Abstract

Blind Source Separation is a problem that occupied the mind of scientists ever since signals could be stored in digital form. Useful solutions to this problem have been found under a variety of conditions. Separating music sources in a monaural recording have not yet been successfully achieved. This thesis presents my own attempt at solving this task. A separating software based on a Long Short-Term Memory Recurrent Neural Network was written along with a platform for using it. The platform is formed by a server and two graphical user interfaces: one for mobile devices and another that is web-based.

In Chapter 1 is described the problem statement along with the motivation for the thesis's subject. Chapter 2 focuses on explaining technical terms used for implementing and understanding the software, and presents scientific discoveries that the application is based on. Chapter 3 examines related work, such as vocal isolation algorithms. Chapter 4 begins by taking a look over different approaches to solving our audio separation problem, and then presents in detail the selected approach and the reasoning behind it. In this same chapter technologies used and data used are mentioned. The last chapter shows the results obtained, and proposes future developments. The chapter also incorporates a personal view on the work and on the subject.

The platform and separation algorithm are implemented and working. The separation software was trained to recognize drums from vocals in a mixture of the two. The algorithm is not as efficient as expected, mostly because training the network is extremely expensive. The future of the software, however, looks bright. The network will keep training until a satisfiable separation is achieved. The scalable structure of the application will help expand it to real platform for music processing, as being able to unmix a song opens the door for a limitless number of other programs.