Part V

CONCLUSIONS

—There are two possible outcomes:
if the result confirms the hypothesis, then you've made a measurement.
If the result is contrary to the hypothesis, then you've made a discovery.
Enrico Fermi

13

Conclusions and Outlooks

The main motivations which underlie this dissertation work spring from the new directions towards which speech telecommunications are going to. *Immersive speech communications* are becoming a reality and soon enough will become part of our daily life. However, immersive communications entail the use of displaced microphones and moreover take place in multisource environments where interfering signals may degrade quality and intelligibility of the desired speech source. Therefore, acquisition of desired signals with high quality is far more difficult and challenging for immersive communications than in the classical telephony environment where the microphone is close to the user.

Thereof the necessity to develop *intelligent acoustic interfaces* is increasingly pressing. An intelligent acoustic interface aims at extracting, from audio signals, desired informations for an acoustic environment, and, at the same time, has to reproduce remote desired acoustic information taking into account the perceptive requirements of a speech communication. To this end an intelligent acoustic interface has to model the acoustic channel, and the more "intelligent" way to do that, looking on the user requirements, is to employ adaptive filtering algorithms.

In this dissertation work we have investigated adaptive algorithms expressly designed for intelligent acoustic interfaces. For this purpose, the work is structured in three main parts.

In the first part we dealt with linear adaptive algorithms in order to tackle acoustic limitations deriving from the modelling of the acoustic impulse response, the presence of nonstationary sources, the presence of interfering phenomena, such as the “double talk”. In this part, starting from the study of new class of adaptive algorithms, such as the *proportionate algorithms*, we have formulated an alternative framework for the derivation of both classic stochastic algorithms and proportionate ones. Moreover, we proposed efficient proportionate algorithms based on the *affine projection* and on the *variable step size*, able to model an acoustic path even in adverse environment conditions.

In the second part we took into account nonlinear limitation, caused by the introduction of loudspeaker distortions in the acoustic path. This is a quite tricky problem, since nonlinearities strongly decrease the quality of a speech communication and due to the fact that commercial nonlinear filtering algorithms are not able yet to satisfy the quality requirements of a speech communication. In order to address this problem we proposed a novel nonlinear filtering model, called *functional link adaptive filter*, that we have used to develop *ad hoc* nonlinear adaptive algorithms for the modelling of nonlinear acoustic paths.

In the last part of this thesis, exploiting the adaptive algorithms proposed in the previous two parts, we developed more sophisticated adaptive filtering architectures which are more robust against adverse conditions of real scenarios. Such architectures have been developed exploiting the capabilities of *adaptive combinations of filters*. The main motivation, which underlies this study, is based on a common problem in the modelling of a nonlinear acoustic path. In fact, in this case, a kind of nonlinearity highly varying, in amplitude or in time, may require to change the filter design during the adaptation. Moreover, another important troubling situation occurs when the desired signal is not known *a priori*, thus it is difficult to choose whether adopting a linear filter or a nonlinear model. This trouble, along with the expensive computational cost of commercial nonlinear adaptive filters, affects the strategies of many companies that provide teleconferencing services, which often choose to employ only linear filters even at the expense of communication quality. In order to tackle this problem we proposed *collaborative filtering architectures* which are able to model an acoustic impulse response apart from its nature, whether it is linear or nonlinear.

The results achieved in this work pave the way for future research. A main relevance could be reserved to the modelling of nonlinear acoustic channel. In fact, the introduction of a new nonlinear model leads to novel interesting scenarios that can be deepened.

First of all, it could be possible to work on FLAF model in order to reduce the drawbacks making it a more consistent model. Moreover, it could be possible to exploits the capabilities of adaptive algorithms to develop more robust nonlinear adaptive filters. For example it is thinkable to apply the sparsity constraints to the modelling of the nonlinearities. This could lead to a further performance improvement.

Another important point is the fact that all the filtering techniques introduced in this work can be extended in the multichannel domain, due to the fact that immersive speech communications are based on the use of MIMO systems.

Moreover, as we have seen, a weak point of such techniques is their not appropriate evaluation. Immersive communications are based on perceived quality of a speech signal, thus the use of performance measures that includes also a perceptive evaluation of the filtering could be more proper.

Furthermore, being these proposed techniques very flexible, their use is not limited only to acoustic application, thus it could be possible to exploit their capabilities to develop *ad hoc* adaptive filtering algorithms.