

SynthFx 12K Dataset

Nelly Garcia¹ and Joshua Reiss²

¹Centre for Digital Music, Queen Mary University of London, UK., n.v.a.garcia-sihuay@qmul.ac.uk

²Centre for Digital Music, Queen Mary University of London, UK.

Abstract— Procedural audio, a method of generating sound from scratch using computational processes, is emerging as a promising approach to creating sound effects. However, limited access to such samples has hindered further research and optimisation of models. This dataset provides both real and synthetic samples for researchers, audio developers and sound designers to use in future projects.

Index Terms— sound effects, dataset, procedural audio, sound synthesis, environmental sounds.

I. INTRODUCTION

Sound effects libraries are basic in audiovisual and audio-only projects, serving as essential resources for tasks such as classification, generation and retrieval. We define sound effects as non-musical and non-dialogue audio elements. The scarcity of publicly available datasets, combined with the cost and time involved in creating custom sound libraries, is a significant barrier to reproducibility and new research in the field.

Most existing datasets consist of pre-recorded sounds, with limited availability of synthetic samples. These synthetic samples, generated using various sound synthesis methods, are particularly important for sound effects and environmental audio. Furthermore, most auditory recognition research has focused on speech and music processing, leaving sound effects comparatively underexplored.

The aim of this dataset is to support open research in sound synthesis by:

1. Contributing with a public dataset containing both synthetic and pre-recorded library samples.
2. Providing detailed information about the sound synthesis methods used in each category, allowing further optimisation of these methods.
3. To support the development of evaluation methods for synthetic sound models.

II. PREVIOUS WORK

A common approach to creating sound effects is procedural audio, also known as digital Foley, which uses computational processes to create sound from scratch. However,

Table 1: Types of sound synthesis used in sound categories. SM = Synthesis Method; P. Informed = Physically Informed

Sound Category	SM 1	SM 2	SM 3
Bubbles	Additive	P.Informed	Modal
Droplets	Additive	P. Informed	Modal
Engine	Additive	-	-
Explosion	Additive	P. Informed	Modal
Fire	Additive	Modal	Subtractive
Gunshot	P. Informed	Additive	Subtractive
Rain	Additive	P. Informed	Subtractive
Waves	Subtractive	Additive	-
Jet	Additive	P. Informed	Modal

there is no standard process for evaluating procedural audio models, and access to procedural audio samples in research is still limited.

Previous studies, such as [1], have compared the perception of different sound textures (e.g. fire and rain) using different sound synthesis methods. However, these studies do not cover all sound categories and the samples used for comparison are not publicly available.

III. DATASET

This dataset contains 12,000 five-second sound samples, evenly distributed across 30 categories of real and synthetic audio. Real samples, taken from sound libraries such as the BBC, Hybrid, Soundsnap and Pixabay, are paired with synthetic samples generated using the Nemisindo online procedural audio engine (Nemisindo). To enhance realism, the synthetic samples have undergone reverberation and equalisation adjustments.

Each sample is 5 seconds long with 44.1KHz, mono and 16bps, with pairs (e.g. samples 0 and 1, 2 and 3) representing the same sound category. Table 1 gives an overview of the synthesis methods used by category, derived from [2]. For a detailed list, see the full documentation: .

IV. REFERENCES

- [1] D. Moffat and J. Reiss, "Perceptual evaluation of synthesized sound effects," *ACM Transactions on Applied Perception*, vol. 15, pp. 1–19, 04 2018.
- [2] D. Menexopoulos, P. Pestana, and J. Reiss, "The state of the art in procedural audio," *Journal of the Audio Engineering Society*, vol. 71, pp. 825–847, 12 2023.