

Zen Foley: Audio Effects

Zen Research Authors
Zen Research DAO
Zoo Labs Inc (501(c)(3) Non-Profit)
San Francisco, California, USA
dev@hanzo.ai
+1 (913) 777-4443

September 2025

Abstract

Video-to-audio Foley generation.

1 Introduction

Video-to-audio Foley generation.

1.1 Key Features

- Video-to-audio synthesis
- 48kHz professional audio
- Multi-scenario sync
- Natural sound effects

2 Technical Specifications

Parameter	Value
Type	Video-to-Audio
Quality	48kHz Hi-Fi
Sync	Multi-scenario
Output	Sound effects

Table 1: Technical specifications

3 Zen AI Ecosystem

Part of the complete Zen AI hypermodal ecosystem:

Language: zen-nano-0.6b, zen-eco-4b, zen-agent-4b

3D & World: zen-3d, zen-voyager, zen-world

Video: zen-director-5b, zen-video, zen-video-i2v

Audio: zen-musician-7b, zen-foley

Infrastructure: Zen Gym (training), Zen Engine (inference)

4 Conclusion

Video-to-audio Foley generation.

Acknowledgments

Thanks to the open-source community and our upstream contributors.