5. Introduction  
 1. Motivation (technical, maybe also personal)  
 2. Objectives  
 3. Structure of the Thesis

Imagine a party. You can hear a wide variety of sounds: music in the background, conversations between people, noises of somebody coughing, maybe even a dog barking outside… These sounds merge into a single stream, which approaches your ears by vibrations in the air and then goes through different physical, biological and psychoacoustical processes to finally come as electrical impulses to the brain. Despite all these sounds from different sources are mixed on the way to your ears, your brain can segregate one (or several) of them. You can focus your hearing on this “target” sound and separate it from the complex mixture, leaving other sounds in the background. This phenomenon has been described as a “cocktail party effect”, and the process of integrating separate sounds into meaningful streams (“auditory objects”) -- auditory scene analysis, or ASA.

In machine perception --- specifically in machine hearing --- a related concept is referred to as Computational ASA (CASA) and is tightly connected to the fields of sound recognition and digital signal processing. CASA systems indeed aim to separate sounds from mixtures, but they differ from BSS (blind source separation) systems in that they try to do this in a way a human ear does. Being based on and trying to combine works from different fields of science, CASA systems can bring new solutions and insights to the complex problem of signal separation.

The main objective for this thesis is to describe the principles and goals of CASA, existing applications and approaches. Another objective is to practically apply the theoretical knowledge and implement a simple CASA system to separate monophonic music from noise. But before all of this, since this thesis is made for an IT-oriented audience, it is needed to make a brief introduction to the underlying physics and biology.

Thus, the thesis is structured as follows:

* Firstly, physical background theory will be provided, including an introduction to what a sound is. Since the implemented system from the practical part aims to segregate music from noise, a special focus in this part will be made on describing harmonic sounds and pitch perception.
* Secondly, having in mind that CASA tries to mimic the human auditory system, a brief introduction to the biological structure of the human ear will be made. Here, auditory scene analysis according to Bregman will be introduced too.
* Next, to cover the math in the implementation part, the basics of digital sound processing will be described. The related mathematical principles and functions used during the implementation will also be given some attention.
* In the following chapter, having all the related theory in mind, an introduction to the main principles and goals of CASA will be made, along with an overview of its applications and selected models.
* Next, in the practical part, the focus will be made on describing the implementation of specific parts of the CASA system built for this thesis (see attached medium).
* Finally, an overview of the experiments made to test the implemented system will be provided.