European Conference on
Speech Technology
Edinburgh, Scotland, UK
September 1987

POSITION REOPTIMISATION FOR A MULTIPULSE EXCITED LPC CODER

I. Boyd*

ABSTRACT

Since multipulse excited linear predictive coding was introduced several improvements to the basic multipulse coding algorithm have been suggested. The most notable of these are the inclusion of a pitch filter and amplitude reoptimisation. This paper presents an alternative to amplitude reoptimisation where not only the pulse amplitudes, but also the pulse positions are reoptimised. A comparison between position reoptimisation and amplitude reoptimisation in terms of complexity and performance is included.

INTRODUCTION

A multipulse LPC (MPLPC) encoder and decoder (ref 1) is illustrated in figure 1. In MPLPC, unlike LPC vocoding, no distinction is made between voiced and unvoiced speech: the same excitation waveform, a series of pulses, is used for all speech segments. For good quality speech several pulses are required per pitch period. However, as the pulse positions and amplitudes must be transmitted a quality versus bit-rate trade-off has to be made. The derivation of the appropriate pulse positions and amplitudes at the encoder is of course crucial to the coder performance and involves an analysis-by-synthesis procedure, i.e. the input and synthesised speech are compared and the excitation is derived to minimise the perceptually weighted error between the two signals. The perceptual weighting filter is included at the encoder so that the error between the input and synthesised speech is minimised, not in a mean-squared sense, but in such a way that it is perceptually less disturbing (ref 1).

10.21437/ECST.1987-162

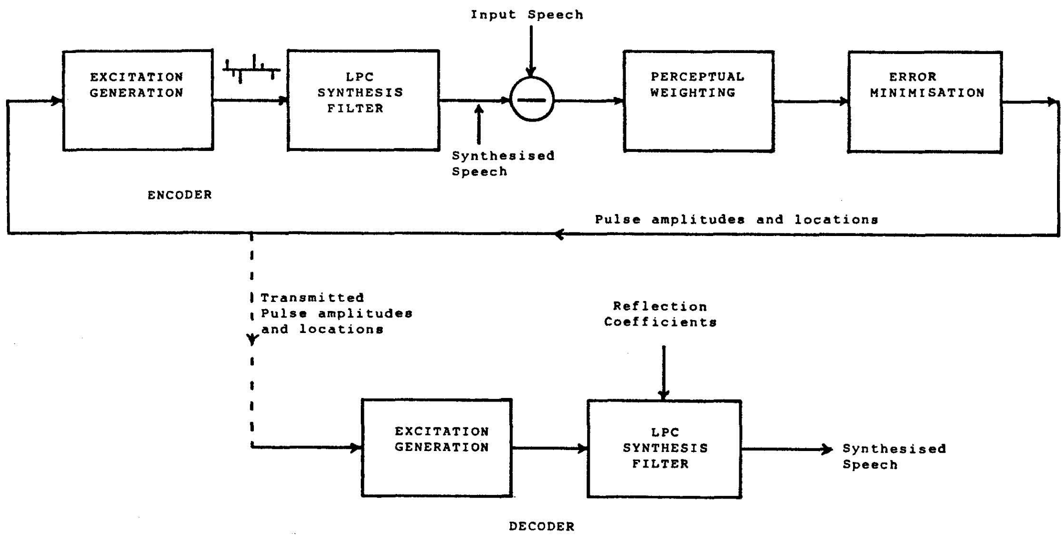

Figure 1 Multipulse Encoder and Decoder

* British Telecom Research Laboratories, Martlesham Heath, Ipswich, IP5 7RE

The excitation analysis procedure requires the partitioning of the input speech into small blocks (32- 128 samples), and a search for the pulse positions and amplitudes which minimise the error over the block. A one-pass solution to finding the optimum positions and amplitudes is a highly non-linear problem and is thus extremely complex. One possible method of finding the excitation is to split the problem into two parts, and try every possible combination of pulse positions. Given the pulse positions, the pulse amplitudes which minimise the error can be found relatively easily. The pulse positions and amplitudes which yield the lowest error over the block form the optimum excitation signal. Unfortunately, even for a small block size and only a few pulses per block, the number of possible combinations is relatively high, and results in an excessive computational load. Sub-optimal methods have thus been developed which find the pulse positions and amplitudes one at a time (ref 1 and 2). These sequential methods reduce the error minimisation process to that of selecting a pulse position as the location at which a maximum occurs in a cross-correlation function. Once a pulse position has been found the calculation of the corresponding pulse amplitude is straightforward. Given both the pulse position and amplitude, the cross-correlation function can be updated and the search for the next pulse can proceed.

The MPLPC coder described above can be extended (ref 3) to include pitch prediction and amplitude reoptimisation. The inclusion of the pitch filter means that advantage is taken of the long-term correlations in the speech waveforms and thus fewer pulses are required per pitch period to achieve the same speech quality (ref 3). Amplitude reoptimisation improves the performance of the sub-optimal pulse search methods and can be applied after all the pulse positions and amplitudes have been found or after each new pulse is found. When it is applied after each new pulse is found it considerably improves the performance of the encoder, at the cost of greatly increased complexity. Amplitude reoptimisation applied after all pulses have been found is much less complex but does not give as large an increase in performance. This paper proposes a new reoptimisation technique, which is applied after all the pulses have been found, where not only the pulse amplitudes but also the pulse positions are reoptimised. The investigations which led to the development of this technique are outlined in the next section.

PULSE POSITION AND SNR DISTRIBUTION ACROSS SUB-BLOCKS

All the results in this section were obtained for a multipulse encoder employing a covariance form of analysis (ref 4) without pitch prediction or perceptual weighting. The pitch and perceptual weighting filters were removed to give as simple an encoder as possible for the investigation. If both are included the distributions discussed below have the same form and the conclusions drawn are still valid.

An investigation into the distribution of pulse positions within multipulse sub-blocks revealed that pulses:

1. Concentrate around the first few sample positions;
2. Are evenly distributed over the middle positions;
3. Are sparse over the last few positions.

This pulse position distribution prompted an investigation into the SNR distribution within sub-blocks. The average SNR distribution across sub-blocks was approximated by evaluating three SNR figures: one for the first few sample positions of each sub-block (SNR_{start}), one for the mid positions (SNR_{mid}) and one for the last few positions (SNR_{end}). The value SNR_{start} is obtained by accumulating the input signal power and the noise power for the first few sample positions in every sub-block and dividing the resulting signal power by the noise power. The values SNR_{mid} and SNR_{end} are similarly obtained. These values indicate how the performance of the analysis-by-synthesis procedure varies within sub-blocks. Figure 2 illustrates the average pulse and SNR distributions for a multipulse coder with 32 samples/sub-block. The illustrated distributions were from 10 seconds of male speech: other male and female speech yields very similar distributions. The SNR distribution indicates that on average, the analysis-by-synthesis waveform matching process results in better matching at the end and middle of sub-blocks than at the beginning.

Both distributions indicate weaknesses in the multipulse encoding technique which result from the block minimisation approach adopted, ie pulses are placed within a block to minimise the error over that block alone. Pulses placed near the start of blocks can minimise the error over almost the whole block, whereas pulses near the end of the block can only minimise the error to the end of the block. Pulses placed near the start of sub-blocks are therefore more effective in minimising the block error and hence pulses tend to concentrate around the first few sample positions. The SNR distribution can be explained by appreciating that the error minimisation effect of all pulses add together near the end of sub-blocks. It is this cumulative minimisation effect which results in better matching near the end of sub-blocks. Another contributing factor, however, is that pulses near the start of sub-blocks attempt to minimise the error over nearly all the sub-block whereas pulses near the end attempt to minimise the error over only the last few block positions and thus can function very effectively.

POSITION REOPTIMISATION

From the SNR distribution it was concluded that pulses near the start of sub-blocks should not waste their waveform matching ability on sample positions beyond the position of the next pulse, as this is really the domain of the next pulse. Ideally each pulse should attempt to minimise the error only up to the sample just before the position of the next pulse. This ideal can be approximated as follows. Firstly, all the pulse positions within a sub-block are estimated using one of the standard sub-optimal search methods (ref 1 and 2). The pulse positions obtained are assumed to be close to those which would result from an exhaustive search (optimum positions). Secondly, starting from left to right, the error minimisation effect of each pulse up to the position just before the next estimated pulse position is calculated, for the current position of the pulse and a few positions to the left and right. The position which gives the lowest error is selected as the new position and the pulse amplitude which minimises the error up to the next pulse position is then calculated. As only a few possible positions are considered, the added complexity of position reoptimisation is not as high as amplitude reoptimisation applied after each pulse is found and its performance is superior at lower pulse rates. It consistently out performs the lower complexity amplitude reoptimisation applied after all pulses have been found.

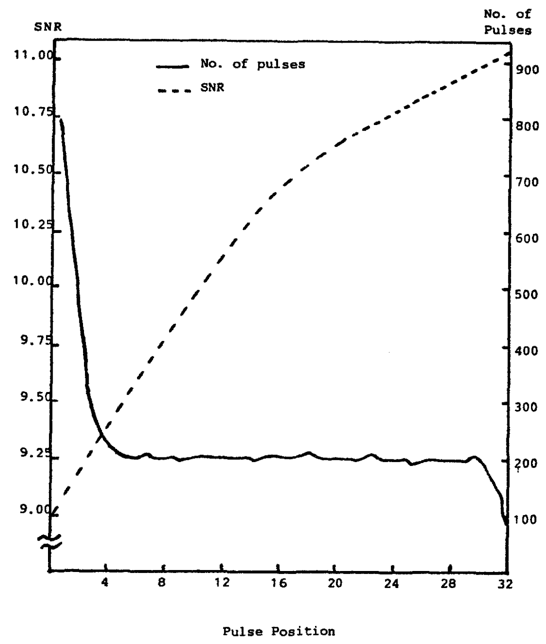

Figure 2 Pulse Position and SNR distributions

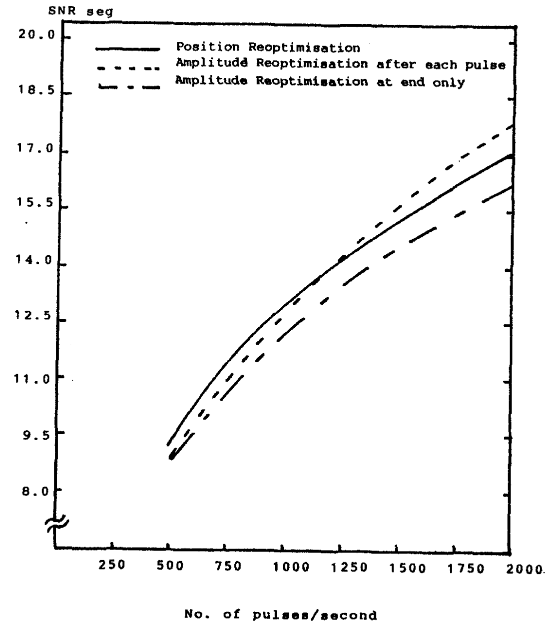

Figure 3 Amplitude and Position Reoptimisation Comparison

CONCLUSIONS

Position reoptimisation offers an alternative to amplitude reoptimisation. It is applied only once, after all the pulse positions have been estimated and results in a moderate increase in complexity. The performance of position reoptimisation is almost identical to the much higher complexity amplitude reoptimisation applied after each new pulse is found. It is possible to implement a multipulse codec employing position reoptimisation in real-time, using some of the faster floating point DSP devices now becoming available.

Unfortunately insufficient space is available to report on the results of combining position and amplitude reoptimisation, the use of two position reoptimisation passes and attempts to effectively overlap sub-blocks. It is hoped that there will be sufficient time during the presentation to present some of these results.

REFERENCES

1. B.S. Atal and J.R. Remde, Proc. IEEE ICASSP, vol 1, pp 614-617, 1982.
2. K. Ozawa, et al, Proc IEEE Global Telecommun. Conf., pp 794-798,1983.
3. S. Singhal and B.S. Atal, Proc. IEEE ICASSP, vol 1, pp 1.3.1-1.3.4, 1984.
4. M. Berouti et al, Proc. IEEE ICASSP, vol 1, pp 10.1.1-10.1.4, 1984.

ACKNOWLEDGEMENT

Acknowledgement is made to the Director of Research of British Telecom for permission to publish this paper.