

## **Περιβάλλον ακουστικής προσομοίωσης χώρων για την αναζήτηση τοποθεσίας κινηματογράφησης**

Νικόλαος Βρύζας<sup>1,\*</sup>, Λάζαρος Βρύσης<sup>1</sup>, Μαρίνα Σταματιάδου<sup>1</sup>, Χαράλαμπος Δημούλας<sup>1</sup>, Γεώργιος Καλλίρης<sup>1</sup>, Ιωάννης Μουρτζόπουλος<sup>2</sup>

<sup>1</sup>Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

<sup>2</sup> Πανεπιστήμιο Πατρών, Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών

\*nvryzas@auth.gr

### **ΠΕΡΙΛΗΨΗ**

*Η επιλογή χώρων αποτελεί ένα αναπόσπαστο κομμάτι της διαδικασίας κινηματογράφησης προπαραγωγής. Το πρόγραμμα SCENE στοχεύει στην υλοποίηση ενός ολοκληρωμένου συστήματος που καλύπτει τις ανάγκες αυτής της διερεύνησης. Στην παρούσα εργασία παρουσιάζεται το εργαλείο Ακουστικής προσομοίωσης χώρων. Πρόκειται για ένα περιβάλλον όπου χρήστες μπορούν να παρέχουν καταγραφές χώρων που σχετίζονται με τοποθεσίες σχετικές με την πολιτιστική κληρονομιά, ακολουθώντας ένα πρότυπο οδηγιών. Οι καταγραφές χρησιμοποιούνται για την προσομοίωση της ακουστικής απόκρισης του χώρου σε διάφορες θέσεις πηγής-ακροατή, το οποίο παρέχεται στους χρήστες μέσα από μια γραφική διεπαφή που παρέχει αναπαραγωγή σε πραγματικό χρόνο.*

### ***A room acoustic simulation tool for filming location scouting***

### **ABSTRACT**

*Location scouting is an integral part of the pre-production filming process. The SCENE program aims to implement an integrated system that meets the needs of this investigation. In the current paper, the room acoustic simulation tool is presented. It is an environment where users can provide space records related to heritage sites, following a set of guidelines. The recordings are used to simulate the acoustic response of the room at various source-listener positions, which is provided to users through a graphical interface that provides real-time rendering and playback.*

## 1. Εισαγωγή

Το έργο SCENE (Searchable multi-dimensional Data Lakes supporting Cognitive Film Production & Distribution for the Promotion of the European Cultural HeritagE) στοχεύει στην υλοποίηση μιας πλατφόρμας βασισμένης σε αρχιτεκτονική λιμνών δεδομένων για την αποτελεσματική συνεργασία και ανταλλαγή γνώσης των ενδιαφερομένων μερών της κινηματογραφικής παραγωγής. Σε αυτό το πλαίσιο, θα είναι προσβάσιμα για εξερεύνηση τρισδιάστατα μοντέλα τοποθεσιών που σχετίζονται με την ευρωπαϊκή πολιτιστική κληρονομιά, παρέχοντας επίσης μια προσομοίωση αλλαγής φωτισμού και ένα εργαλείο προσομοίωσης ακουστικής απόκρισης χώρων. Ο βασικός στόχος του εργαλείου προσομοίωσης ήχου είναι ότι να παρέχει σε παραγωγούς και οι σχεδιαστές ήχου πρόσβαση σε ένα εικονικό περιβάλλον όπου μπορούν να ακούσουν μία εκτίμηση του ηχητικού αποτελέσματος σε συγκεκριμένο χώρο χωρίς τη φυσική μετάβαση εκεί. Αυτή η λειτουργικότητα, σε συνδυασμό με τα υπόλοιπα εργαλεία που αναπτύσσονται στο έργο, διευκολύνει τη διαδικασία εντοπισμού και επιλογής τοποθεσίας κινηματογράφησης, προσφέροντας ταυτόχρονα μια πιο οικονομική και οικολογική εναλλακτική από την επιτόπια έρευνα.

Ιδιαίτερη συμβολή στην εξέλιξη του πεδίου της ακουστικής προσομοίωσης χώρων, αλλά και στην αξιοποίηση σχετικών προσομοιώσεων για την παροχή εμπειρίας χρήσης κατέχουν τα διαθέσιμα σύνολα δεδομένων. Η βάση ReverbDB [1] περιλαμβάνει ένα σύνολο πραγματικών Κρουστικών Αποκρίσεων Χώρου (KAX), που χρησιμοποιούνται για την εκπαίδευση εύρωστων μοντέλων αυτόματης αναγνώρισης ομιλίας. Η βάση dEchorate [2] περιλαμβάνει πολυκαναλικές καταγραφές KAX σε ένα σύνολο δωματίων. Το σύνολο δεδομένων OpenAir [3] περιλαμβάνει αποκρίσεις χώρων που σχετίζονται με μνημεία πολιτιστικής κληρονομιάς και υποστηρίζει διαφορετικές τεχνικές καταγραφής, καθώς και οπτικοακουστικό υλικό. Η συμπερίληψη γεωμετρικής πληροφορίας του χώρου μπορεί να αποτελέσει μια επιπλέον σημαντική πληροφορία για την ακουστική προσομοίωση χώρων [4]. Μία επιπλέον επικρατής τεχνική για την καταγραφή της ακουστικής απόκρισης των χώρων αποτελούν οι αποκρίσεις με χρήση μικροφώνων ηχητικού πεδίου και συνιστωσών του προτύπου B-Format Ambisonics [3], [5], [6].

Για την αξιοποίηση σχετικών τεχνικών σε εφαρμογές που απευθύνονται σε μη εξειδικευμένους χρήστες, κομβική είναι η ύπαρξη εργαλείων και υπηρεσιών. Το Pyroomacoustics [7] είναι μια βιβλιοθήκη σε γλώσσα python για την προσομοίωση χώρων, όπως και το GruRIR [8], που παρέχει επιπλέον δυνατότητες παραλληλοποίησης των υπολογισμών μέσω κάρτας γραφικών για αποδοτικότερη αξιοποίηση υπολογιστικών πόρων. Δυνατότητες ενσωμάτωσης πληροφορίας ακουστικής χώρων παρέχουν επίσης τα σύγχρονα περιβάλλοντα δημιουργίας βιντεοπαιχνιδιών, με περισσότερες ή λιγότερες δυνατότητες παραμετροποίησης [9]. Η χρήση προγραμμάτων φυλλομετρητών είναι μία πολύ βολική επιλογή για την παροχή περιβαλλόντων ακουστικής προσομοίωσης σε τελικούς χρήστες μέσω διαδικτύου, χωρίς να απαιτεί πρότερη γνώση ή υποδομή από αυτούς [10].

## 2. Το πρόγραμμα SCENE

Όπως αναφέρθηκε, στόχος του ερευνητικού προγράμματος είναι η ενσωμάτωση νέων τεχνολογιών για την ενίσχυση των διαδικασιών κινηματογραφικής προπαραγωγής, παραγωγής και μεταπαραγωγής. Στο επίπεδο της προπαραγωγής, οι βασικές εργασίες που υποστηρίζονται είναι η αναζήτηση τοποθεσίας βάσει κριτηρίων και η τρισδιάστατη μοντελοποίηση χώρων.

### 2.1 Αναζήτηση τοποθεσίας

Κατά τη διαδικασία αναζήτησης τοποθεσίας υπάρχουν διάφορα δεδομένα κριτήρια που πρέπει να πληρούνται. Αυτά αφορούν γεωγραφικά και μορφολογικά χαρακτηριστικά (π.χ. πόλη, θάλασσα, δάσος) που αναφέρονται στις σκηνές που πρέπει να γυριστούν. Άλλα κριτήρια μπορεί να είναι η προσβασιμότητα, το κόστος, η εγγύτητα σε άλλες επιθυμητές τοποθεσίες κλπ. Η πλατφόρμα του SCENE σκοπεύει να προσφέρει δυνατότητες ευφυούς αναζήτησης με χρήση φυσικής γλώσσας μέσω μεγάλων γλωσσικών μοντέλων για την αποτελεσματικότερη ανακάλυψη υποψήφιων τοποθεσιών.

### 2.1 Μοντελοποίηση χώρων

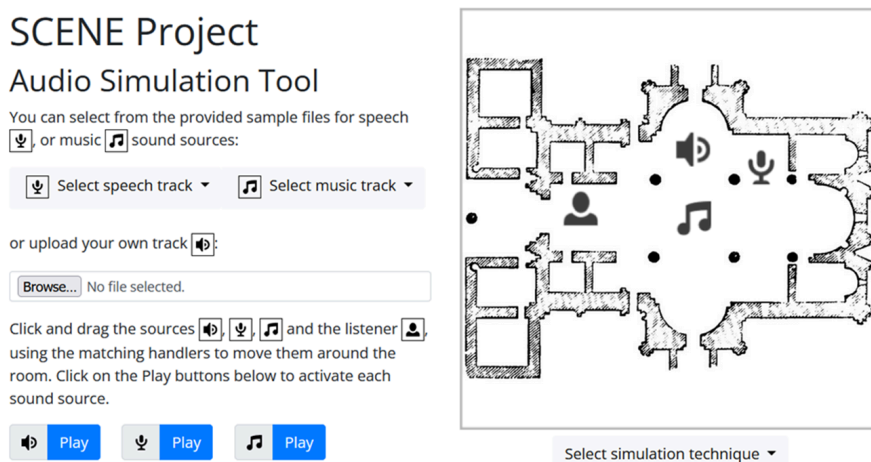
Η τυπική διαδικασία για την τελική επιλογή της τοποθεσίας κινηματογράφησης, όπως προκύπτει από την ανάλυση απαιτήσεων χρηστών που προηγήθηκε περιλαμβάνει μετάβαση στην τοποθεσία για αυτοψία και λήψη φωτογραφικού υλικού. Το SCENE αποσκοπεί στη δημιουργία μια βάσης (λίμνης) δεδομένων όπου να υπάρχουν τρισδιάστατες μοντελοποιήσεις χώρων με δυνατότητα πλοήγησης και παραμετροποίησης. Αυτό αρχικά αφορά τη διάθεση τρισδιάστατων οπτικών μοντέλων, όπου οι παραγωγοί μπορούν να έχουν μια πλήρη εποπτεία του χώρου και τον σχεδιασμό των σκηνών, καθώς και να προσαρμόσουν τις συνθήκες φωτισμού, χωρίς την απαίτηση για εξειδικευμένο εξοπλισμό, λογισμικό ή εκπαίδευση. Παράλληλα παρέχονται μοντέλα ακουστικής προσομοίωσης για την εκτίμηση της ακουστικής συμπεριφοράς του χώρου και την αξιολόγηση της καταλληλότητάς του από αυτή τη σκοπιά.

## 3. Περιβάλλον ακουστικής προσομοίωσης χώρων

### 3.1 Γραφική διεπαφή και είδη χρηστών

Το πρώτο είδος χρήστη αφορά τη λειτουργικότητα της πλοήγησης στο υλικό της βάσης δεδομένων και της διάδρασης με αυτό [11]. Έχει σχεδιαστεί ένα γραφικό περιβάλλον το οποίο βασίζεται στη βιβλιοθήκη JavaScript Google Resonance Audio [12]. Πρόκειται για μια βιβλιοθήκη που επιτρέπει την αναπαραγωγή αμφιωτικού ήχου σε πραγματικό χρόνο με βάση συγκεκριμένη παραμετροποίηση. Πιο ειδικά, ο χρήστης μπορεί να επιλέξει το είδος και τον αριθμό των ηχητικών πηγών (μουσική, ομιλία κλπ). Το περιβάλλον παρέχει κάποια προεπιλεγμένα αρχεία ήχου για δοκιμές,

αλλά δίνεται στον χρήστη η δυνατότητα να ανεβάσει τα δικά του αρχεία. Αυτό είναι βολικό ιδίως για την ακρόαση της προσομοίωσης της επίδρασης του χώρου σε πραγματικές ηχογραφήσεις ηθοποιών που έχουν πραγματοποιηθεί σε στούντιο. Παράλληλα ο χρήστης μπορεί να επιλέξει τη θέση των πηγών και του ακροατή στον χώρο. Για τη σχετική παραμετροποίηση αξιοποιείται κάτοψη του χώρου που προκύπτει από το τρισδιάστατο μοντέλο. Τέλος, ο χρήστης μπορεί να επιλέξει ποια από τις διαθέσιμες τεχνικές προσομοίωσης επιθυμεί.



Σχήμα 2.1 Στιγμιότυπο του εργαλείου ακουστικής προσομοίωσης χώρων

Το δεύτερο είδος χρήστη αφορά τη συνεισφορά περιεχομένου. Η συνεισφορά γίνεται μέσω ηχογραφήσεων σε καινούργιους χώρους, με βάση ένα σύνολο οδηγιών. Ανάλογα με την τεχνική προσομοίωσης που θα χρησιμοποιηθεί ακολουθείται μια διαφορετική διαδικασία, ενώ ο χρήστης δηλώνει την τεχνική που χρησιμοποίησε. Στόχος του έργου είναι η δυνατότητα συνεχούς επέκτασης με νέους διαθέσιμους χώρους. Για αυτό τον λόγο παρέχονται τα εργαλεία για την υλοποίηση προσομοιώσεων με διαφορετικές τεχνικές, δίνοντας στους χρήστες τη δυνατότητα συνεισφοράς ανάλογα με τον διαθέσιμο εξοπλισμό τους.

### 3.2 Τεχνικές ακουστικής προσομοίωσης χώρων

Η πρώτη επιλογή απαιτεί τη χρήση φορητής ψηφιακής συσκευής εγγραφής B-Format χαμηλού κόστους. Αυτό επιτρέπει την καταγραφή κρουστικών αποκρίσεων για κάθε συνιστώσα  $X$ ,  $Y$ ,  $Z$ ,  $W$  του προτύπου B-Format. Η επιλογή αυτή προτιμάται τόσο για το πλεόνασμα πληροφορίας που προσφέρει, και τη συμβατότητά της με περιβάλλοντα εμπύθισης, όσο και γιατί είναι σε σύμπτωση με δημοφιλή υπάρχοντα αποθετήρια, όπως το γνωστό σύνολο δεδομένων OpenAir [3]. Για την ακουστική διέγερση του χώρου χρησιμοποιούνται ημιτονοειδής σαρώσεις ή και σκάσιμο μπαλονιού. Η δεύτερη προσέγγιση βασίζεται σε μοντέλα βαθιάς

μάθησης. Ως είσοδος χρησιμοποιείται μια καταγραφή ομιλίας στον χώρο ενδιαφέροντος και μια καταγραφή ομιλίας σε συνθήκες στούντιο. Το μοντέλο στοχεύει στη μεταφορά ύφους (style transfer) από την ηχογράφηση του χώρου στην κατά προσέγγιση ανηχική ηχογράφηση. Αυτό μπορεί να γίνει είτε απευθείας (end-to-end) είτε μέσω ενός ενδιάμεσου σταδίου τυφλού υπολογισμού παραμέτρων αντήχησης και ενός σταδίου φίλτραρίσματος της ανηχικής ηχογράφησης βάσει αυτών των παραμέτρων.

#### 4. Συμπεράσματα και μελλοντικές επεκτάσεις

Στην εργασία παρουσιάστηκε το πλαίσιο λειτουργίας που υλοποιείται στο πλαίσιο του προγράμματος SCENE, και ειδικότερα το εργαλείο προσομοίωσης ακουστικής χώρων. Πρόκειται για μία γραφική διεπαφή που επιτρέπει στον τελικό χρήστη να πλοηγηθεί στη βάση δεδομένων της πλατφόρμας και να ακούσει την ακουστική απόκριση διαφορετικών χώρων παραμετροποιώντας το είδος, τον αριθμό και τις θέσεις των ηχητικών πηγών. Παράλληλα μπορεί να συμβάλλει στην επέκταση της βάσεις παραχωρώντας καταγραφές νέων χώρων, ακολουθώντας ένα σύνολο βασικών οδηγιών

Στις ερευνητικές κατευθύνσεις του έργου είναι η υλοποίηση και αξιολόγηση των τελικών μοντέλων ακουστικής προσομοίωσης και η ενσωμάτωσή τους στο περιβάλλον, καθώς και η εξέλιξη της γραφικής διεπαφής με στόχο την αναβάθμιση εμπειρίας χρήστη.

#### 5. Ευχαριστίες



Η παρούσα έρευνα είναι μέρος του ερευνητικού έργου SCENE που έχει λάβει χρηματοδότηση από το πρόγραμμα της ΕΕ Horizon Research and Innovation με αριθμό No 101095303.

#### 6. Αναφορές

- [1] Szöke, I., Skácel, M., Mošner, L., Paliesek, J., & Černocký, J. “Building and evaluation of a real room impulse response dataset.” *IEEE Journal of Selected Topics in Signal Processing*, 13(4), 863-876 (2019).
- [2] Carlo, D. D., Tandaitnik, P., Foy, C., Bertin, N., Deleforge, A., & Gannot, S. (2021). dEchorate: a calibrated room impulse response dataset for echo-aware signal processing. *EURASIP Journal on Audio, Speech, and Music Processing*, 2021, 1-15.
- [3] Murphy, D. T., & Shelley, S. (2010, November). Openair: An interactive auralization web resource and database. In *Audio Engineering Society Convention 129*. Audio Engineering Society.
- [4] Vryzas, N., Vrysis, L., Stamatiadou, M. E., Kotsakis, R., & Dimoulas, C. “The effect of geometry information in blind estimation of room acoustic parameters.” In *Audio Engineering Society Convention 154*. Audio Engineering Society (2023, May).

- [5] Vryzas, N., Stamatiadou, M. E., Vrysis, L., & Dimoulas, C. “Multichannel mobile audio recordings for spatial enhancements and ambisonics rendering.” In 2023 Immersive and 3D Audio: from Architecture to Automotive (I3DA) (pp. 1-6). IEEE (2023, September).
- [6] Müller, K., & Zotter, F. “Auralization based on multi-perspective ambisonic room impulse responses.” *Acta Acustica*, 4(6), 25 (2020).
- [7] Scheibler, R., Bezzam, E., & Dokmanić, I. “Pyroomacoustics: A python package for audio room simulation and array processing algorithms.” In 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 351-355). IEEE (2018, April).
- [8] Diaz-Guerra, D., Miguel, A., & Beltran, J. R. “gpuRIR: A python library for room impulse response simulation with GPU acceleration.” *Multimedia Tools and Applications*, 80(4), 5653-5671 (2021).
- [9] Kamaris, G., Giannatsis, E., Kaleris, K., & Mourjopoulos, J. “Suitability of game engines for virtual acoustic experiments.” In Audio Engineering Society Convention 146. Audio Engineering Society (2019, March).
- [10] McArthur, A., Van Tonder, C., Gaston-Bird, L., & Knight-Hill, A. “A survey of 3D audio through the browser: practitioner perspectives.” In 2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA) (pp. 1-10). IEEE (2021, September).
- [11] Vryzas, N., Vrysis, L., Stamatiadou, M.-E., Dimoulas, C. & Kalliris, G. “An acoustic simulation interface for cultural heritage locations scouting” in Proceedings of the 156th Audio Engineering Society Convention (2024)
- [12] Gorzel, M., Allen, A., Kelly, I., Kammerl, J., Gungormusler, A., Yeh, H., & Boland, F. “Efficient encoding and decoding of binaural sound with resonance audio.” In Audio Engineering Society Conference: 2019 AES International Conference on Immersive and Interactive Audio. Audio Engineering Society (2019, March).

