

# Preface

Depending on your source of reference, there are between 7 and 8 billion mobile phones in use at the time of writing this work. It is a breathtaking technological success. In less than 20 years, mobile telephony has reached all parts of the world and there is an almost universal access to mobile phones for all people in the world. That is not to say that all 7 billion people of the world would have their own phone since many have multiple phones, but a large majority of people does have *access* to a mobile phone. Almost all people in the world have the option of making a (mobile) phone call.

The consequences of this technological leap are as astonishing as the speed with which this leap has happened. In the industrial world, practically every adult has a mobile phone. It is a defining characteristic of our age—to be always reachable, to be always online. It has fundamentally changed our attitude towards communication. Being within reach is no more a thing that needs to be planned, gone are the days when we would agree the time and place when we would talk the next time. We are simply always there.

Improved telecommunication solutions have the potential of increasing the amount of communication between people. This is almost only a good thing. Increased communication improves understanding between people, it increases awareness of society and your surroundings, and has thus an impact on the level of democracy.

In the industrial world, democracy can of course be *nice*, but where mobile telephony has had a much larger impact on democracy is the developing world. It is fairly common for small, poor villages and communities in Africa to share one mobile phone for the whole village. It has proved to be more expensive to build and maintain land-lines than mobile networks, and thus hardly any new land-lines are being built. There are plenty of villages that never had a phone, never had an instantaneous communication channel, before the mobile phone. Where before, people had to walk to the next village to get news from their relatives, they can now simply make a phone call. They can even talk to their relatives who have moved abroad, perhaps even overseas, and get the latest news. Imagine the societal and human impact of that: For the first time in human history, any one person in the

world can connect with anybody else in an instant, independent of geographic location!

As with every success, there is a drawback; the huge steps forward have proved to be the biggest enemy of the technology. For example, it would be easy to make the conclusion that since 7 billion people have access to phones, *it works already* and there is thus not much to be improved. Alas, how often have I not heard this claim, even from scientists in the field!

Fortunately, the same argument can be turned upside down to demonstrate the importance of further improvements to mobile telephony. Imagine the impact that the tiniest of improvements would have, when multiplied by 7 billion users! The largest fixed costs of mobile networks surely lie in the network, where efficiency and costs are directly proportional. A 1% improvement in coding efficiency, in the actual amount of data sent, does not have a large impact on the individual phone, but worldwide, already the reduced energy consumption due to a 1% improvement in efficiency would have a huge impact.

At the same time, as the number of devices capable of speech coding is exponentially increasing, the amount of energy required is also increasing rapidly despite improvements in energy efficiency of CPUs. To make it feasible to use speech codecs on such a large number of devices, it will be necessary to reduce the complexity of speech coding algorithms. This is especially evident with the emerging concept of the internet-of-things, where it is envisioned that all electronic devices would be connected and capable of communication. It is obvious that we need to limit the complexity of codecs to reduce power consumption if they are to be implemented on such a large number of devices.

Moreover, even if only a fraction of the improvement yields lower costs for the end-users, it is not difficult to see its impact on that small African village. 50 cents saved per month is good money.

The other obstacles for mobile telephony imposed by its own success are the economic incentives within standardisation processes. Regrettably, however, the current author is not authorised to pursue an academic discussion of the economic incentives of standardisation any further.

It is then not surprising that during the last decade academic research in speech coding has been scarce. It seems that most of the cutting-edge research happens behind closed, corporate doors and little if any information is leaked out. All that is visible to the outside are the (sometimes) brilliant engineering art-works of speech coding standards, which include hardly any new science. Speech codecs have become very finely tuned machineries which are so complex, that the tiniest modification in any part inevitably breaks something somewhere else.

The purpose of this book is to take a step towards science. The objective is to formalise the most common speech coding tools into a scientific framework, such that their strengths and weaknesses can be assessed in isolation, without the complex interconnections of a full-scale codec. The hope is that this work will give a stable scientific framework which allows for new innovations and paves the way for new break-through technologies.

The emphasis in this book is in understanding *why* and *how* the commonly used methods work. Specific details about standards have for the most parts been intentionally left out to keep the overall structure clear, but ample references are provided for the experts.

The book was originally created as a compendium for the course “Speech Coding” which I taught at the Friedrich-Alexander University Erlangen-Nürnberg (FAU). It is therefore designed to be useful as learning material for students working on a master’s degree in speech and audio processing, as well as information or communications technology. In parallel, the book is also meant to provide in-depth information for engineers and scientists working in the field and can thus be used as a handbook on speech coding.

Erlangen, Germany  
October 2016

Tom Bäckström

Speech Coding

with Code-Excited Linear Prediction

Bäckström, T.

2017, XXI, 240 p. 76 illus., Hardcover

ISBN: 978-3-319-50202-1