An Overview of Structural Design Issues in Theoretical and Practical Implementations of Analog Speech Scrambling Systems

Max Power
Power Broadcasting
HireMe.geek.nz

Spring 2003 (Original content)
Fall 2008 (Revision and Reissue)

This 2008 revision fixes numerous broken web links as well as addressing other minor latent flaws in the original content. The 2008 version is being transitioned into being a general purpose reference research paper. Changes in the document's original content have been minimised, so as to retain as much as possible of the research paper in its original form.

Thesis

Analog speech scrambling systems (as opposed to digital voice encryption systems) have existed since the mid-1920s. Analog speech scrambling technology (as opposed to more modern voice encryption technology) has a checkered past with respect to overall voice security.

Modern analog speech scrambling systems have successfully tackled the issues of voice security by using new technological methods that have been borrowed from high fidelity digital audio bitrate reduction systems, namely MPEG audio.

Modern methods of making speech secure over analog telephone and radio circuits involves digitizing speech. Manipulating digitized speech in the frequency and time domains by using digital encryption methods is the core of modern voice scrambling systems.

Digitization and encryption methods used by these secure voice scrambling systems must be step with psychoacoustic research: historically analog voice scrambling systems have been more easily broken.

Low complexity digital audio based psychoacoustical processing models and techniques coupled with low complexity digital encryption methods are adequate enough when coupled together properly to make scrambled audio extremely unintelligible (and thus secure) under normal telecommunications conditions.

Most importantly, it is possible for people with modest engineering talents and expertise to design a reasonably secure yet relatively simple analog voice scrambling system.

Extensive use of existing international telecommunications standards, coupled with common design interfaces and principals can aid in the construction of secure telecommunications systems.

Analog scrambled voice systems have been around for at least 80 years. These systems have been of strategic asset to their users by providing immunity from either casual eavesdropping or full blown surveillance as practiced by security intelligence agencies of domestic and foreign origin.

Analog voice scrambling has a permanent place in the global telecommunications infrastructure:

> http://www.tccsecure.com/voicetx.htm#analogvoice
>
> The principle disadvantage of the analog voice encryption technique is in its retention of a finite number of signal permutations. When the number of signal permutations is limited, it may be possible (with a reasonable amount of effort) to achieve some degree of success using signal analysis countermeasures.
>
> This approach requires the use of sophisticated signal analysis of the individual encrypted audio segments in an attempt to characterize each to a degree where they can be reconstructed and re-organized in their original orientation and sequential order.
>
> However, the ability to reconstruct the signal using brute force methods is very limited if sophisticated encryption techniques are used, plus the process is too slow to achieve anywhere near real time signal reconstruction.
>
> It is therefore an excellent approach for achieving a "tactical" level of voice security, and (depending on the sophistication of signal processing used) can achieve 'strategic' (long term) levels of signal protection.

Analog scrambled speech is inherently less secure (from the viewpoint of information theory) than fully digitized and encrypted speech. This is true because of the inherently massively parallel supercomputing nature of many parts of the human auditory system.

The auditory system can learn to cope with continuous inconsistencies within the human auditory range, allowing a conversation between two people to be picked up out crowd of hundreds. No human designed computer system has been able to achieve this.

Analog scrambled speech systems however are better suited for "less than ideal" telecommunications environments. Analog scrambled speech systems can go where digital encrypted voice systems cannot go, an often missed strategic advantage.

> http://www.tccsecure.com/voicetx.htm#analogvoice
>
> The analog voice encryptor can be viewed as a hybrid between a digital encryptor and a voice scrambler. It also digitizes the voice signal (often at a data rate much higher than

the typical Vocoder), but handles the voice processing in a manner that allows digital-to-analog reconstruction in a bandwidth constrained manner.

This means that although the analog voice signal is digitally processed, it retains sufficient voice-like characteristics, that when transmitted out over the channel, maintains its energy within the original voice channel.

(...)

The principle advantage of this approach is the voice quality which is typically much higher than a vocoder generated product for a given channel bandwidth. Additionally, it operates on far worse channels (noise, multipath, phase distortion, etc.) than the digital equivalent encryption system. The degree of security is to a large degree dependent on the level of signal processing and the security of the key stream generator used to set the signal processing's permutation attributes.

On one hand, it's extremely difficult to attack the key stream used, particularly if hashing functions are used (that hide the actual key stream output) and the fact that any key stream 'visibility' is very limited. As noted above in the Digital Voice Encryption discussions, this makes a break of the key stream through key analysis extremely improbable.

The first digital speech encryption systems date to WWII. The vast majority of these digital encryption systems since WWII have avoided decipherment of their cryptographical systems by all but a handful of very well funded intelligence agencies.

http://www.nsa.gov/wwii/papers/sigsaly.htm

Efforts to create a secure voice system had existed since the 1920s. Some progress had been made, but as with the A-3, no device was able to offer complete security. In the early 1940's however, the situation began to improve. Bell Telephone Laboratories, under the direction of A. B. Clark (who later headed up the research and development effort at the fledgling NSA), and assisted by British mathematician Alan Turing, began work on what would become known as "the Green Hornet."

The design of the system was based on earlier 1930s-era research on the transforming of voice signals into digital data. The device earned the nickname for the buzzing noise heard by someone attempting to eavesdrop on the conversation. The "buzz" closely resembled the theme song of the popular serial radio show of the time that went by the same title. In time, however, it acquired the more formal moniker of SIGSALY. (...)

However, digital speech encryption systems were always very finicky and by their nature not very robust nor as portable as analog systems typically have been. In the late 2000s, IP telephony came along and fixed most of the problems of portable speech encryption systems – but the fix has only worked for the telecom sector.

Although I cannot offer any concrete proof, one can easily surmise that US and UK military forces based in the former Eurasian states of the former USSR, Afghanistan, and Iraq use analog speech scrambling systems extensively. These regions have very difficult telecommunications environments.

The amount of “engineering systems analysis” these complex modern analog scrambled speech systems receive has typically been very little. Yet, as long as the way that the human brain's speech recognition functions is taken into account, designing a secure analog scrambled speech system may indeed be very simple.

Since the end of the Cold War traditional methods of analog speech scrambling have been abandoned.

The traditional methods of speech inversion that (as time went on) used more and more instances of “programmable frequency inversion” within the speech band -- made the analog circuits that performed these functions more complex over time.

There was little “real world security” gained with most of the increases of complexity.

> http://www.cescomm.co.nz/encryption/scrambling.html
>
> One of the most common scrambling methods used is frequency inversion. As the name suggests, this process takes the frequency of the input speech signal, and inverts it to provide a mirror image at a different frequency range.
>
> For transmission purposes, the human voice spectrum ranges from about 300 … 3000 Hz, and the voice signal is more powerful at lower frequencies. By inverting the signal, the relative power at each frequency level changes. The original spectrum is also repeated at a higher frequency.
>
> The inversion frequency used affects the frequency of the resultant signal.
>
> (…)
>
> This process relocates the power level of each input frequency to a new position, calculated as the difference between the original frequency and the inversion frequency.
>
> In this way, the values originally at 3000 Hz are relocated to 300 Hz [3300 – 3000 = 300] and the values originally at 300 Hz are relocated to 3000 Hz [3300 - 300 = 3000]. However, the inversion process also creates a second signal block, identical to the input signal but moved to higher frequencies.
>
> These frequencies for this higher block are calculated as the sum of the input signal and the inversion frequency (so the block starts at 300 + 3300 = 3600 Hz and ends at 3000 +

3300 = 6300 Hz).

These high frequencies are outside the radio's normal transmission range, so the signal is filtered before transmission and only the lower block is transmitted. As long as the transmitting and receiving radios process the signal using the same inversion frequency, the signal can be scrambled and recovered with good quality.

(...)

In a dynamic inversion system, both the sending and receiving radios must change their inversion frequency simultaneously and in an identical manner, so the change is generally controlled by an algorithm. The algorithm dictates the location of the new inversion frequency, and the rate at which the frequency changes.

Some scramblers use a "sweeping" technique, where the inversion frequency changes rapidly, but in a predictable manner. Another technique is pseudo-random hopping from one inversion frequency to another. This is less predictable, but the time spent at each frequency tends to be considerably longer.

The two techniques can be combined, resulting in a scrambler which sweeps across frequencies with periodic hops to a new position in the cycle. After the input signal has been scrambled the altered signal can be transmitted openly and cannot be understood by a casual eavesdropper.

In all these cases, the signal change is controlled by an algorithm, in combination with a programmed key or code. Both scramblers must be able to operate in a matching pattern, which is normally agreed at the start of the scrambled transmission. Synchronization signals may also be sent periodically during the communication, to ensure that all radios continue to work in a matching pattern.

(...)

In terms of protection, it is not simply the case that the more difference there is between the original speech and the scrambled signal, the better the protection. As with all forms of communications security, if there is a constant and predictable relation between the original message and the "secure" transmission, then this pattern can be uncovered over time and the scrambling will provide very little protection. A system with some element of unpredictability provides much better protection.

Traditional analog speech scrambling systems are mostly now toy exhibits at science museums and electrical engineering schools. Yet, analog voice scrambling systems are still around -- but in a revised form.

In the 1990s, analog speech scrambling systems abandoned their traditional analog circuit design for a hybrid digital / analog design. The only change was an overall more

systematic design that showed itself in its use of narrower audio filter bandwidths.

How do these modern voice scramblers function? Speech is digitized, put through a Fast Fourier Transform (FFT), broken into subbands, cryptographically manipulated in digital domain, and then finally inverse-FFTed into speech band audio.

All of these details may sound overly complicated but they are not. A telephony tool kit running on a PC (with a graphical user interface) can allow anyone to design and debug such a system with little difficulty or background Electrical Engineering knowledge.

Designing a secure analog scrambled speech system is not easy, and the National Security Agency's comments on the A-3 analog voice privacy system and its successor speak volumes:

> http://www.nsa.gov/wwii/papers/sigsaly_story.htm
>
> Before the full involvement of the United States in WWII, the United States and the United Kingdom were using transatlantic high-frequency radio for voice communications between senior leaders.
>
> The analog voice privacy system in use, called the "A-3," provided reasonable protection against the casual eavesdropper, but it was vulnerable to anyone with sophisticated unscrambling capability.
>
> This system continued to be used during the early part of the war, and government officials were warned that they could be overheard. In fact, it was later discovered that a German station in the Netherlands was breaking out the conversations in real time.
>
> This situation was intolerable, but neither the U.S. nor the U.K. had a ready solution. Fortunately, the technical groundwork for a solution was already in place.
>
> About 1936, Bell Telephone Laboratories (BTL) started exploring a technique to transform voice signals into digital data which could then be reconstructed (or synthesized) into intelligible voice.
>
> It was called a "vocoder," short for voice coder. An early demonstration of the voice synthesizer portion of the vocoder was even a part of the 1939 World's Fair in New York. The approaching war stimulated the investigation of true voice security.
>
> The BTL staff soon discovered that there were about eighty patents issued on the general topic, but analysis indicated that all of the methods were really unsatisfactory from a national security viewpoint.
>
> New technology was required. Spurred on by added interest from the U.K. and some

early research results, the vocoder was selected as the basis of a new high-tech voice security system. BTL proceeded on its own to develop this much-needed capability and was soon able to demonstrate it to the satisfaction of the Army.

A US Army contract was awarded in 1942 for the production of the first two systems. This system eventually came to be called SIGSALY2 and was first deployed in 1943.

If one is conceptually designing an analog speech scrambling system, it is possible to ignore initially certain aspects of the design -- microphone, etc -- and to concentrate on telephone level digitized audio as the digital input.

Before I go into a deeper explanation of the computational mechanics of scrambling analog speech, I must refer you to some of the fundamental information theory aspects of speech:

http://www.cescomm.co.nz/encryption/voice.html

Nearly fifty years have passed since Claude Shannon wrote his paper which established him as the father of modern communications. He maintained that human voice can be represented completely by as few as 500 bits per second. This means that in theory we could transmit voice digitally in as little as 250 Hz bandwidth, yet we normally allocate more than ten times that to transmit telephone quality analog voice.

Hence it is claimed that voice has very low entropy (information content) compared to the bandwidth it occupies. This in turn, leads us to ask the question "so what is voice?" In the frequency domain we can define voice as a series of harmonically related sine waves, starting with a fundamental frequency of about 100 Hz for males and about 200 Hz for females and extending throughout the 3000 Hz or so frequency range.

It is interesting to note at this point that although we 'hear' the fundamental frequency of the speaker on the telephone, the fundamental is not actually there at all, since it is not passed by the telephone circuit. This happens because the human ear fills in the fundamental frequency by listening to the harmonics. Obviously the mechanism of hearing is very complex, and even now it is still not fully understood.

Not all of these harmonically related sine waves are equal in amplitude, however. Peaks occur in each of 3 frequency regions known as 'formant' frequencies. The first formant frequency peaks at about 800 Hz in the telephone circuit, the second at about 1500 Hz, and the third at about 2800 Hz.

The first formant frequency is dominant, with the second being more than 6 db lower, and the third another 6 db lower yet. Most of the voice power is in the lower frequency vowel sounds, yet most of the information content exists in the higher frequencies.

In the time domain voice varies in amplitude with time, but even in active conversation

the gaps between words on average occupy more than 50% of the conversation time. To make the situation even more complex, the envelope of voice alone carries most of the information.

Consider the case of whispered voice. A whisper is composed of random noise, shaped in amplitude by the voice mechanism. But a whisper is still understandable.

This is because most voice information is not carried in the harmonically related frequencies, but in the envelope of voice itself. In fact, most voice information is carried in the fricative, or short, sharp sounds.

At the same time, hearing the lower frequencies or vowel sounds gives the illusion of intelligibility. This is easily observed by filtering out the higher frequencies, leaving only the low frequency vowel sounds and observing the results. The listener feels that he is hearing and even recognizing the speaker, yet he cannot repeat the words being spoken.

Confused? Shannon wouldn't be. He would also have told us that, despite our advancements in technology over the intervening years, the successful securing of analog voice still presents an enormous challenge.

Even to date, this challenge has not been met very well by voice scramblers, hence their rather poor reputation. The complexity of analog voice, as briefly described here, demand a very sophisticated technology to successfully secure it.

Let us return to the essential details of our design… We can assume that our input is telephone audio digitized at 8000 samples / second with 8 bits of resolution. This sampling rate is based on ITU (International Telecommunications Union) standards.

This sampling rate is adequate for medium quality speech, but not for digital audio like on a CD nor in a digital radio broadcasting application. Typically one would use 16000 samples per second (at 10 or 12 bits per sample) for this kind of application to increase systemic fidelity.

A modern Digital Signal Processing "Unit" (aka DSP) or CPU with of 10 bits or more of internal processing resolution can take these 8000 samples and instantaneously provide an FFT of 1/10 of a second of this audio (800 samples) in 10 milliseconds or less. Once the FFT is completed, a whole psycho-perceptual can of worms can be opened.

In other words our design problems now get messy very quickly. Let us assume for simplicity that the FFT audio sample is converted into an MPEG-Layer II (MPEG II) audio sample.

MPEG (Motion Picture Experts Group) is intimately linked to the IEEE (Institute

of Electrical and Electronics Engineers) and ACM (Association for Computing Machinery) and the AES (Audio Engineering Society). These standards bodies are involved with the treatment and processing of audio with respect to digital audio transmission and storage systems.

Each MPEG formatted audio sample has been bandwidth limited to the telephone range of 300 Hz to 3300 Hz.

What is this MPEG sample? It is now a vector and magnitude dataset that describes 1/10 of a second of speech using approximately 16 individual bands of audio divided into the range. Each audio subband in this system contains about 180 Hz of speech spectrum (180 Hz ~= 6% of 3000 Hz).

Now what does one do with this band filtered digital audio? The answer is: you have to scramble it based on psychoacoustic and auditory principals.

We must ask ourselves: what psycho-acoustical goals need to be reached with respect to our scrambled audio for our scrambling to be successful?

Here is a summary of what elements of speech must be rendered meaningless by the scrambling process:

- f0, or the primary frequency emitted by the voice box, must be lost
- f1 and to a lesser degree f2 must be lost as well; f1 and f2 provide clues about the state of f0 -- even though f1 and f2 they are typically 6 db and 12 db below the spectral intensity of f0
- the scrambling must work with a wide variety of speakers: f0, f1, f2 never fall in the same place even with the same speaker during a conversation
- formants and formant transitions must not be discernable
- vowels (including high, middle, and low vowels and front, centre, and back vowels) must be muddied
- consonants (hard and soft) must be spectrally scattered, noting that hard consonants provide valuable clues to speech content across many transmission mediums
- co-articulation must be spectrally scattered as much as possible
- there must not be sufficient clues about speech content over a prolonged time frame – typically more than 5 minutes; this is a crypto keying issue which requires changing the key often
- the audio quality of descrambled speech must have the same perceptual quality as normal speech over the same communication channel

Doing the above spectral manipulations in the cryptographical domain may sound hopelessly complex, but in reality this is far from the truth.

Most keying algorithms for scrambled speech are based on slowly changing keys or else on fixed keys [that may or may not be changed at regular intervals]. Almost

universally robust (error free) self-correcting cryptographical algorithms are preferred for this kind of application.

There are about 12 common cryptographical keying systems specifically designed for scrambled speech applications. However, the world we live in is not perfect with respect to using digital keys within an analog technology:

> http://www.tccsecure.com/voicetx.htm#analogvoice
>
> Early methods of "analog" encryption were nothing more than voice scramblers with little security to any aggressive attack. The advent of more powerful voice processing circuitry and software allowed for more sophisticated voice processing techniques which use a key generator's secure key stream for selecting the given sound segment's permutations.
>
> These permutations include band segmentation, subband frequency inversions (or non-inversions), and subband segment interleaving. The more combinations used the harder to reconstruct the signal without knowledge of the key generator's key stream.
>
> This technique will generally provide a near-plain mode level of voice quality while containing the encrypted channel to within the plain modes voice channel bandwidth. It is common in the newer 'analog' techniques to digitize the signal, but it processes (in many respects) like an analog signal. In this respect, it is a bit of a misnomer to call it 'analog' encryption; however it is done primarily to differentiate it from 'digital' voice encryption techniques.

With our theoretical voice scrambling system, we can assume that just by changing the key once every 6 seconds (an action consisting of altering the order of the subbands [example: (1,2,3,4,5,6,7,8,9) --> (1,8,2,7,3,6,2,5,4,9)] based on the output of a keying algorithm) *that we will achieve adequate signal security*. Six seconds is about the size window for short term auditory memory, so key changes less than this time constant are of paramount consideration.

More or less the only questions that remain to be solved by your system design are: key frequency change, crypto key synchronization and crypto key exchange using analog modem technologies. For the moment we will ignore the cryptographical key exchange issues.

This cryptograpical keying idea is based on the physiological premise that humans have a window of about 10 seconds of echoic memory. Parametric audio subband changes in this temporal window could selectively keep the hearing system from being able to lock onto the predictable coherencies within speech spectra.

- As long as the order of the subbands is randomized on a short term basis subband swapping is quite effective against most forms of casual eavesdropping.
- Usually f0, f1, most formants, consonants, and coarticulations are easily destroyed beyond recognition by this method.

It is possible to enhance this method of scrambling, as there are many inherent weaknesses with this kind of simplicity.

There is also the alternate option of time division subband scrambling the speech band audio, albeit this process has many different names. This added security feature creates an issue of time delay and synchronization with respect to the electronic decoder circuit at the other end. The encoding / decoding delay can be increased to 1/2 second with this secondary scrambling option.

There is a human perceptual cost for this time delay, or any time delay longer than 100ms for that matter. This increased encoding / decoding delay time may make holding a sensible conversation over satellite telephone circuits (where 1/2 second time delay is normal) more difficult or almost impossible.

> You can never assume that your secure telephone conversation will not be relayed by satellite. Even 4 hop shortwave links have 110ms delays. Delays are the bane of the telecommunication engineer.

After the final stage of spectral subband disordering, the MPEG audio encrypted dataset is sent to another CPU chip to be converted back into normal audio for transmission.

Additional analog filtering components may be added to the outgoing audio, so as to eliminate square waves that may cause distortion. Distortion may cause the scrambling / descrambling units to loose synchronization with each other, as well as lower the perceptual quality of the audio for the listener.

Assuming you have an ideal encoder and decoder, the simplest of digital speech scrambling systems should work flawlessly. However, each scrambling system must go through perhaps 3 - 6 months of field trials to work out any design flaws.

Many commercial voice scrambling systems are the end product of 5 years or more of research and have a much higher price to reflect this fact.

Most modern digital and analog speech scrambling systems can fail under less than ideal conditions. My research has only uncovered one analog system robust enough for use on HF (shortwave) communications links.

The theoretical analog speech scrambling system presented here would probably not work very well on shortwave or on low quality telephone lines. Almost universally modern voice scrambling systems have a synchronization subsystem that uses Radioteletype (RTTY) level signals.

Scrambled speech systems are sold to both commercial and military users, but the telephone features are generally identical for both. Military secure voice systems typically have only minor alterations made in their encryption subsystems (and different product design packaging) versus the systems sold to the civilian market.

The minor alterations in the cryptographical subsystems for military users are generally required to keep commercial units from being used to break into or interface with autonomous military voice traffic networks.

This simplification of voice scrambling subsystems has purposefully skipped over the analog synchronization features needed to keep the scrambled speech units locked onto the same cipher key, and each other.

There is currently only one voice scrambling system on the market that does not need to use an analog band synchronization signal. It goes without saying that this theoretical voice scrambling system is quite primitive by commercial standards.

One must remember that when individuals or institutions buy secure telephone systems, these are their primary concerns:

- What level of security is really needed?
- What device type best suits the specific application?
- What level of device reliability and supportability is needed?
- What key management approach is desired for the application?
- Is key management really needed for the application?
- What cost constraints exist to limit product selection?

Here is a brief excerpt from a product brochure (Codan: 12-20148-EN):

> The Codan Voice Encryptor option is specifically designed for the needs of business, government, and non-government organizations, which require safe and secure communications when carrying out vital operations. Such operations may include anti-narcotic operations, international truck convoys, and humanitarian aid organizations.
>
> (…)
>
> The Codan Voice Encryptor uses the unique patented SAFE encryption technology, which does not require synchronization and allows communication to be conducted successfully and efficiently even across poor channel conditions.

What are speech scrambling systems up against?

The modern voice scrambling system must evade Personal Computers (PCs) with audio analysis software. Modern audio analysis software possesses the ability to manipulate the audio signal in a programmable manner. An off-the-shelf PC can easily be upgraded to decode some kinds of scrambled speech.

Most analog scrambling systems are ahead of the technological curve, but it is really impossible to predict how long this technological edge will last. PC audio analysis software does not consume very much CPU time on a multi-core CPU, because modern 32 (and 64) bit PCs typically run at 1.0 Gigahertz or more.

All that the audio processing software must do is to render the scrambled audio into some form that is readily usable by the human auditory system. All the audio processing software must do is alter the scrambled audio just enough to recover the secured conversation.

Conclusion

It is possible to design a secure telephone communications system using hybrid digital-analog scrambling technology with readily available telephony and computer components and systems. Said system can be reasonably secure from compromise from the majority of opponents that may eavesdrop on the telephone or radio circuit.

For said system to be viable the designers must not only take into account the basic needs of securing speech signal, but also the user interface and electronic standardization issues that may compromise system security.

Such a design easily falls within the intellectual abilities of most telecommunication system designers, but the design must be subject to step by step evaluation from acoustical and linguistic experts to ensure the highest level of security, because the human auditory system is so highly adaptive.

Bibliography

http://www.codan.com.au/
Codan is an Australian manufacturer of telecommunications equipment that happens to implement SAFE voice scrambling technology. It uses the unique patented SAFE encryption technology, which does not require synchronization.

http://www.tccsecure.com/
For over 35 years TCC has specialized in securing communications networks. TCC's products have been installed in over 100 countries and on six continents. Companies like TCC are typically called on to secure communications networks in difficult environments.

http://www.itu.int/
The Union was established last century as an impartial, international organization within which governments and the private sector could work together to coordinate the operation of telecommunication networks and services and advances the development of communications technology.

http://www.ieee.org/
The IEEE and its predecessors, the AIEE (American Institute of Electrical Engineers) and the IRE (Institute of Radio Engineers), date to 1884. From its earliest origins, the IEEE has advanced the theory and application of electro-technology and allied sciences, served as a catalyst for technological innovation and supported the needs of its members through a wide variety of programs and services.

http://mpeg.telecomitalialab.com/
The Moving Picture Experts Group (MPEG) is a technical working group in charge of the development of international standards for compression, decompression, processing, and coded representation of moving pictures, audio and their combination. MPEG usually holds three meetings a year. These comprise plenary meetings and subgroup meetings on Requirements, Systems, Multimedia Description Schemes, Video, Audio, Synthetic Natural Hybrid Coding, Test, Implementation Studies and Liaison. MPEG meetings are attended by over 300 experts from over 20 countries.