EMBEDDED REAL-TIME EXPRESSIVE GUITAR TECHNIQUE RECOGNITION

Domenico Stefani
Creative, Intelligent & Multisensory Interactions Laboratory
University of Trento
domenico.stefani@unitn.it

Recent advancements in deep learning and embedded platforms for real-time audio processing have made it possible to consider embedding advanced algorithms into musical instruments. Among the many possible applications of such systems, this offers the opportunity to further develop the 45-years-old concept of *guitar synthesizer*: a real guitar, paired with a sound synthesizer that can enable guitarists to sound nothing like a guitar.

Albeit innovative at the time of their inception, these instruments quickly lost interest from many guitarists as the tracking systems used to interface the instrument with the synthesizer were rather limited. Throughout the years, these were improved in terms of the accuracy achieved in tracking both the pitch and envelope of the guitar notes. However, these systems did not manage to track many of the nuances introduced in the sound by guitar players through the use of *expressive techniques*, such as palm-mute, harmonics, the use of different picking positions, or even percussive fingerstyle in acoustic guitars. This results in the lack of ability to transfer this expressiveness from the "controller" to the synthesizer, requiring players to detach their hands from the instrument to shape the synthetic sound through a conventional and inadequate synthesizer interface composed of knobs and sliders.

The recent developments in and embedded platforms for audio make it possible to consider embedding into a regular guitar a system that can track the expressive technique used by the player, along with pitch and temporal evolution of notes, and reuse this information in real-time to shape a synthetic sound or audio effects. Moreover, the generation of the synthetic sound can be embedded into the instrument too, like in *smart musical instruments* [1]. Finally, real-time information about the expressive techniques used can be exploited to control other systems external to the instrument, such as stage lighting setups or video generation for music performances.

In this presentation, I will report on the progress of our research group in developing an embedded, real-time *expressive technique recognition system* for the guitar. I will briefly introduce the challenges of real-time Music Information Retrieval (MIR) on embedded systems, along with the initial results that we obtained with a monophonic technique classifier, developed to recognize pitched and percussive techniques on the acoustic guitar [2]. Our initial efforts resulted in an accuracy of 99.1% in recognizing percussive acoustic guitar techniques, with an average onset-to-result latency of 30.7 ms. However, we are working towards a reduction of latency to 20ms, with promising results. This first system yielded lower accuracy results with pitched techniques, which is to be largely attributed to the lack of refinement and tuning of the different stages of classification, which include onset detection, feature extraction, and a deep classification network.

I will discuss how such a multi-stage classification pipeline can be beneficial to the latency of real-world embedded implementations. The technique recognition system was implemented as an audio plugin and deployed to a Raspberry PI4 board with the real-time operating system *Elk Audio OS* [3], which provides quick and high-resolution audio processing capabilities.

Then I will present the progress that was made to improve the accuracy and latency of our embedded classifier, which includes fine-tuning the onset detection stage to reduce latency variability [4] and investigating the performance differences between four deep learning inference engines for embedded CPUs [5].

In our work we aim at developing both theoretical advancements and embedded implementations, as well as conducting measurements of performance at computational and memory levels, which contributes to the research field of embedded real-time MIR and in general to the field of embedded AI for NIME.

1. REFERENCES

[1] L. Turchet, "Smart Musical Instruments: vision, design principles, and future directions," *IEEE Access*, vol. 7, pp. 8944–8963, 2019.

[2] D. Stefani and L. Turchet, "On the Challenges of Embedded Real-Time Music Information Retrieval," in *Proceedings of the 25-th Int. Conf. on Digital Audio Effects (DAFx20in22)*, vol. 3, pp. 177–184, Sept. 2022.

[3] L. Turchet and C. Fischione, "Elk Audio OS: an open source operating system for the Internet of Musical Things," *ACM Transactions on the Internet of Things*, vol. 2, no. 2, pp. 1–18, 2021.

[4] D. Stefani and L. Turchet, "Bio-Inspired Optimization of Parametric Onset Detectors," in *Proc. 24th Int. Conf. on Digital Audio Effects (DAFx20in21)*, vol. 2, pp. 268–275, Sept. 2021.

[5] D. Stefani, S. Peroni, and L. Turchet, "A Comparison of Deep Learning Inference Engines for Embedded Real-Time Audio Classification," in *Proceedings of the 25-th Int. Conf. on Digital Audio Effects (DAFx20in22)*, vol. 3, pp. 256–263, Sept. 2022.

Video Recording from the NIME Workshop Session:
https://youtu.be/H99MRsQ6U3Y?t=4760