obtained for the photodetected current in the SSB case:

$$I_{out}(L_1+L_2,t) \propto \frac{L_1}{L_1+L_2} \cdot \exp\left(-\frac{2t^2}{\left[\frac{\tau(L_1+L_2)}{L_1}\right]^2}\right) \cdot \left[1+2a\cos\left(2\pi f_m t\frac{L_1}{L_1+L_2}+2\pi^2 L_2\beta_2 f_m^2\frac{L_1}{L_1+L_2}\right)\right] \quad (2)$$

It may be observed that while the time-stretch phenomenon is still present, the dispersion factor is now encountered in the phase term of the electrical photodetected signal. Therefore, when SSB modulation is employed no power penalty is obtained, as shown in Fig. 3, although some phase distortion is produced. However, this phase distortion may be easily equalised in the electrical domain, as discussed in [4].

Conclusion: The bandwidth limitation of photonic ADCs based on time-stretch techniques has been analysed. It has been shown that, when SSB modulation is employed, the chromatic dispersion power penalty which causes the bandwidth limitation is completely overcome. A dispersion factor remains in the phase term, but the resulting phase distortion may be eliminated in the electrical domain. Therefore, the authors propose SSB modulation as a simple and affordable technique which overcomes the main drawback of photonic ADCs based on time-stretch techniques.

13 October 2000
Electronics Letters Online No: 20010046
DOI: 10.1049/el:20010046

J.M. Fuster and J. Marti (*Fibre Radio Group, ITACA Research Institute, ETSI Telecomunicacion, Universidad Politecnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain*)

E-mail: jfuster@dcom.upv.es

D. Novak and A. Nirmalathas (*Australian Photonics Cooperative Research Centre, Photonics Research Laboratory, Department of Electrical and Electronic Engineering, The University of Melbourne, Grattan Street, Parkville, Vic. 3052, Australia*)

References

1 COPPINGER, F., BHUSHAN, A.S., and JALALI, B.: 'Photonic time stretch and its application to analogue-to-digital conversion', *IEEE Trans. Microw. Theory Tech.*, 1999, **47**, pp. 1309–1314
2 SMITH, G.H., NOVAK, D., and LIM, C.: 'Technique for optical SSB generation to overcome dispersion penalties in fibre-radio systems', *Electron. Lett.*, 1997, **33**, pp. 74–75
3 MARTI, J., FUSTER, J.M., and LAMING, R.I.: 'Experimental reduction of chromatic dispersion effects in lightwave microwave/millimetre-wave transmissions employing tapered linearly chirped fibre gratings', *Electron. Lett.*, 1997, **33**, pp. 1170–1171
4 CORRAL, J.L., and MARTI, J.: 'Single sideband optical modulation on chirped fibre grating based delay lines for optically controlled phased array antennas', *Electron. Lett.*, 1999, **35**, pp. 761–762

Voice activity detector based on CAPDM architecture

Chia-Horng Liu and Chia-Chi Huang

A voice activity detector based on the controlled adaptive prediction delta modulation architecture is presented. The proposed detector utilises two parameters, a prediction gain and a prediction indicator. Computer simulation results show the proposed voice activity detection algorithm to be very robust in a noisy environment.

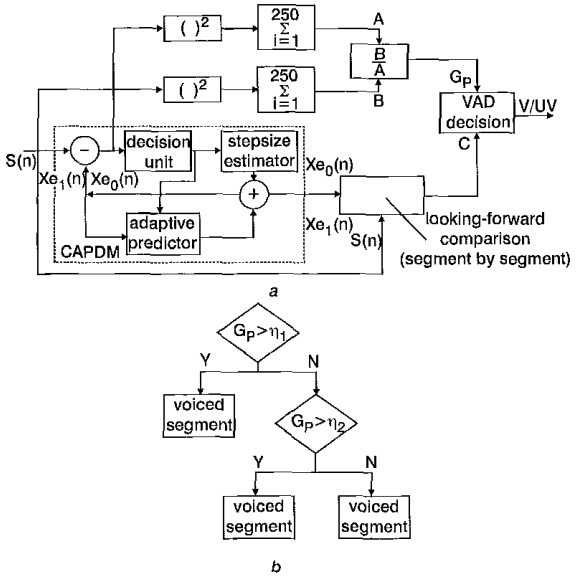
Introduction: Discontinuous transmission (DTX) is an efficient and simple technique which makes use of the on-off characteristics of human conversations. Using DTX, only the active periods of a conversation are transmitted. The inactive periods are detected and removed at the transmitter, and only about 40% of a conversation is transmitted. Therefore, system capacity can be increased and transmission power reduced, at the same time.

To implement the DTX mechanism, a voice activity detection (VAD) algorithm is needed to group the input speech samples into voiced and unvoiced segments. Conventional VAD algorithms are based on monitoring short-time energy, $E_s(n)$, and the zero-crossing rate, $Z_s(n)$, of a speech signal [1]. Based on these two parameters, commonly used VAD decision rules for deciding the presence of voiced segments are [2]

$$E_s(n) > \eta_e \quad (1)$$

$$Z_s(n) < \eta_{zl} | Z_s(n) > \eta_{zu} \quad (2)$$

where η_e is an energy threshold, and η_{zl} and η_{zu} are the lower bound and upper bound, respectively, of the zero-crossing rate. It has been observed that the performances of the traditional VAD algorithms are poor at low signal-to-noise ratio (SNR) [3].



726/1

Fig. 1 *Speech coder and voice activity detector, and VAD decision rules*
a Block diagram of CAPDM speech coder and proposed voice activity detector
b VAD decision rules

Basic structure of CAPDM: A block diagram of a CAPDM encoder is shown in the dashed box of Fig. 1*a* [4]. Basically, it consists of a decision unit, a stepsize estimator, and an adaptive predictor. The function of the decision unit is to compare the distance between the current speech sample and the two estimated samples, assuming bit 0 or bit 1. It decides whether 0 or 1 is to be sent, according to which estimate is closer. This is the feature of one-step looking forward decision.

The function of the stepsize estimator is to produce both syllabic and instantaneous stepsize estimation. These two estimates are then combined to generate the current stepsize estimation, $\Delta_0(n)$ and $\Delta_1(n)$, for bit 0 and bit 1. The function of the adaptive predictor is to generate two estimates of the current speech sample, $Xe_0(n)$ and $Xe_1(n)$, for bit 0 and bit 1, as shown in the following equations:

$$Xe_0(n) = \sum_{i=1}^{N} a_{0i}(n) * Xe(n-i) + \Delta_0(n) \quad (3)$$

$$Xe_1(n) = \sum_{i=1}^{N} a_{1i}(n) * Xe(n-i) + \Delta_1(n) \quad (4)$$

where $Xe(n-i)$ is the previously chosen estimate at time instant $(n-i)$. The filter coefficients, $a_{0i}(n)$ and $a_{1i}(n)$, are adapted recursively using a simplified stochastic approximation of the gradient method.

VAD algorithm designed for CAPDM architecture: Two parameters extracted from the CAPDM architecture for VAD usage are the prediction gain G_p and the prediction indicator C, as shown in

Fig. 1*a*. The prediction gain [3] is defined as the ratio of the power of the input speech signal and the residual signal over a speech segment (~16ms). The residual signal is the difference between the input speech samples and the predicted speech samples. The prediction gain works well at a low SNR (> 5dB). However, when the SNR decreases below 5dB, the misclassification probability increases rapidly. Therefore, in addition to the prediction gain, we suggest use of a prediction indicator to reduce the misclassification probability at a very low SNR.

The prediction indicator is the number of occasions that two consecutive speech samples are not bounded by the two estimated speech samples for both the assumed bit 0 and bit 1, as expressed in the following equations:

$$\text{If } Xe_1(n) < S(n) \text{ and } Xe_0(n+1) > S(n+1) \qquad C++ \tag{5}$$

$$\text{If } Xe_0(n) < S(n) \text{ and } Xe_1(n+1) > S(n+1) \qquad C++ \tag{6}$$

Since CAPDM has good speech waveform tracking ability, the prediction indicator for a voiced segment is smaller than an unvoiced segment. Furthermore, when the background noise is large, the prediction indicator still has good classification capability. Therefore, the prediction indicator is used along with the prediction gain for VAD. Overall, the VAD operation consists of two steps, as shown in Fig. 1*b*. The first step is to compare the prediction gain G_p with a threshold η_1. A segment with the prediction gain larger than η_1 is classified as a voiced segment. Otherwise, we compare the prediction indicator C with another threshold η_2. A segment with C smaller than η_2 is classified as a voiced segment. Otherwise, it is considered as an unvoiced segment. It is noted that η_1 and η_2 are derived by simulation to achieve the smallest misclassification probability.

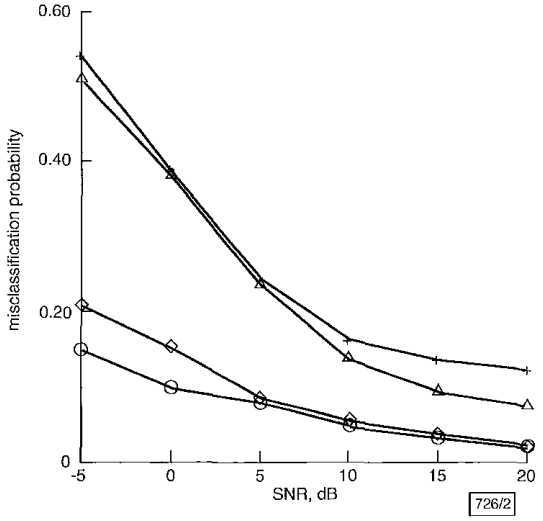


Fig. 2 *Comparison of misclassification probability of proposed and conventional VAD algorithms*

○ prediction gain + prediction indicator
◇ prediction gain
△ zero-crossing rate
+ energy levels

Performance comparisons: We evaluated the performance of the proposed VAD algorithm by comparing the misclassification probability with two conventional VAD algorithms [2]. The misclassification probability is the sum of miss detection and false alarm probability. The input speech signal was sampled at 16kbit/s and processed by a CAPDM speech coder. The background noise was an additive white Gaussian signal. The length of a segment was 250 samples, which is ~16ms. The η_1 is 0.9 and the η_2 is 80, both formed for our simulation.

Fig. 2 shows the misclassification probability for the proposed VAD algorithm and the conventional ones, with SNR varied from −5 to 20dB. It shows that the misclassification probability of the proposed VAD algorithm is much smaller than that of the conventional algorithms over the whole SNR range, especially at low SNR. Fig. 2 also shows that the misclassification probability is reduced at very low SNR by the use of prediction indicator.

Conclusions: We present a new VAD algorithm using both the prediction gain and the prediction indicator based on the CAPDM architecture. This new VAD algorithm can classify speech segments well at a very low SNR. The results obtained from our simulation show that the misclassification probability of the proposed method are much smaller than the conventional methods. Through the use of the proposed VAD algorithm with DTX, the effective data rate of a 16kbit/s CAPDM speech coder can be further reduced.

19 October 2000
Electronics Letters Online No: 20010030
DOI: 10.1049/el:20010030

Chia-Horng Liu and Chia-Chi Huang (*Wireless Communication Laboratory, Department of Communication Engineering, National Chiao-Tung University, No. 1001, Ta Hsueh Road, Hsinchu, Taiwan, Republic of China*)

E-mail: huangcc@cc.nctu.edu.tw

References

1 RABINER, L.R., and SAMBUR, M.R.: 'An algorithm for determining endpoints of isolated utterances', *Bell Syst. Tech. J.*, 1975, **54**, (2), pp. 297–300
2 GÖKHUN TANYER, S., and ÖZER, H.: 'Voice activity detection in nonstationary noise', *IEEE Trans. Speech Audio Process.*, 2000, **8**, (4), pp. 478–482
3 HATAMIAN, S.: 'Enhanced speech activity detection for mobile telephony'. Proc. IEEE 42nd Veh. Technol. Conf., 1992, pp. 159–162
4 LIU, CHIA-HORNG, and HUANG, CHIA-CHI: 'A packet-based CAPDM speech coder for PCN applications', *IEEE Trans. Veh. Technol.*, 2000, **49**, (3), pp. 753–765

Improving direct torque control by means of fuzzy logic

A. Arias, L. Romeral, E. Aldabas and M.G. Jayne

Investigations are carried out on a fuzzy logic direct torque control system, which not only reduces the torque ripple, but also reduces the reactive power taken from the mains by choosing the optimum reference flux value. Additionally, this fuzzy controller is adaptive and may be applied to any induction motor.

Introduction: Classical direct torque control (DTC) has inherent disadvantages, such as problems during starting resulting from the null states, the compulsory requirement of torque and flux estimators, and torque ripple. This Letter is focused on the means of reducing torque ripple, which is the most undesirable of the disadvantages mentioned [1, 2]. This is achieved by the application of a duty cycle, which will impose an inverter switching state which lies between the selected state of the classical DTC system and a null state [3].

A stator flux reference optimum controller is also designed, which not only helps to achieve a smaller torque ripple, but also reduces the reactive power consumption of the drive. This is achieved by changing the stator flux reference value with reference being made to the correspondent torque reference value. Therefore, the stator flux reference value chosen is to be just of sufficient value to produce the desired torque.

New controller: In classical DTC induction motor drive systems, there are torque and flux ripples which are attributable to the none existence of the voltage source inverter (VSI) states which can generate the exact voltage value which will make the electromagnetic torque error and stator flux error both zero [4].

This new technique is based on supplying the inverter with the selected active states for sufficient time to achieve the torque and flux references values. For the rest of the switching period, a null state is selected that will not almost change either torque or flux