some musi- cal instruments differ significantly in their dynamic and frequency range, direction characteristics, and timbre. Such large differences can undoubtedly lead to masking one group of instruments to another. More- over, the concert hall brings significant changes to the overall sound picture. Only a proportional selection ofthe number of instruments and the correct location on the stage can give the desired result.

Therefore, it is advisable to analyse the listener's location depending on the instrument he wants to hear. In addition, good visibility of the performance andsound quality are also important factors.

5.1 Materials and methods

To determine the best location, multi-criteria para-meters were taken into account, including [7]:

- The convenience of viewing the play (two number characteristics – upper and lower border),
- The viewing angle,
- The sound level of the group of instruments (the number of parameters is determined by the number of instruments and their location - in the orchestra pit or on stage),
- Physical parameters of the room,
- Frequency ranges of instruments,
- Physical parameters of the spectator placement quadrant.

The task of classification is set. Class labels are the quadrant numbers of the listener's placement.

A fairly simple neural network architecture was experimentally selected, which was a compromise be-tween the accuracy of the classification and its speed. The architecture of the network will change for each concert hall and the number of musicians involved in the performance. 98 traits (input parameters) were se- lected for the pre-workout.

A fully connected neural network to complete the task is chosen. Four fully connected layers with the number of neurones 256, 128, 64, 10, respectively, were used, followed by batch normalisation [8] and the dropout technique with probabilities of 20%, 20%, and 50%, to avoid retraining. The architecture of the neural network is presented in Fig. 5.

5.2 Implementation

The dataset was collected based on Lviv Opera House [9]. The Artificial intelligence module (AI module) was developed in Unity using C# and has the following file structure:

- Constants - a file that contains constants that willnot change during the program.
- Features_extracting - a file containing functions for direct work with audio files. The get_features

function picks up audio in digital format and returns a list of selected features.

• Model - a file in which data is pre-processed, normalised, created and trained neural network. Contains only one main function, which takes the input path to the dataset and then normalises it, creating and training the neural networks based on the specified dataset. The result of the function is a trained neural network, which is stored on a disk, and then used in the main part of the program.

• Main - implements the basic algorithm of the program. Contains a predict_place function that classifies the best user's location (quadrant).

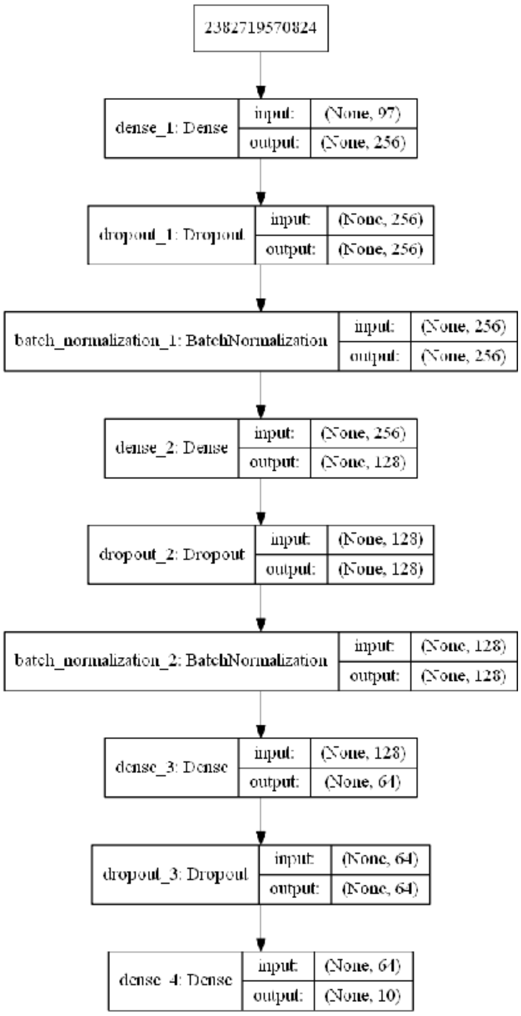

Figure 5 – The architecture of the artificial neural network

The developed neural network was trained on the dataset of the Lviv Opera House but will be applied to the Konzerthaus Berlin and Teatro del Maggio Florence during the further course of the project.

6. CONCLUSIONS

The analysis gave a clearer view of the differences between both model types. It confirmed the choice to use the laser-scanned version for further development and testing. It gives the most realistic result and is best suited for virtual reality applications.

The developed neural network will be used for the conduction of a case study, which will be the main focus for the future months of the AURA project. Business models will be created using the instrument placement analysis to give visitors the best seat location based on their preferences.

The finished auralised models will be used for testruns with experts and laymen in the next steps. Case studies for educational purposes are planned and will take up most of the remaining running time of the project.

ACKNOWLEDGEMENTS

This paper portrays the work carried out in the context of the AURA project that is generously co-funded by the Creative Europe Programme of the European Union (grant number: 101008547).

7. REFERENCES

1. IDEA Consult, Goethe-Institut, Amann S. and Heinsius J. 2021, Research for CULT Committee – Cultural and creative sectors in post-Covid-19 Europe: crisis effects and policy recommendations, European Parliament, Policy Department for Structural and Cohesion Policies, Brussels
2. TRONCHIN, Lamberto, et al. Validation and application of three-dimensional auralisation during concert hall renovation. *Building Acoustics*, 2020, 27. Jg., Nr. 4, S. 311-331.
3. Schauer, S., Bertocci, S., Cioli, F., Sieck, J., Shakhovska, N. & Vovk, O. (2022). Auralization of Con- cert Halls for Touristic Purposes. i-com, 21(1), 95-107. https://doi.org/10.1515/icom-2022-0008
4. POURU, Lasse. The Parameters of Realistic Spatial Audio: An Experiment with Directivity and Immersion. 2019.
5. VOLZONE, R.; CIOLI, FEDERICO; BIGONGIARI, M. The church of St. George in the Kyrenia castle in North Cyprus: Bringing out the shape of architecture. The church of St. George in the Kyrenia castle in North Cyprus: Bringing out the shape of architecture, 2018.
6. Kersten, T., Drenkhan, D., & Deggim, S. (2021). Virtual Reality Application of the Fortress Al Zubarah in Qatar Including Performance Analysis of Real-Time Visualisation. KN-Journal of Cartography and Geo- graphic Information, 71(4), 241-251.
7. Hanifa, R. M., Isa, K., Mohamad, S., Shah, S. M., Nathan, S. S., Ramle, R., & Berahim, M. (2019). Voiced and unvoiced separation in malay speech using zero crossing rate and energy. Indonesian Journal of Electrical Engineering and Computer Science, 16(2), 775-780.
8. Cao, B., Zhao, J., Lv, Z., Gu, Y., Yang, P., & Hal- gamuge, S. K. (2020). Multiobjective evolution of fuzzy rough neural network via distributed parallelismfor stock prediction. IEEE Transactions on Fuzzy Systems, 28(5), 939-952.
9. Войтович О. Критерії оцінювання звучання оркестру в концертних залах. Українська музика: Науковий часопис ЛНМА ім. М. В. Лисенка. Львів,2018. Ч. 3(29). С. 94–98.