Speech Communication 7 (1988) 113–123
North-Holland

PAN-EUROPEAN SPEECH CODING STANDARD FOR DIGITAL MOBILE RADIO

Jon E. NATVIG
Norwegian Telecommunications Administration Research Establishment, N-2007 Kjeller, Norway

Received 13 October 1987

Abstract. A new Pan-European digital cellular mobile radio system has been under planning since 1982 within CEPT GSM (Groupe Speciale Mobile). The paper presents the progress of the studies that were initiated by the joint Speech coding Experts Group created by CEPT TR3 and COST 207 to define an algorithm to encode the speech at around 16 kbit/s for this purpose.

In a coordinated evaluation programme, 6 candidate codecs were compared with respect to speech quality, transmission delay and ease of implementation. As a result of this evaluation, a Regular Pulse Excitation/Long-Term Prediction LPC (RPE-LTP) coding algorithm has been selected as the basis for the standard for the Pan-European cellular system.

Zusammenfassung. Ein neues zellulares, digitales Mobilfunksystem, welches den Gesamt-Europäischen Raum umfassen soll, ist im Planungsstadium seit 1982. Die Planung wird abgewickelt innerhalb der Spezialgruppe Mobilsysteme (CEPT GSM). Der Beitrag zeigt den Fortgang der Arbeiten auf, die initiiert wurden durch eine Gruppe von Sprachkodierungs-Experten; diese Gruppe wurde durch CEPT/TR3 und die "COST Action 207" gemeinsam gegründet mit dem Ziel, einen Sprachkodier-Algorithmus im Bereich von 16 kbit/s zu definieren.

Im Rahmen eines koordinierten Testprogrammes, wurden sechs Kodierer verglichen was ihre Sprachqualität, Verzögerung durch die Signalverarbeitung und Einsetzbarkeit betraf. Aufgrund der Resultate wurde ein LPC-Algorithmus mit Regulärpuls-Erregung und Langzeit-Prädiktion als Standard für das europaweite Zellularsystem ausgewählt.

Résumé. Un nouveau système de radio mobile cellulaire et numérique à couverture européenne a été planifié depuis 1982 à l'intérieur du CEPT GSM (groupe spécial mobile). L'article présente la progression des études qui ont été initiées conjointement par le groupe d'experts en codage de la parole créé par CEPT TR3 et le projet COST 207 pour définir un algorithme afin de coder la parole à environ 16 kbit/s.

Dans un programme coordonné d'évaluation, six codeurs candidats ont été comparés sur base de la qualité de la parole, du retard de la transmission et de la facilité d'implantation. A la suite du résultat de cette évaluation, l'algorithme d'excitation à Impulsions Régulières et de Prédiction LPC à Long Terme (RPE-LPC) a été choisi comme base pour les standards du système cellulaire pan-européen.

1. Introduction

Automatic cellular mobile radio in Europe has experienced a tremendous growth in the recent years. In many countries in Europe the projection of the demand for this service has indicated the need for a second generation system in the early nineties.

Until now, the European countries have established separate and incompatible systems on a national basis. However, considering the prospects of mass production of equipment resulting in lower costs and the possibilities for a more attractive service in a common system, the European Telecommunications Administrations created in 1982 the working group CEPT-CCH-GSM (Groupe Speciale Mobile). The task given to GSM was to study and develop a standard for a future Pan European mobile system to be introduced in 1991/92. At the same time CEPT recommended to reserve the frequency bands 890–915 MHz and 935–960 MHz for this purpose through-out Europe.

At an early stage in the development, the working assumption was established that the system would be based on digital transmission of voice and data and that it would, as far as possible, extend the functions and services offered by the integrated services digital network (ISDN) into the mobile sector. The following basic system requirements have been established by GSM:
- The system should allow automatic international roaming. This means that the system will keep track of the mobile subscribers locations so that calls from the fixed network can be routed to them directly with no need for the fixed telephone subscriber to know the location of the mobile subscriber.
- No significant changes to the fixed telephone network or the existing infrastructure for cellular mobile services should be required by the new system.
- The system should be compatible with the evolving ISDN network and should support as many ISDN services as possible, including the possibility to offer encryption of the user information on the radio path. The system will support car mounted as well as handheld mobile units.
- The average speech quality perceived by subscribers should be equal to or better than the quality in 900 MHz analogue mobile systems of today.
- The system should be as spectrum effective as possible, the aim being to offer a capacity superior to that of existing systems.

According to the current action plan, outline versions of the recommendations will be available by the end of 1987, and stable versions are expected at the end of 1988.

To establish the characteristics of the system, a number of research programmes have been carried out in the European countries. A number of scenarios have been studied for the modulation and access techniques as well as for the speech coding algorithm and in order to coordinate and evaluate these technical studies, GSM has established a number of specialised sub-working groups within the CEPT organisation.

In October 1985, the Speech Coding Experts Group (SCEG) was formally created as a joint group between the CEPT sub-group TR/SG3 and the COST project 207, a Pan-European research programme in the field of digital mobile telephony. This paper is based on the results obtained by this group.

Taking into account the deadlines defined by GSM, a selection procedure consisting of three distinct phases was established by the SCEG. In the first phase, the main task was to establish the design requirements and test methodology for the comparison of candidate coding algorithms.

In the second phase, starting in September 1986, 6 selected codecs were assessed in a coordinated experiment [1]. In addition to speech quality, the transmission delay and implementation aspects were evaluated for each of the candidates.

On the basis of the results from phase two, GSM decided in early 1987 to select one basic scheme for further optimization in a third phase which is presently going on (June 1987). The result of this optimization was the selection of a simplified Regular Pulse Excitation LPC algorithm with long-term Prediction (RPE-LTP) for the European DMR system. Continued work in phase 3 will include detailed optimization of a channel coding scheme for error protection, hardware validation experiments, subjective characterisation experiments, and the detailed specification of the algorithm.

2. Performance objective for the GSM system speech coder

During the course of the studies a number of basic design objectives have been identified for the GSM speech coder:

2.1. *Speech quality*

The basic requirements for speech performance is that from the subscriber's point of view, the average quality for voice telephony should be at least as good as that achieved by the first generation 900 MHz analogue systems over the range of practical operating conditions.

Medium and low bit rate codecs will have their performance optimized for speech, and it was foreseen that certain schemes might have problems with other signals normally encountered in

public telephony. Therefore, the following specific requirements have been identified:
- The coding algorithm should be robust to variations in voice spectra and levels. A wide range of signal spectra can be expected due to variation in talkers, microphones and transmission effects of the telephone network. Furthermore, the average speech levels in the telephone network are known to vary over a range of about 30 dB.
- The codec should be robust to environmental noise and multiple voice signals. The mobile units have to operate in a variety of locations. In particular noisy environments like moving cars have to be considered. Multiple voices can occur as background "noise" as well as originating from loudspeaking telephones.
- The tandeming of two coder/decoder combinations is envisaged in the case of a mobile to mobile connection. In this case, a certain degradation in quality could be accepted, however, intelligible communication should be possible in all cases.

2.2. Bit rate

The sampling rate will be 8 kHz to allow simple interfacing to the Public Switched Telephone Network (PSTN). Based on a preliminary analysis of the conflicting requirements of spectrum efficiency and speech quality, a bit rate around 16 kbit/s was taken as a working assumption at an early stage.

2.3. Transcoding

The basic speech coding standard for the GSM system will probably be defined as a transcoder between 13 bit uniform PCM (8 kHz sampling rate) and the bitstream of about 16 kbit/s delivered to the radio system.

On the network side the GSM speech coder will have to provide transcoding to and from A- or μ-law PCM. The conversion between 64 kbit/s PCM according to CCITT G.711 and uniform PCM has already been fully defined in CCITT Recommendation G.721. This conversion will result in 13 bit uniform PCM in the case of A-law PCM.

In the mobile station the manufacturers will have the choice between readily available A/μ-law codecs followed by the conversion to uniform 13 bit PCM, or they may use a linear A/D–D/A to provide the uniform 13 bit format directly.

The linear PCM interface will probably form the basis for transcoding to other standards, since the uniform PCM level is likely to be included in any low rate encoding method. Transcoding to 32 kbit/s is already defined in principle through the existence of the 13 bit uniform interface in both algorithms.

2.4. Transmission of non-voice signals

No voice-band data requirements have been specified for the speech codec. However, it is required that the speech coder should be capable of transmitting the audio tones provided to the subscriber by the network, such as dialing tones, ringing tone, busy tone etc.

The GSM system will support a number of digital data services as well as the most common voice band data services. Considering that the design of medium and low bit rate codecs rely on the exploitation of basic properties of the speech signal, it must be expected that a compromise coding scheme to accommodate both speech and voice-band data would result in reduced voice quality. Therefore, in the speech coder design, priority has been given to voice performance and it has been decided that special terminal adapters will be implemented to transmit voice-band data.

2.5. Transmission delay

Because of impedance mismatch at the 4- to 2-wire conversions at the extremities of a telephone connection, reflections will occur. In normal national and short-range international calls, these reflections are sufficiently attenuated in the network, and will not cause any problems. However, if the one-way delay of the connection exceeds about 50 ms, the attenuation in the network is not sufficient and echoes may cause serious disturbances to customers if not removed by other means. In such cases, echo control devices (echo cancellers or suppressors) have to be utilised. The overall delay in a DMR system will have two main contributions:

Speech coder delay

To achieve an acceptable speech quality at bit rates below 16 kbit/s, block-adaptive coding schemes have normally to be used. Consequently, the transmission delay expected in medium and low bit rate coders may be substantial, depending on the block size, the complexity of the algorithm, the processing speed of the DSP etc.

Delay of the radio-subsystem

To combat the effect of burst errors, bit interleaving techniques have to be used by the radio system. This is a technique which requires that the transmission bits are buffered and transmitted in a new sequence to spread the errors. The delay introduced in this way may be several speech coder frames.

In order to eliminate excessive delay alternatives, an upper limit of 65 ms has been fixed for each of these components.

The delay introduced by the DMR system has to be added to any delay in the telephone network. Therefore, the situation may arise that echo control may be needed in DMR connections that would otherwise not need it. This situation is difficult to detect in the PSTN, and it is therefore envisaged that the DMR system will have to provide its own echo control to prevent unacceptable performance due to echo.

2.6. Ease of implementation

The implementation requirements for the speech coder derive from the basic requirement that hand-held portables should be supported by the system. In a hand-portable unit, maximum values for size and weight have to be fixed as a compromise between human factors considerations and the operational duration of a handheld unit. The operational duration is highly dependent on the battery capacity and the transmitting power, and the battery capacity will of course be limited by the weight limitation.

The implementation requirements for the speech coder are:
- implementation on one single VLSI chip;
- minimum power consumption.

3. Selection of a speech coder for the GSM system

Initially, more than 20 speech codec developments directed toward DMR applications were reported in the participating countries. After an initial screening on a national basis, 6 candidate algorithms were presented to the European Speech Coding Expert group for consideration.

In order to be able to evaluate factors such as transmission delay and implementation aspects, all candidate codecs had to be presented as hardware laboratory models operating in real time. By this requirement a certain maturity of the schemes could also be ensured. For comparison purposes, a common gross bit rate of 16 kbit/s was agreed. Within this bit rate, codec designers had the liberty to allocate the bits to speech coding or channel coding as they deemed necessary to achieve a satisfactory average performance with Bit Error Rates (BER) up to 1%.

The 6 candidate codecs can be grouped into two classes: Sub-band coders and pulse excited coders. A brief description is given below.

Sub-band coders

Four sub-band coders were presented:
- Three variants of sub-band coders with block-adaptive PCM coding of the sub-band signal (SBC-APCM), and
- one sub-band coder using backward-adaptive ADPCM coding of sub-band signal (SBC-ADPCM).

The main parameters of the sub-band coders have been summarized in Table 1.

All 3 SBC-APCM schemes [2, 3, 4] can be regarded as variants of the following basic scheme:

The signal is split into a number of (8 or 16) sub-bands, a sub-set of these (6 or 14 respectively) are transmitted. Forward estimation and adaption on a frame by frame basis is used. The side information consist of block power information and the spectral information represented by the sub-band signal levels (i.e., 6 or 14 values). The sub-band signals are quantized by individual Max-type (pdf optimized) quantizers. The number of bits to be used in the coding of the signal samples in each sub-band is determined from frame to frame by an adaptive scheme based on the distribution

Table 1
Algorithm parameters of sub-band coders

Coder	A	B	C	D
Scheme	SBC-APCM[1]	SBC-APCM[2]	SBC-APCM[3]	SBC-ADPCM[4]
Bitrate budget				
Sub-band signals	14 kbit/s	14 kbit/s	10 kbit/s	15 kbit/s
Side information	1 kbit/s	1 kbit/s	3 kbit/s	0 kbit/s
Net bit rate	15 kbit/s	15 kbit/s	13 kbit/s	15 kbit/s
Other (sync etc)	–	–	–	1 kbit/s
Error protection	1 kbit/s	1 kbit/s	3 kbit/s	0 kbit/s
FEC scheme	Golay + Hamming	BCH truncated	Extended Golay	–
Frame length	15 ms	20 ms	16 ms	1 ms
Filter bank data				
No. of sub-bands total/used	8/6	16/14	16/14	8/6
Filter type	QMF tree structure	Parallell	Parallell	QMF tree structure
Filter taps	32, 24, 16	64	80	32, 16, 12
Sub-band signal coding				
Quantizer type	Max	Max	Max	Forward adaptive ADPCM
Bit allocation	Adaptive 1–4 bits	Adaptive 0, 2–5 bits	Adaptive 0–5 bits	Band 1–6: 4, 3, 2, 2, 2, 2 bits
Coding of side information				
Spectral representation	vector quantization	vector quantization	logarithmic PCM	–
Code book size	128 × 6 bits	128 × 6 bits + 64 × 8 bits	14 × 3 bits –	–
Block gain	7 bits A-law PCM	4 + 4 bits Log PCM	6 bit Log PCM	–

Sources
[1] Centro Studi e Laboratori Telecommunicazioni (CSELT).
[2] Electronics Laboratory (ELAB), Technical University of Trondheim, Norway.
[3] Ellemtel Utveckling AB, Sweden.
[4] British Telecom, United Kingdom.

of the sub-band levels, to improve the perceived quality.

In the SBC-ADPCM, adaptive differential PCM [5] is used for the coding of each sub-band signal. 8 sub-bands are used, 6 are transmitted. In the four lower bands, backward-adaptive quantizers and predictors are used. In the two upper bands only backward-adaptive quantization is used. The number of bits for the quantization of each band is fixed. Consequently, the coder does not rely on transmission of any side information. In the rightmost column in Table 1 the algorithm parameters of the SBC-ADPCM are summarized.

Pulse excited coders

Two LPC-based codecs participated in the tests:
– a simplified regular-pulse excited LPC codec (RPE-LPC) [6].
– a multipulse excited codec with long-term prediction (MPE-LTP) [7].

The main characteristic of these coders are

Table 2
Algorithm parameters of pulse-excited coders

Coder	E	F
Scheme	RPE-LPC[5]	MPE-LTP[6]
Bit rate budget		
Net bit rate	14.77 kbit/s	13.2 kbit/s
Error protection	1.23 kbit/s	2.8 kbit/s
FEC scheme	Reed–Solomon	Extended Hamming
Framing		
Speech frame	19.5 ms	20 ms
Window	25 ms Hamming, overlapping	20 ms Rectangular, no overlapping
LPC-analysis		
Filter order	12	8
Algorithm	Schur	Le Roux-Gueguen
Coef. coding	52 bits	28 bits
Inverse filtering	Lattice type	Direct form
Pulse modelling		
Method	Regular-pulse (simplified procedure)	Multi-pulse with long-term prediction
No. of pulses/frame	52	24

Sources
[5] Philips Kommunikations Industrie AG, Nürnberg, Fed. Rep. Germany.
[6] IBM laboratory, La Gaude, France.

listed in Table 2.

There are no major differences between the two codecs concerning the following basic functions:
- Framing,
- LPC-analysis, and
- inverse filtering.

The main difference appears in the pulse modelling and quantization methods.

The MPE algorithm assumes the representation of the residual signal samples by a small number of pulses, the positions and amplitudes of which are optimized to minimize a weighted error criterion. This algorithm also includes a long-term prediction loop around the quantizers to increase coding efficiency.

The RPE algorithm assumes that the residual signal is represented by a pulse train, chosen with a reduced set of positions, but with more pulses than the MPE scheme.

4. Results of evaluations

The candidate speech codecs have been compared with respect to the following criteria:
- Subjective opinion ratings under different conditions;
- informal evaluations of effects of multiple voices and lorry noise;
- transmission delay;
- complexity and implementation aspects.

4.1. Subjective quality

4.1.1. Test conditions

The basic speech quality requirement for the speech coder has been formulated as an average taken over the range of practical operating conditions (Section 2).

An experiment was designed to evaluate this requirement, by selecting test conditions to cover the range of conditions considered to be essential. Furthermore, test conditions from an existing companded FM system were included in the subjective tests. In this way, it was possible to evaluate the requirement by direct comparison of an average MOS for each codec computed over all test conditions to the corresponding average obtained for this analogue reference system. Based on the performance requirements in Section 2, the following test conditions were selected:

Input levels: 12, 22, 32 dB below overload of the coder

A range of 20 dB was chosen to represent the practical range of speech input levels that the codecs would experience in an operational environment taking into account the variations in vocal level between individuals as well as the level distributions to be expected from the PSTN.

Bit error probabilities: 0, 10^{-3} and 10^{-2}.

The tests were conducted in an early stage in the planning process, and consequently the radio and access techniques were still under study. Therefore, no reliable estimate of the burst error patterns to be expected at the speech coder interface could be found. For simplicity, it was agreed to assess the codecs using randomly distributed errors, the assumption being that if a speech coder performs well with random errors, then it would perform well also with burst errors, provided a suitable channel coding adapted to the error statistics is provided.

Transcodings: 1 and 2
The condition of two codecs in tandem represents the situation of a mobile-to-mobile call. In this condition, the transcoding was performed with interconnection on the analog level, i.e., with full A/D–D/A in both codecs.

The companded FM conditions were:
– Carrier-to-noise (C/N) ratios 18 and 26 dB combined with simulated Rayleigh fading corresponding to a vehicle speed of 10 m/s.
– A C/I value of 18 dB is a commonly used limiting value for planning of analogue cellular systems. This condition is normally assumed to be exceeded for 90% of the locations in the cell. The corresponding limiting value for the digital system has been assumed to be a BER of 10^{-2}. Similarly, the value C/I = 26 dB is assumed to be exceeded for 50% of locations in an analogue system whereas the corresponding limit in the digital case was taken to be BER = 10^{-3}.

4.1.2. Results of subjective experiments
To provide the information needed for the selection, an international subjective experiment was carried out. In a joint laboratory session, hosted by CSELT (Italy), test recordings were made in a number of different languages. Subjective tests were then conducted with subjects listening to their native language [1]. The following laboratories participated in the experiment: British Telecom (United Kingdom), CNET (France), CSELT (Italy), Dr. Neher Laboratories (Netherlands), FTZ (Germany) and the Norwegian and Swedish Telecommunications Administrations.

On the basis of the results, global scores (averaged over all experiments and conditions) have

Table 3
Average speech quality – global results

Codec		Average mean opinion score
A	SBC-APCM	2.98
B	SBC-APCM	2.46
C	SBC-APCM	3.14
D	SBC-ADPCM	2.92
E	RPE-LPC	3.54
F	MPE-LTP	3.27
	Companded FM	1.95

been calculated for all six codecs in the experiment, as well as for the analogue FM system. The average is based on 10 representative test conditions, assessed with about 20 talkers and 80 listeners. The differences seen in Table 3 are therefore considered to be statistically highly significant.

Table 3 shows that all codecs on the average performed clearly better than the analogue reference FM system. It is observed that the average score obtained by this system corresponds to the scale value "poor". This is clearly not a correct description of the subscribers impression of existing mobile systems. A probable explanation for the low score obtained by this system may be that the subjects may have had little or no experience with the degradations to expect in a mobile call. Consequently, it is probable that they have made their judgement with reference to their experience with telephone connections resulting in relatively low scores. The scores in Table 3 are however believed to reflect a correct rank order, although the absolute scores should be treated with some caution.

The experiment did not include any "ideal" condition for the companded FM system to correspond to the BER = 0 condition of the digital codecs. However, previous results and experience [4] have shown that the speech quality in the analogue FM system saturates at a C/N value of about 35 dB, which would represent the "ideal" condition for companded FM. A reasonable estimate for the ideal condition in this experiment was found by linear extrapolation from the values obtained at C/N = 18 and 26 dB in the experiment. This extrapolated value was included in the average score.

4.2. Transmission delay

The candidates were implemented as hardware equipments operating in real time, and the delay of each coder could have been measured directly. However, being prototype versions, many of the codecs were far from optimum with respect to delay. The evaluation of the delay had therefore to some extent to be based on theoretic analysis assuming the use of commercially available digital signal processors. The delay figures shown in Table 4 have all been normalized to reflect the net coder delay consisting of the theoretical algorithm delay (frame or window delay), a realistic processing delay and the delay due to the 16 kbit/s serial transmission.

The results given in Table 4 show that the codecs fall into two distinct classes: The SBC-ADPCM codec with only 7 ms delay and the remaining 5 codecs with delays in the range 35–45 ms. Taking into account the uncertainties associated with the estimation of the delay parameter, no discrimination was made between the 5 coders.

The GSM digital mobile system will be based on a so called narrow-band TDMA system. This implies that the speech coder information will be transmitted in bursts. Depending on the TDMA frame length, on the peak processing power of the processing device used, and on the structure of the algorithm (e.g. if it can be divided into sub-blocks), the codec delay will be between the frame length and the serial transmission delay given in Table 4.

Table 4
Delay properties of the candidate coders

Codec		Frame length	Delay assuming serial transmission
A	SBC-APCM	15 ms	< 40 ms
B	SBC-APCM	20 ms	<45 ms
C	SBC-APCM	16 ms	35 ms
D	SBC-ADPCM	1 ms	7 ms
E	RPE-LPC	19.5 ms	<40 ms
F	MPE-LTP	20 ms	36 ms

4.3. Implementation aspects

The aim of the evaluation of the implementation aspects was to assess the possible realization of the different candidate codecs as single VLSI chips and to estimate the power consumption requirements.

The basic data available for the evaluation were:
- the memory requirements for parameter storage (Random-Access Memory, RAM), the program size (Read-Only Memory, ROM);
- the computational complexity in terms of MOPS (Millions of arithmetic operations per second).

The data collected on the 6 speech coder prototypes are summarized in Table 5.

The random-access memory (RAM) is 16 bit for all processors used. The read-only memory (ROM) have been normalized to 16 bit words for comparison purposes. The figures for the number of multiplications and additions reflect the algorithm complexity to some extent. Not reflected in the table are the additional operations needed to implement the algorithms, branch operations, loop instructions, address calculations etc.

It is recalled that all candidate codecs were presented as laboratory prototypes implemented in most of the cases on several commercial DSPs. The implementations, basically aimed at a demonstration of the achievable quality were generally not optimized for effective implementation. The figures in the table are therefore not suited for detailed comparisons between the coding

Table 5
Memory requirements and computational complexity

Codec		Memory requirements		Computational complexity		
		RAM	ROM	ADDs	MPYs	Sum
A	SBC-APCM	0.820 k	3.4 k	0.6 M	0.6 M	1.2 M
B	SBC-APCM	1.660 k	5.5 k	0.9 M	0.5 M	1.4 M
C	SBC-APCM	0.350 k	3.5 k	0.8 M	0.7 M	1.5 M
D	SBC-ADPCM	0.450 k	2.2 k	1.0 M	0.9 M	1.9 M
E	RPE-LPC	0.550 k	4.5 k	0.8 M	0.7 M	1.5 M
F	MPE-LTP	0.340 k	12.6 k	3.5 M	1.4 M	4.9 M

schemes. However, it is clearly seen that codec F seems to be significantly more demanding than the remaining 5 both in terms of arithmetic operations and in memory requirements.

Comparing the figures listed in Table 5 to the capabilities of existing digital signal processors, it can be concluded that many of the candidate codecs are very close to or can in fact be implemented on one single DSP of todays generations.

A precise estimate of the power consumption of a chip can only be made reliably when the actual chip has been defined and the number of transistors can be counted. No detailed analysis of this nature was possible in the relative short period of time available. However, considering that the codec probably would be implemented using 1 μm CMOS technology and that application specific design would reduce the transistor count compared to a DSP solution, the preliminary conclusion was that any of the less complex algorithms in Table 5 would result in a power consumption in the order of 150–300 mW.

5. The GSM speech coder: RPE-LTP

In Table 6 the rank orders found in Tables 3–5 with respect to the main factors speech quality, delay and implementation aspects have been summarized. In the design of each speech coder, a number of trade-offs involving speech quality, delay and complexity had to be considered by the codec designers. Therefore, it was no surprise that no single coder was found to be superior in all respects.

However, the RPE-LPC (Coder E), was found to have the best average quality and was rank ordered in the "best" group as regards complexity. In the delay aspect, only the SBC-ADPCM (Coder D) achieved a siginificantly better performance – however at the expense of a certain reduction in speech quality, particularly in bit-error conditions. On this basis, the decision was taken to adopt the RPE-LPC, as the starting point for further studies.

Following a detailed analysis of the technical solution applied in the two LPC-type codecs, the possibility was seen to include the long-term prediction feature of the MPE-LTP into the RPE-scheme resulting in a new "compromise" scheme termed RPE-LTP.

A study by simulation of several options for this scheme was carried out and a new scheme was proposed. In Table 7, the basic features of the new scheme are compared to the original RPE-LPC.

The new scheme was reported to have the following properties:

– At a bit rate of 13.0 kbit/s this coder achieves a

Table 6
Rank order of candidate codecs. Entries in brackets cannot be discriminated

Parameter	Rank 1 (Best)	2	3	4	5	6
Voice quality	[E]	[F]	[C]	[A]	[D]	[B]
Delay	[D]	[A,B,C,E,F]				
Implementation	[A,B,C,D,E]	[F]				

Table 7
Comparison of RPE-LPC and RPE-LTP

Coder	RPE-LPC (Basic scheme)	RPE-LTP (New scheme)
Bit rate budget		
Net bit rate	14.8 kbit/s	13.0 kbit/s
Framing		
Speech frame	19.5 ms	20 ms
Window	2.5 ms Hamming, overlapping	20 ms Rectangular, no overlapping
PC-analysis		
Filter order	12	8
Algorithm	Schur	Shur
Coef. coding	52 bits	36 bits
Inverse filtering	Lattice type	Lattice type
Pulse modelling		
Method	Regular pulse (simplified procedure)	Regular pulse (simplified procedure)
No. of pulses	52/frame 4 bits/pulse	52/frame 3 bits/pulse

speech quality equivalent to that of the RPE-LPC at 14.8 kbit/s in error free conditions.
- The inclusion of the LTP loop results in a minor increase in the sensitivity to errors. However, if the bit rate reduction of 1.8 kbit/s is utilised for additional error protection, a net increase in error robustness can be achieved.
- The potential delay of the new scheme is 5 ms lower than the original one since no overlapping of analysis windows is required.
- The complexity of the new scheme is similar to that of the original one.

On this basis, it was decided to adopt the new RPE-LTP scheme for standardization in the Pan-European DMR system.

6. Further work

6.1. Error protection

The studies reported in section 3 and 4 of this paper were conducted in an early stage of the planning process, when both the speech coder and the radio and access techniques were still under study. Having fixed the technical solution for the GSM system, the problem of defining an optimized channel coding scheme has just begun (June 1987). The goal of the optimization is to achieve a satisfactory average performance with a minimum number of redundancy bits added to the 13 kbit/s stream from the speech coder.

The error protection envisaged will consist of two elements:

"Selective" forward error correction
It is well known that the various parameters of a medium/low bit rate speech coder have different sensitivities to bit errors. Consequently, for an effective utilization of a limited number of redundancy bits per speech coder frame, the output bits of the speech encoder have to be divided into classes according to their importance and given different degrees of protection. The problem to be resolved is to determine: the overall number of redundancy bits necessary, the number of classes n, the number of bits and the degree of protection in each class, and finally the error protection algorithm itself. This is a complicated multidimensional problem, and in order to reduce the number of variables, the following preliminary choices have been made:
- the number of classes will be 3;
- a 3-level protection will be achieved using punctured convolutional codes at the rates $\frac{1}{2}$, $\frac{3}{5}$ and $\frac{2}{3}$.

Error detection or "frame erasure" information
The residual error statistics expected at the speech coder interface, even after interleaving, are expected to be of burst nature, with strong correlation between the bit classes. In particular, when there is a non-correctable error in the most sensitive class, it is highly probable that the complete speech coder frame will be unrecoverable or "erased". By including a few parity bits in the speech coder frame, this situation can be detected, and the subjective effect of the erased frame can be ameliorated by extrapolating parameters from preceding speech frames. Of course, this procedure is only applicable for a limited period of time, and the strategy for dealing with this situation is under study.

A similar situation where frame extrapolation would be needed, may arise in handover situations, where very rapid signalling is necessary. In this situation, the possibility to "steal" speech coder frames for signalling purposes is being studied.

6.2. Future speech coder developments

In a mobile cellular system, the frequency spectrum represents a restricted resource that cannot be readily expanded if the need arises to increase the capacity of the system. In the GSM system two technical solutions will be studied to improve the spectrum efficiency.

Half rate codecs
In the TDMA access system now being defined, the speech coder frame (20 ms) will be transmitted in 4 TDMA bursts. In the planning of the radio and access systems, provisions will be made to accommodate "half rate" speech codecs without any modification of the TDMA system. A half rate codec will be transmitted in 2 TDMA frames, which implies that the gross bit rate (data

rate of speech coder + channel coder) will be half of that of the full rate channel. Considering that a low rate codec may produce a larger proportion of sensitive bits, the actual speech coder bit rate aimed at may be lower than 6.5 kbit/s.

Discontinuous transmission

In a two way conversation, the average speech activity of a subscriber will be less than 50%. If this could be utilized to avoid transmission in periods of silence, the interference level could be reduced significantly which would considerably improve the spectrum efficiency of the system. The essential component in such a system is the voice activity detector. In the speech coder scheme selected, interesting possibilities exist for reliable voice detection on a frame-by-frame basis. However, the problems of speech activity detection in noisy environment, noise switching contrasts and possible tandeming of voice-activated devices in other parts of the network will have to be studied.

7. Conclusions

A digital cellular mobile-radio system offering international roaming is under planning in the European countries. The progress of the speech coding studies have been presented: Following an introductory phase where performance requirements and testing methodology were defined, a coordinated evaluation programme was carried out. As a result of this, a 13 kbit/s Regular-Pulse Excitation LPC speech coding algorithm including Long-term Prediction, has been selected. The detailed specification of this algorithm has just begun (June 1987) and will be completed in early 1988.

Acknowledgments

The work reported here represents the outcome of contributions from a number of European companies and Telecommunications Administrations in the CEPT TM3/COST 207 Speech Coding Experts Group. Without the joint efforts and cooperation of these organisations, and not the least, of the individuals representing them, the successful completion of this multi-national development programme would not be possible.

References

[1] A.E. Coleman, N. Gleiss and P. Usai, "A subjective testing methodology for evaluating medium rate codecs for digital mobile radio applications", *Speech Communication*, this issue.

[2] V. Lazzari, R. Montagna, D. Sereno and A. Rusina, "Comparison of two speech codecs for DMR systems", *Speech Communication*, this issue.

[3] T.A. Ramstad, "Sub-band coder with a simple adaptive bit-allocation algorithm: A possible candidate for digital mobile telephony?" *ICASSP* 82, Paris (1982).

[4] T. Mattson and J. Uddenfeldt, "Digital vs. Analog Cellular Radio – Subjective Quality", 2nd Seminar on Land Mobile Digital Radio Communication, Stockholm, October (1986).

[5] R.B. Hanes and P. Atkins, "The UK candidate 16 kbit/s codec for the GSM pan-european study on digital cellular land mobile radio", *Speech Communication* , this issue.

[6] P. Vary, R. Hofmann, K. Hellwig and R.J. Sluyter, "A regular-pulse excited linear predictive codec", *Speech Communication*, this issue.

[7] C. Galand, M. Rosso, Ph. Elie, E. Lancon, "MPE/LTP coder for mobile radio application", *Speech Communication*, this issue