A look at 2003

Anders Askenfelt, chairman

Department and the associated Centre for Speech Technology continued an extensive research program with 25 projects running during 2003. An all-time-high in PhD degrees was noted as six research students successfully defended their dissertations. The financial situation for the Department is good. The closing of the accounts for 2003 showed a result of +1.1 MSEK, giving an accumulated capital of 6.2 MSEK at the entry of 2004 (about 17% of the turnover). The financial setting is, however, presently being repainted with brisk brush strokes as regards EU-funded research, and the situation may change quickly.

We are strongly dependent on external funding, in particular from the European Commission. The projects within the Fifth European Framework Programme, in which the Department has been remarkably successful, are successively being brought to an end. According to official statistics, we have held a top rank among the departments at KTH in attracting European funding per researcher in FP5. Our establishment in the pending Sixth Framework Programme, which operates with much larger projects and constellations of partners, is in a preliminary stage. The formation of alliances of research groups for the proposals within FP6 has been a new type of activity we learnt during the year. So far the outcome for the Department consists more of well-founded expectations than of signed documents. It is of vital importance for the future financing of our research that we get on board for at least a couple of the central Integrated Projects in FP6 dealing with speech technology, sound-gesture communication, and multimedia-audio.

Two major events related to the organization of KTH left their marks in the year that passed. After a year of cooperation with research groups

from other departments in the research center *Wireless@KTH*, our Speech Signal Processing group found a move to the Department of Signals, Sensors and Systems (S3) strategically motivated. According to chair Professor Bastiaan Kleijn's judgement, the subject would develop more richly in a signal-processing-oriented environment like S3. Discussions had been going on since the summer of 2002 and the Rector made the move effective at the beginning of 2003. The change in organization has led to our department losing important competence in speech signal processing, a field which was taken up as a separate subject at KTH in 1996. We wish the speech signal processing group good luck in their work at the new department and look forward to future cooperation in suitable projects.

The other event was an instruction from Rector to investigate the possibilities of a merging of our department with the Department of Numerical Analysis and Computer Science (NADA). NADA is a very large department with nine major laboratories and research centers and close to 300 employees. After many lengthy discussions with representatives for the two departments it became clear that a merging was not motivated, and the question was removed from the agenda.

Leaving the financial and organizational matters, past and future, aside, it is assuring to observe that our research activity is strategically positioned in fields which will have a profound influence on future everyday life, comfort, and culture. The expectations are developing as samples of the new technology are becoming available. We are exploring the basic conditions for human communication with sounds – speech and music – and developing demonstrators of emerging technology. It is very reasonable to assume that the human voice will become the main communication channel for control of devices of all kinds: robots, computers, cars, and more. Besides the actual information in your spoken sentences, also your mood (as reflected in your voice) will be detected and used by machines. Everyone knows that a sad person needs to be met differently than a happy one, and machines will also learn to respond accordingly, in action, and in a spoken dialogue with you. Speaker identification and verification will be refined and regularly used in banking and other situations where personal identification is necessary. Naturally, voice communication will be supported by vision in many applications. A picture, or an animated agent, will add to the rate and clarity in information. 'A picture says more than a thousand words' is still valid.

The development of better hearing aids is a field that unfortunately is of interest to an increasing number of people. Our acoustical environment has changed dramatically over as a short period as a century, but not our delicate hearing mechanism. Noise levels at workplaces and in communities are subject to strict regulations, but exposure to high sound levels for pleasure is a subtler problem. Like smoking, information about the potential risks is only a partial solution to the problem. The inflation of sound levels at concerts, above all in rock music, has also spread to other genres. The main concern for working injuries among young symphony orchestra musicians is no longer pains in the arms or the back, but permanent hearing loss.

It will take at least a decade or two before actual repair work can be made in the inner ear (cochlea), or efficient audio-electrical cochlea implants are in regular use. Advanced external hearing aids are therefore of vital importance to an increasing number of people, old and young. The modeling of the signal processing in the ear, which is a core field of research in our hearing technology group, is basic for the development of more efficient hearing aids, automatically adapting to the patients' needs and the current environment.

A memorable event during the year was Stockholm Music Acoustics Conference, SMAC 03, which was organized for the third time. The conference recurs every 10th year with an unsurpassed punctuality during the first week in August. During the 20 years since the first SMAC in 1983, an impressive evolution from the research field of 'classic' music acoustics, primarily dealing with the acoustics of musical instruments, the singing voice, and the perception of musical sounds, has occurred. SMAC 03 covered a number of new academic disciplines including music psychology, music perception, music technology, computer music theory, and physics-based modeling of instruments. The expansion of the activity probably reflects the importance of music in society and everyday life, and an accompanying increased interest in explaining why music has such an impact on humans, how it is performed and perceived, and how new sounds and instruments could be designed.

Admittedly, the commercial aspects of music have another dignity today than a couple of decades ago. The easily accessible tools for composing, performing, processing, distributing, and retrieving music by computers add to the driving spirit in the field. The principal instrument of many a professional musician as well as of a multitude of non-professional music lovers is not longer a piano or guitar but the computer.

The Department will move to new facilities in April 2004 as the current building is passed over to the Swedish National Defense College. The plans for the reconstruction of our new building, originally raised as a physics laboratory in 1948, were finalized during the year. The new facilities will offer the Department an attractive research environment in the very heart of KTH campus, adapted to suit our needs. We are looking forward to the change and the year of 2004 with great expectations!-

Read more about the research and teaching at the Department on the following pages, and meet us also for current news, scientific articles, and upcoming events at *http://www.speech.kth.se.*