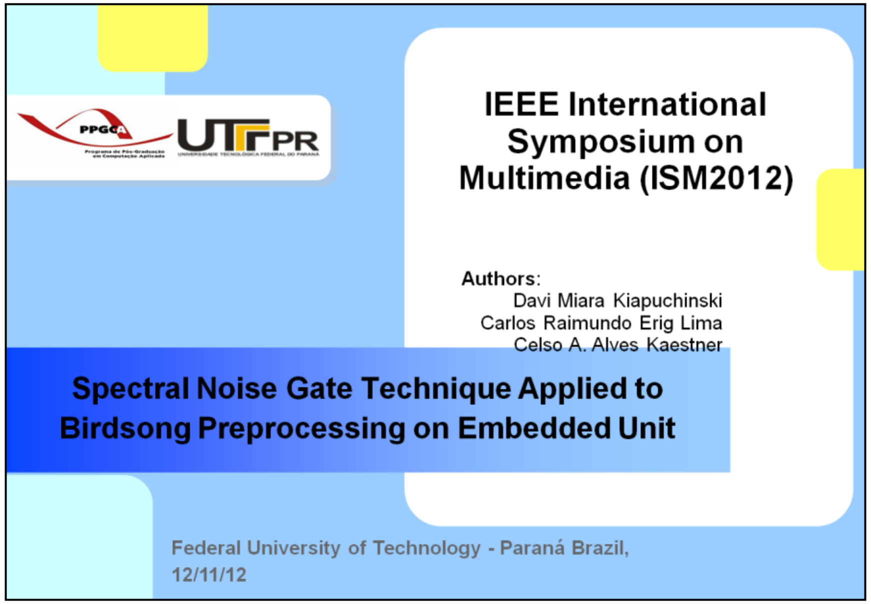


Hello my name is Davi Miara Kiapuchinski, I am an engineer, student of the Master's computing program in Federal University Technology of Paraná and on of the author of the work that I will present.

Firstly I apologize for any possible failures in my English and in my vocabulary.

Our work, called Spectral Noise Gate Applied to Bird Song Classification, is part of a university project entitled BRBird who cares about improving the state the art and of the technique in research with birds.

Specifically this work is based on promoting the research concerned with the bird song automatic classification.

This branch has presented impressive successes in recent years, over 95%, for example, using algorithms such as J4.8.

Another example is the work of some of the participants of BRBIRD, which achieved hits over 99% using some sound manipulations made after the classification phase.

The research is interesting because even with these successes, there are still some problems that the automatic bird song classification can remedy. Motivates this research factors such as: concern animal, wildlife control, the lack of additional information on the birdsongs to help in classification process, the safety of human beings life and especially difficult to amplify the good results of the classifiers in real environments, with real and noisy samples. These samples that carry, in addition to the song, sounds of wind, rain and a wide type of noises.

This fact demands a repetitive work and a manual preparation of samples before inserting them into the classify.

With the intention to help solving this problem this paper aims to: build an embedded environment capable of collecting and processing the sounds of the environment in which it operates. After process it, to extract its features to be used in a classifier.

The process of sound processing applied to the birdsong classification has some peculiarities. The main one is the fact that these animals most often live in environments with a lot of noise that 'blurs' the spectrum.
Another feature is the wide spectral range that these animals can emit sounds. While humans speak in a frequency band ranging from 1k to 3kHz, the birds sing in frequencies of 100 to 10kHz.

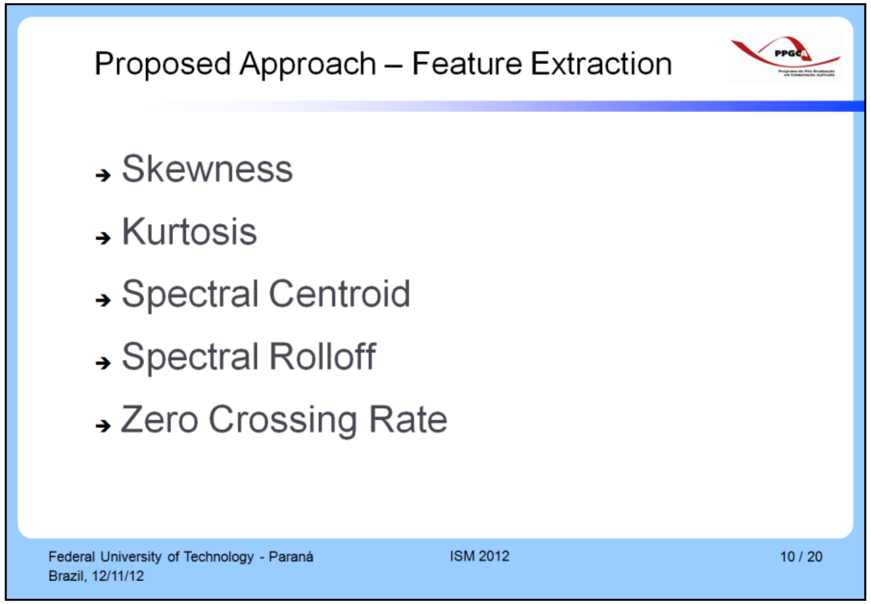
Several communities, companies and researchers work so that these peculiarities can be applied with the automatic classification algorithms that achieve good results with samples previously prepared.

The Xeno-Canto website is a community that gathered birdsongs in its natural form in several locations of the world. The Cornell Lab of Ornitology, is a lab that develops products (hardware and software) for the area. The Brazilian CENIPA is a commission to investigate the causes of aviation accidents.

The figure shows a result of an important researcher in the area. Agranat in one of his works had a noise filter that eliminates blurred in the sound spectrum and emphasizes the song. For this, he manually apply various filters and a filter with logarithm transformation of frequency.

The proposed approach aims to deliver to the automatic birdsong classifier a sound clean and prepared. Therefore, filters were programmed and a process of extracting features of this sound was implemented and materialized in an embedded environment.

The filters routine, based on the sound manipulation algorithm called Spectrum noise gate, works using a sample noisy acoustic signature. This signature, drawn from every window of the spectrum, is compared with the windows in the sound to be filtered. This creates indexes of smoothing that are also influenced by the values of smoothing windows of nearby neighbors. The resulting sound is subjected to inverse Fourier transform and saved in. WAV file.



From this new sound, some features are extracted to obtain a template acoustic of the signal. Functions such as skewness, kurtosis, Spectral centroid, Spectral

The hardware architecture used in this work consists in a Sansung processor with a ARM920T (nine hundred twenty) core with 32 bits. This processor includes interfaces such as SPI, USB Host, RS232, PWM, Direct Memory Access, Real Time Clock, sound interfaces like AC97 and I2C. The processor also has an 8 channel AD converter, 10 bits of resolution and five hundred thousand samples per second.

We used a version of embedded Linux with kernel on 2.6.32.2 version. The cross compiler used was GCC 4.3.2.

The experimental phase was divided into three parts: the sound recording on equipment, spectral analysis in the pre-processing and feature extraction.

The phase of the sound recording was tested in an environment with a mono microphone (only a single channel), the samples are stored in the FLASH memory device with 8 or 16 bits of precision, with sampling rate up to 44100 Hz (forty-four thousand and one hundred).

The analysis phase of the sound included tests of Fast Fourier transform in the embedded environment using three windowing functions: Hamming, Hanning and Bartlett or triangular function.

The table shows an example of a feature extraction test conducted with a sample of 3 seconds duration, and 16 bits per sample, using the Bartlett windowing function.

This is an example of use of the filter. Sets up a small sample of the signal and this will be the noisy profile to be deleted from the original sound. This sample may be a noise pattern previously inserted in the equipment or a part of the sample that is only noise.

The figures show the same spectrum. First on time domain and then on frequency domain. The first sample is the noise signal and the graph of half is the original signal. The third signal shows the result of the filter. The spectrum is a signal without unwanted noise founded in the noisy profile.

Additional tests were made and it was noted the usefulness of filter to remove sounds contained in the sample, as the sound of rain or winds. The sign above presents a signal on the time domain and also on the frequency domain with the sound of a waterfall being eliminated from the sample. Leaving only the interesting to be analyzed, ie the birdsong.

Additionally, we analyzed the efficiency of processing time of the routines implemented in embedded software, to analyze the feasibility of a real implementing with the equipment. The tests showed an average time of 1s for performing an FFT on a signal of 3 seconds with sixteen bits of precision. The total time for the entire filtering process was up to 24 seconds (at worst case) to a Hanning windowing function, and Forty-Four thousand and one hundred Hz of sampling rate. A regular case that guarantees a good sound quality and a reasonable response time is a case of using a sampling rate of Twenty-two thousand and one hundred Hz with the Bartlett windowing function. This situation can provide a result in less than 8 seconds of processing.

Our research continues, other results are being written in other articles. This is a image of the equipment ready to perform real tests. A battery was incorporated in with the intention to increase the autonomy of the equipment

It was used for the sound processing and automatic classification (in other studies) tests, tools like Audacity, the Marsyas and Raven Lite.

Thanks Araucaria Foundation and the National Research Council that funded the study.

Here are some key references, I am now open to questions, my contacts are shown and until a next opportunity.

THE END

Davi Miara
davi.miara@ctim.mar.mil.br