**Aims & background material (student)**

Acoustical timbre is defined for the purposes of this project as the qualities characteristic to a sound that allow it to be recognised as having been produced by a particular voice or instrument. Thus, timbral characteristics allow sounds of the same pitch and volume to be differentiated when they are produced by a different body.

The aim is to develop a system that automatically extracts the timbral qualities from a digital audio signal, so that recordings of an instrument or voice can be identified in absolute (by labelling) or relative (by comparison) terms. Accurate description of the timbre specific to a sound, such as a particular model of instrument or a person's voice, is a challenging problem not usually tackled by a single method; my research so far has shown that a large variety of methods have been applied to timbral analysis in the literature. Thus a combination of signal processing and machine learning methods may provide interesting results in this context.

The applications of timbral analysis are wide-ranging, from automatic audio metadata labelling and music technology to assisting automatic speech recognition in diarization.

**Student Summary of project deliverables, fallbacks & extensions (student)**

The first task being undertaken is researching existing work in the field of timbral analysis, in order to gain an understanding of the current state of the art, main developments in the field, and to start sourcing potential methods to implement, or reference for the project implementation. This research will allow me to learn about and compare a wide array of approaches to the problem, which will inform the decisions I make at the start of the project in selecting a method to develop myself. I will then be able to decide whether to implement a novel, more ambitious method inspired by the literature review, or to combine or re-purpose an existing approach in the literature, perhaps applied to a new context; depending on how ambitious the methods are and my informed appraisal of my knowledge, resources and ability. This research phase also includes sourcing relevant datasets of labelled audio for potential use in training and evaluating a timbral analysis system.

As a first step in developing a timbral analysis system to carry out alongside the literature review, I plan to test and implement open-source specifications or codebases for relevant signal processing tools for analysing timbre, which will help me understand these methods as well as allow me to develop a “toolbox” for later use in the project as part of a larger system. This will allow me to immediately start incrementally building the project’s components before having chosen a particular direction to specialise the project, and eventually to develop a basic proof-of-concept system as a fall back in case the full system I plan (for instance a more end-to-end machine learning system outputting labels) is infeasible in the scope of the project.

Following this iterative development process, I will ensure I leave enough time to evaluate the method(s) developed. This will require devising a scheme for quantifying/qualifying my system’s performance, and selecting suitable baselines (implemented by me or found in the literature) for comparison. As a possible extension, if promising results are achieved in the way of timbral analysis it may be interesting to explore applying the method to downstream tasks and evaluate the gains brought to them (for instance, perhaps improvements in speech diarization).

**Summary of Risks (student)**

I anticipate a potential risk of the project is the breadth of the aim, since timbral analysis does not correspond to any one task in audio processing, and timbre is to an extent a subject perceptual quality of sound. In my research, I will attempt to mitigate this issue by forming a more specific definition of timbre and how to represent it for the purposes of this project.

Over-ambition of the selected approach to timbral analysis is another potential risk; I can imagine trying to develop a method beyond the scope of the project and falling short due to time constraints and inexperience, and thus not obtaining any results. Hopefully, careful research and the incremental approach described above will help avoid the issue of over-stretching my ability. I also imagine that the scope of the project as described previously may have to be reduced depending on how challenging the problem is in reality (as it is still difficult for me at this point to appraise whether or not promising results could be realistically achieved in the scope of the FYP). In this case, the technical aims may have to be adjusted or restricted, still within the context of processing audio with the goal of extracting information providing some insight into the timbre of a sound.

I also expect to narrow the range of sounds (to short solo recordings of a single musical instrument for example) considered for the project, depending on the datasets available, as I don’t realistically expect to develop a system complex enough to handle any type of sound.