Concurrent Evolution of a Discipline and its Technical Society: The IEEE Signal Processing Society

James L. Flanagan

Advances in technology are stimulated through exchange and dissemination of basic, non-proprietary information. Technical societies provide the neutral forum whereby collaboration and knowledge creation can occur. The paths of technology and society are therefore closely linked. And, both are jointly molded by new-won knowledge.

In electrical engineering, the IEEE is the international umbrella for some 38 societies----each representing a recognized discipline within electrical engineering. I want to speak about the evolution of one of these societies----the Signal Processing Society (SPS). To quantify history, it's useful to recall the origin, which in this case is 1884 with the founding of the American Institute of Electrical Engineers (AIEE) [Figure 1]. This founding was prompted by the remarkable discoveries of the 1800's----exemplified by Oersted's relationship between electricity and magnetism, Faraday's electromagnetic induction, Morse's telegraph, Maxwell's prediction of electromagnetic propagation (later confirmed experimentally by Hertz), Bell's telephone, Edison's light bulb, and numerous others.

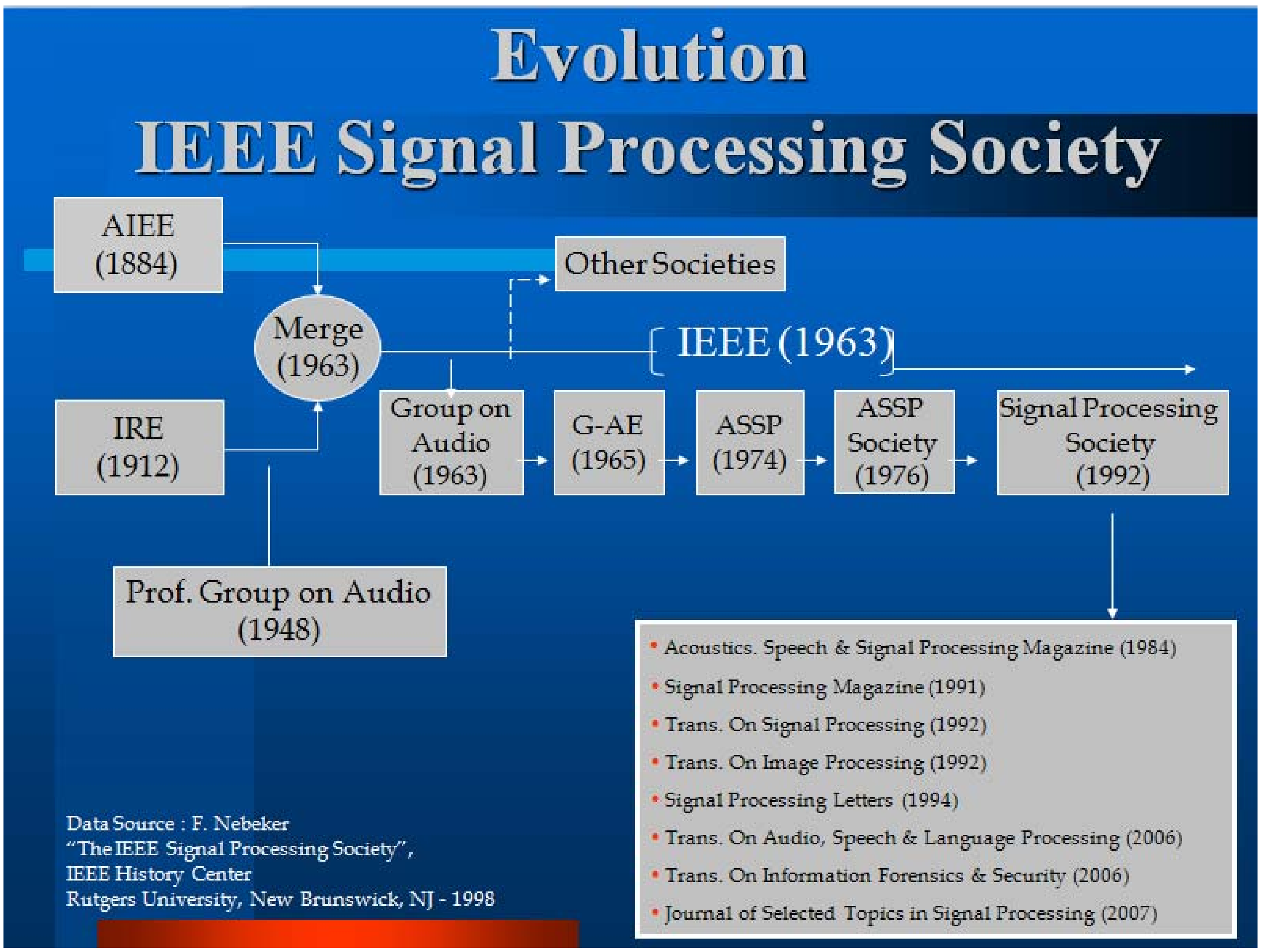

Figure 1. Evolution of the IEEE Signal Processing Society

Whereas AIEE was initially focused on power generation, transmission, and conversion, by 1912 electromagnetic propagation was being exploited. Radiotelephony was becoming an intense activity---- heralded by

Marconi's wireless telegraphy and DeForest's vacuum-tube amplifier----leading eventually to commercial radio broadcast. The Institute of Radio Engineers (IRE) was consequently formed to serve this sector.

With further advances, some of which were prompted by WWII, sound recording and transduction were becoming entertainment factors, as well as communication components. In response to this development, the IRE formed a Professional Group on Audio in 1948. **(**Interestingly, this was the year that solid-state amplification emerged----namely, the transistor**)**.

Eventually, cognizant of economy of scale, and a remaining overlap in technical focus, and a growing international interest, AIEE and IRE merged in 1963 to form the Institute of Electrical and Electronic Engineers (IEEE). In this merger, the Group on Audio was preserved. Over the years, this specialty area has undergone a series of transformations, which constitute the focus of this report. The initial era under IEEE brought us technologies such as disc recording **(**first 78 rpm, then 33 1/3 rpm Long Playing, and small 45 rpm**)**, and wire and tape recording. Very soon, by 1965, new methods for electro-acoustic transduction **(**microphones, loudspeakers, disc and tape recording heads**)**; for room acoustics; and, for stereo and high fidelity urged a transformation to the Group on Audio and Electro-acoustics (G-AE).

Audio technology continued strong, driven especially by telephony. Sampled-data theory and binary computation (profiting from pulsed-circuit techniques developed for radar during WWII) began to support algorithmic studies of traditional analog concepts. Digital processing of speech began to allow computer simulation of complete coding and transmission systems. And, an important link was made between digital filtering and spectral analysis through re-discovery of a digital algorithm, called the Fast Fourier Transform (FFT). In fact, in 1966, G-AE convened a select workshop on this topic to consolidate knowledge. This conference was held at Columbia University's Arden House, in upstate New York.

Advances in computing triggered a shift of interest---- from efficient speech transmission (or bandwidth compression) to voice interaction with information systems. This shift expanded research in speech----in its synthesis and automatic recognition. The emphasis on speech and digital simulation encouraged, in 1974, a re-naming of the group to Acoustics, Speech and Signal Processing (ASSP). Rapidly increasing activities in these topics created a significant growth in membership. Concomitantly, progress in microelectronics and integrated circuits began to make commercial applications of speech systems increasingly attractive. Very quickly, in 1976, this growth, both in membership and opportunity, qualified the group for Society status, and resulted in a re-titling to Acoustics, Speech and Signal Processing Society (ASSP Society). The expanded purview included instituting an annual technical conference----entitled International Conference on Acoustics, Speech and Signal Processing (ICASSP). This forum has enjoyed remarkable success, and endures to the present as the premier world conference for its subject fields.

As methods for representing analog signals by discrete-time data progressed, and as the speed of computing increased with constantly diminishing costs of memory, the applications of digital signal processing (or, DSP) became pervasive. And, DSP became a unified and recognized discipline. Consequently, in 1992, the ASSP Society was transformed into the broader Signal Processing Society (SPS), which it remains today. The Society sponsors a variety of specialty publications, which are among the most popularly subscribed in IEEE (note the lower right block in Figure 1). The Society also continues to sponsor the now-famous ICASSP annual conferences [Figure 2]. Contributed papers now typically exceed 1,000 in number and are presented in multiple simultaneous sessions----sometimes as many as seven parallel sessions, and with a comparable number of poster sessions. (Data for numbers of conference papers are not available for 1995-2000.)

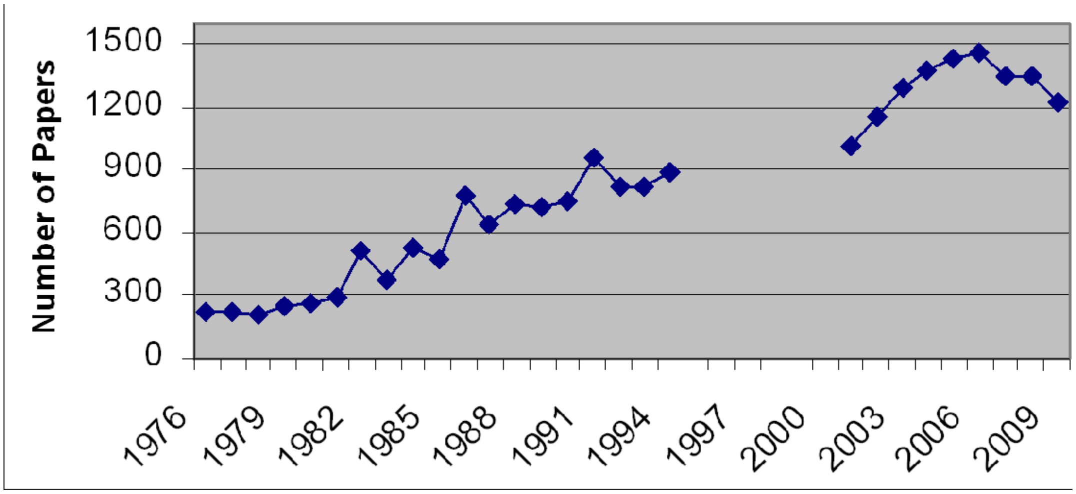

Figure 2. Number of papers accepted at IEEE ICASSP conferences.

Over the years, the ICASSP meetings have become more international, with the international sites drawing notable spurts of papers. Note in Slide 2, the "local peaks" that correspond to: 1982-Paris; 1986-Tokyo; 1991-Toronto; and, 2006-Toulouse. The most recent meeting---2009-Taiwan---drew 1,225 papers. In keeping with the growing international flavor of ICASSP, future meeting sites are projected to include: Prague, Kyoto, Vancouver, Florence, and Brisbane.

The various incarnations of SPS have correlations in the growth of membership [Figure 3]. Commencing under IRE in 1948, with a membership of about 1,000 engineers, mostly domestic, the organization and its subsequent forms grew to a peak of 20,000 in five decades. It is now composed of a broad international cadre of experts----- many of whom are seen annually at the ICASSP conferences.

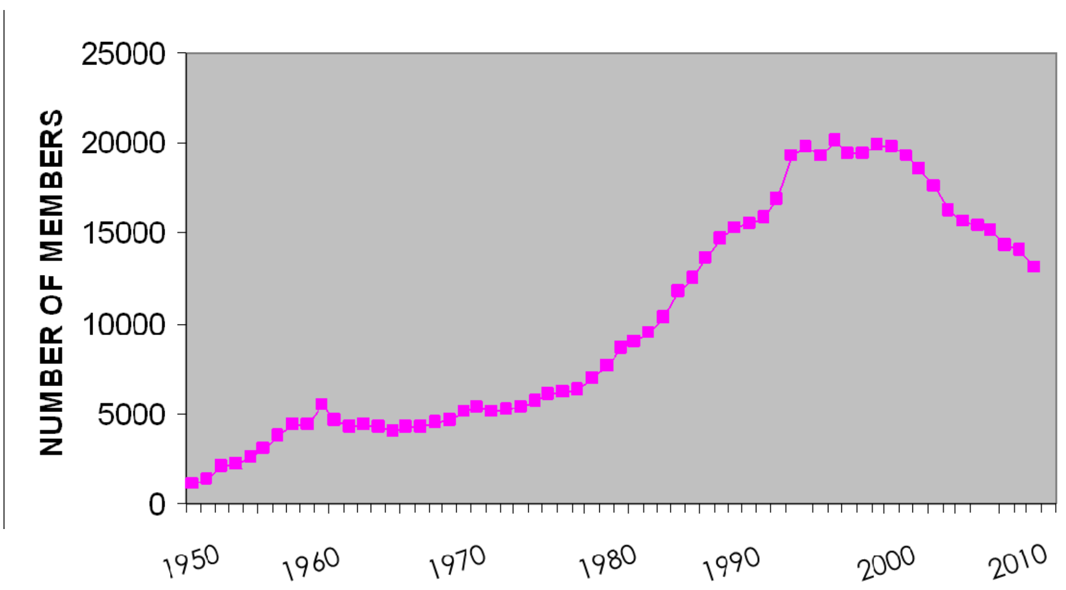

Figure 3. Membership in the IEEE Signal Processing Society.

The pattern of membership also reflects developments in technologies. Membership experienced a plateau during the first two decades (1950-1970). Here, analog audio interests dominated and matured----while the first tools for digital simulation and coding were being conceived. A driver then was bandwidth conservation in speech telephony. But, interests broadened to additionally include topics such as video coding, language processing, security and human-computer interaction. This particular shift contributed to explosive growth in membership over the next three decades (1970-2000). Note particularly the "step increase" (of about 2,000 members) just from 1991 to 1992. Several important advances were emerging about this time: the World Wide Web was created; compressed audio coding contributed the MP3 standard; large-scale commercial deployment was made of speech recognition for routing of telephone calls; video coding research generated the MPEG1 standard. And, natural environments for group teleconferencing were coming under serious study.

The last few years have shown some diminution in formal membership. Interestingly, however, no significant lessening in participation in the ICASSP international conferences has been seen. In fact, the strongest increase in meeting papers has actually occurred in the interval *after* the year 2000. One interpretation might be that the membership pattern could be influenced by two factors. First, the IEEE began to make publications available to members over the Web. While reaching a greatly expanded readership, this access tempts increased sharing of information (such as among professors and students), and may lessen the attractions of formal membership. Then, secondly, the trend may also reflect a normal life cycle of technical societies. As technologies mature, they become more "commodity-like", and there tends to be a fractionation into new directions, some of which may grow and eventually require societies of a new specialization---- as we have already seen in this account. These paths are usually difficult to predict. They depend upon the creation of new knowledge---- which is sometimes convoluted by serendipity; by response to market; and, by perceived benefits to society.

One of these new directions, I believe, is likely to focus on natural user environments for machine-mediated communication, and on multi-modal human interaction with networked information systems [Figure 4]. One would like user interfaces that approximate the ease of face-to-face communication, whether among individuals, or groups, or with machines. Such designs use, and build upon, "commodity" signal processing. Most of our communications, even now, are machine mediated. Interfaces are becoming more capable of capturing and displaying information (witness the capabilities of high-end cell phones), and the interfaces largely accommodate the senses of sight, sound, and touch. Ideally, we desire 3D realism in these sensory dimensions, but, as yet, this can only partially be realized. Mediation functions at the user site, or immediately accessible at a server, are constantly increasing in sophistication---- offering expanded computation, economical storage, and new methods for security.

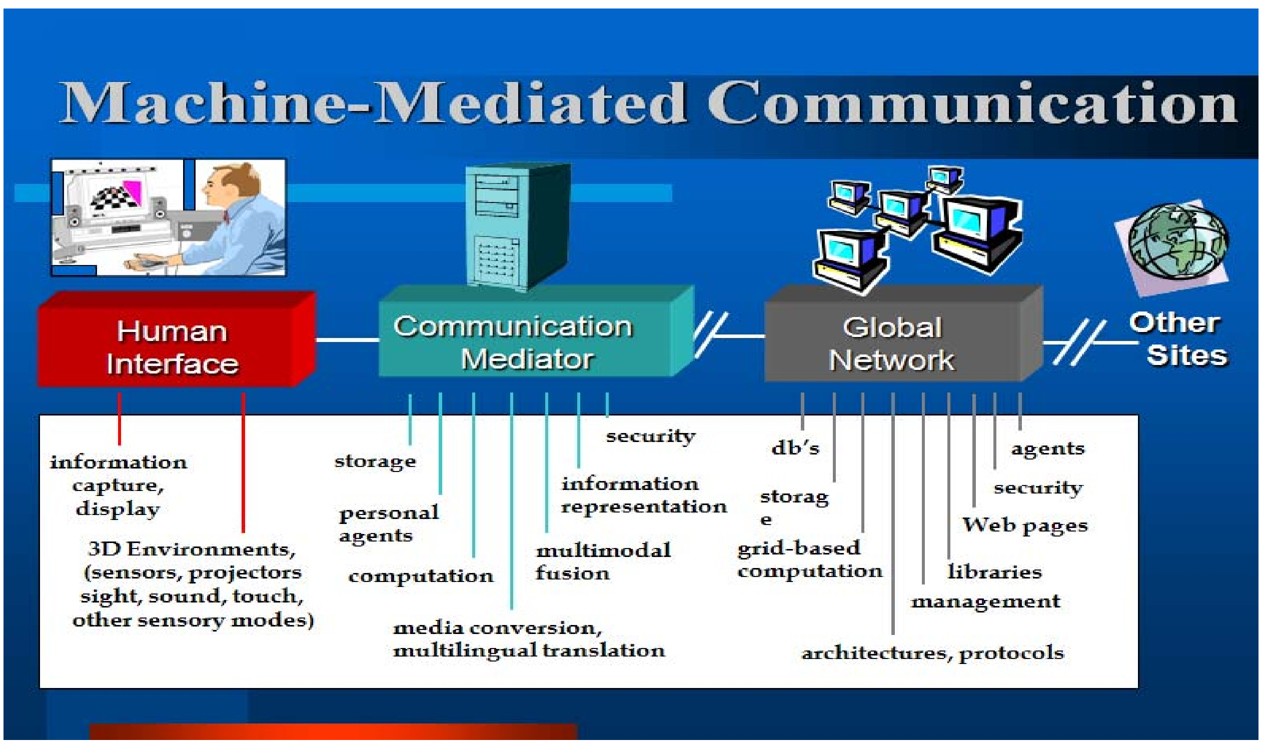

Figure 4. Machine-mediated communication.

Once inside the global network, a variety of embedded intelligence facilitates access to huge databases and libraries, to large storage, and to grid and cloud computing. But most particularly, the network provides versatile connectivity to other users or groups, or to other machines---- transcending time and geographical location.

Present research strives to approximate some of the naturalness of face-to-face exchange by so-called "multi-modal communication" [Figure 5] The approach seeks to exploit simultaneous use of sight, sound, and touch. In the particular case here, a gimbaled camera and a co-located infrared light source constitute a gaze tracker. The device locates the users face, illuminates the eye with safe infrared, and computes the angle between the corneal reflection and the centroid of the pupil. It therefore can position a cursor on the terminal screen at the point where the user is looking. An auto-directive microphone array focuses on the user's mouth, and captures high-quality speech signals----thus enabling interactive conversation by automatic speech recognition and voice synthesis. And, a snugly fitting tactile glove permits transmission of gesture, and even grasp and movement of objects in the visual display. The glove incorporates a Polhemus coil on the back of the wrist to signal hand position. It has miniature thrusters to provide force feedback to the fingers, and optical sensors to detect finger closure.

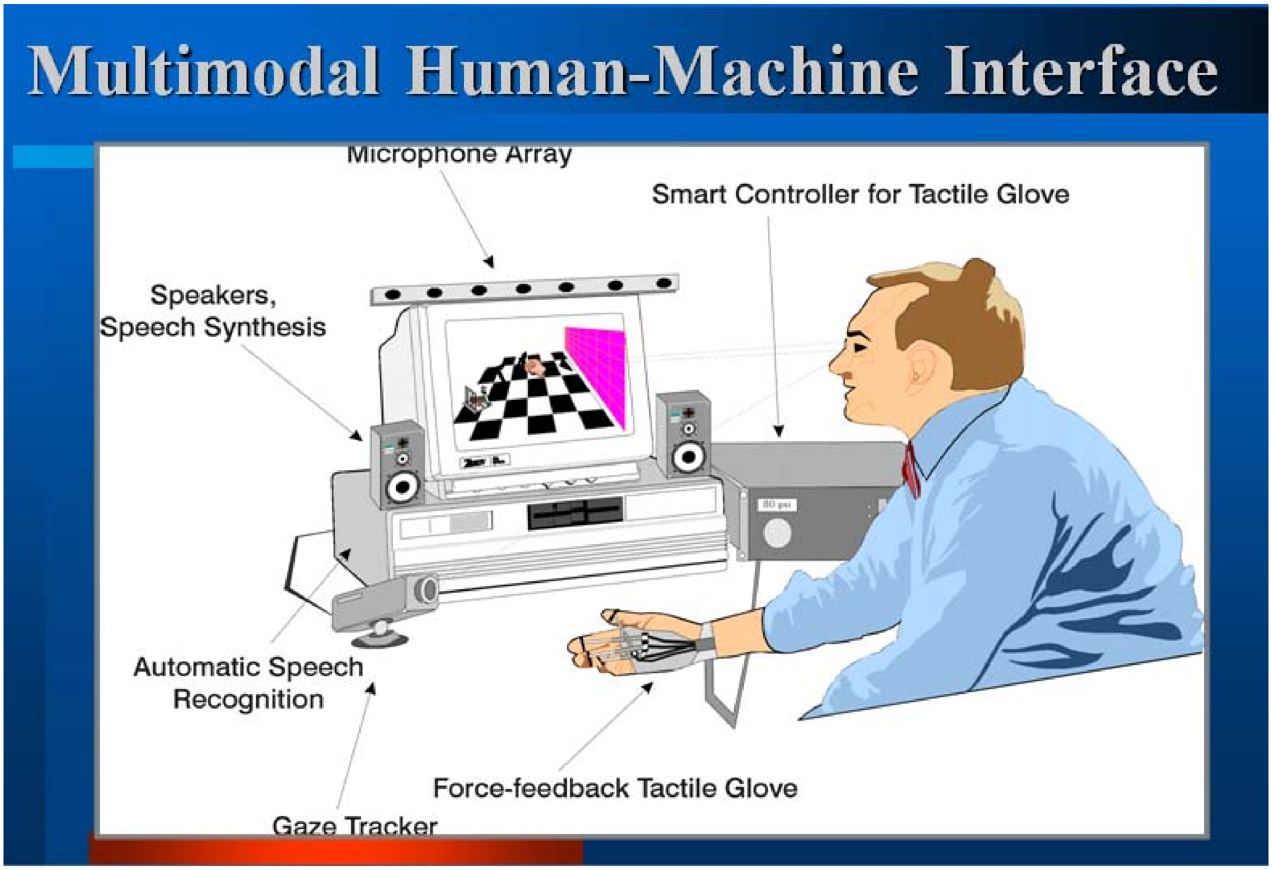

Figure 5. Multimodal human-machine interface.

While still primitive, these implementations have been created in the laboratory [Figure 6]. They have been exercised favorably in experimental collaborative tasks---- such as planning the response to a hypothetical natural disaster. As often occurs, such experiments point up a further research challenge: namely, the need for a quantitative framework for multi-modal language. We already have reasonably useful models for spoken language. For a specific language, we can establish the distinctive sounds of speech----the phonemes. We largely understand the rules for concatenating these sounds into words and the syntax for producing sentences. To a certain extent, we can extract and interpret semantic relationships over moderately long intervals. But, we have no comparable framework for communication by the simultaneous use of pointing, grasping, speaking, and looking. What are the "phonemes" for such simultaneous signaling?

Figure 6. Laboratory implementation of multimodal human-machine interface.

In these and similar multi-modal implementations, Signal Processing is pervasive. It certainly underpins machine-mediated communication. Nay, indeed, it now permeates most electrical communication and control systems. Signal Processing has truly achieved commodity status. It will, no doubt, continue to be central to electrical applications far into the foreseeable future. And, I believe, the IEEE Signal Processing Society will continue to adapt and serve these new advances---- as it has so admirably done in the past.