

**Correspondence Author**

**Licence**

Jack Hoggard

CC BY 4.0

**New Vistas**

Volume 12, Issue 1, 2026, UWL Annual Doctoral Students' Conference 2025

**Abstract**

In music production, spatial audio has seen slow progress, despite the advantages it offers with respect to the reduction of masking and listener envelopment. Producing spatial audio is significantly more difficult than traditional stereo mixing, it requires specifying several parameters for each sounds location and spread in three dimensions. This has opened up the opportunity for intelligent tools to assist users in implementing the technical aspects of a spatial mix while leaving room for artistic expression.

The aim of this research is to develop and validate an intelligent machine learning model that is designed to reduce the number of parameters needed for the production of spatial audio. To accomplish this, a machine learning model will be trained on datasets compiled of Room Impulse Responses (RIR), 3D multitrack mixes and spatial panning parameters. Prior to training, we plan to devise and assess different ways to objectively measure the quality of spatial audio.

The first approach will be to optimise a reduced parameter set for an individual audio mix with a defined instrument setup. The second approach will be to optimise the model to generalise to unseen audio mixes with unknown sources. The interface itself will be assessed through user studies to measure usability and perceptual quality. A dedicated vocabulary will guide the user reviews, providing a structured framework for evaluating the system's immersive capabilities. The success of this research will allow the production of spatial audio to become simpler and more accessible to non-specialists, which could pave the way for widespread adoption of spatial audio.

**Keywords**

Intelligent user interfaces, Spatial audio