Table 2: Number of Sub-ANNs and corresponding range

Number of sub-ANNs	The positive range
8	0-512-1024-1536-2048
14	0-128-256-384-512-1024-1536-2048
20	0-128-256-384-512-768-1024-1280-1536-1792-2048
26	0-32-64-96-128-256-384-512-768-1024-12801536-1792-2048

Table 3: Number of sub-ANNs and its SNR values

Number of sub-ANNs	SNR
	dB
8	14.59
14	16.29
20	16.41
26	17.43

Multi-ANN system design considerations: Because the sub-ANNs are used for different sample level-bands, the structure of each sub-ANN needs to be optimised with respective to the appropriate band and for this purpose several structures were tried to achieve optimisation, prior to the design of the multi-ANN system. The number of sub-ANNs used in the multi-ANN system corresponding to different level-bands affeets the performance. Experiments were performed using different numbers of sub-ANNs: 8, 14, 20 and 26. A system with eight sub-ANNs has the distribution of level-bands given in Table 1. The use of sub-ANNs in this manner improves the possibility of a defined output for obviously similar input data sets. The width of the level-bands are chosen with regard to the graph in Fig. 1. Table 2 lists the number of sub-ANNs and the corresponding bands (positive), used in the experiments. For each of these sub-ANN configurations, the average quantised error (with respect to the dynamic speech range) and SNR values were computed. Fig. 3 shows the average quantised error for a system employing 26 sub-ANNs. Table 3 gives the SNR values for the system.

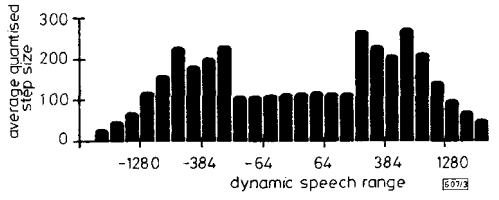

Fig. 3 *Error distribution graph for multi-ANN system with 26 sub-ANNs*

Discussion and conclusions: As mentioned earlier, a single-ANN system yields an SNR value of 12dB and an average quantised error of 1600. Fig. 2 and Table 3 show the improvement due to the inclusion of sub-ANNs in the multi-ANN system, the maximum quantised error now being less than 300 and the SNR being 17.43dB. With a large number of sub-ANNs catering for the lower region of the speech range [–128 to 128], the error is less than 115 pointing to a reduction by a factor of 14 with respect to the single-ANN system. The prediction and compression can be further improved by designing a more efficient system for data set distribution to the sub-ANNs.

7 December 1994
Electronics Letters Online No: 19950173

Y.K. Peng, S.Q.A.M.A. Hossain and D.A. Rankin (*The Queen's University of Belfast, Department of Electrical and Electronic Engineering, Ashby Building, Stransmillis Road, Belfast BT9 5AH, United Kingdom*)

References

1 LI, J., and MANIKOPOULOS, C.N.: 'Non-linear prediction in image coding with DPCM', *Electron. Lett.*, 1990, **26**, (17), pp. 1357–1359

2 LIPPMANN, R.P.: 'An introduction to computing with neural nets', *IEEE ASSP Magazine*, April 1987, **4**, pp. 4–22

DSA type secure interactive batch verification protocols

L. Harn

Indexing terms: Smart cards, Cryptography

The author proposes some DSA type secure interactive batch verification protocols, where the signer generates n signatures through interactions with the verifier, and the verifier validates all these n signatures.

Introduction: Naccache *et al.* [1] proposed various methods for optimising the DSA [2] for smart card applications. These included an interactive batch verification protocol, where the signer generates n signatures through interactions with the verifier, and then the verifier validates all these n signatures at once based on the batch verification criterion. Lim and Lee [3] pointed out that the interactive DSA batch protocol proposed by Naccache *et al.* is insecure. The fundamental problem of this protocol is that the batch verification criterion used to verify n signatures is insecure, although it is derived by combining multiple secure DSA signature equations. In other words, Lim and Lee show ways to generate individual signatures to satisfy the batch verification criterion; but the individual signatures cannot satisfy the DSA signature verifications. Because the purpose of the protocol is to verify all the multiple signatures at once, we would like to propose some secure batch verification criteria to preserve security.

DSA type interactive batch transaction: This involves two procedures. The signer follows the signature collection protocol to generate n signatures through interactions with the verifier. Then the verifier validates all these n signatures at once through the batch verification criterion.

(a) Signature collection protocol: Let p, q and α be the DSA type public parameters: p a large prime, q a large prime divisor of $p-1$ and α an element in Z_p of order q. Let H denote the secure hash algorithm (SHA) and m_i the ith message to be signed. The signer has a secret key $x \in Z_q$ and a public key $y = \alpha^x \bmod p$.

(i) The signer randomly picks $k_i \in Z_q$, and computes $r_i = \alpha^{k_i} \bmod p$, for $i = 1, 2, ..., n$. The signer also computes $R = \Pi_{i=1}^{n} r_i \bmod p$. The signer sends R to the verifier.

(ii) For i = 1, 2, ..., n, the following steps are performed:

(a) The signer sends r_i to the verifier.

(b) The verifier replies with an e-bit message randomiser b_i.

(c) The signer sends $s_i = xRH(m_i \| b_i) - k_i \bmod q$.

(b) Batch verification criterion: The verifier checks that

$$R = \prod_{i=1}^{n} \bmod p \text{ and } y^{R\sum_{i=1}^{n} H(m_i \| b_i)} = R\alpha^{\sum_{i=1}^{n} s_i} \bmod p$$

and replaces $\{r_i, s_i, b_i, m_i | \text{for } i = 1, 2, ..., n\}$ by the DSA type triples $\{R \bmod q, \Sigma_{i=1}^{n} s_i \bmod q, (m_i \| b_i | \text{for } i = 1, 2, ..., n)\}$.

Theorem: If all signatures are generated by the legitimate signer, it will satisfy the verification criterion.

Proof: According to (ii)(c), each signature satisfies

$$y^{RH(m_i \| b_i)} = r_i \alpha^{s_i} \bmod p$$

By multiplying the above equation repeatedly for $i = 1, 2, ..., n$, we have

$$\prod_{i=1}^{n} y^{RH(m_i \| b_i)} = \prod_{i=1}^{n} r_i \alpha^{s_i} \bmod p$$

Thus, we obtain

$$y^{R\sum_{i=1}^{n} H(m_i \| b_i)} = R\alpha^{\sum_{i=1}^{n} s_i} \bmod p \qquad QED$$

This scheme is essentially as fast as a single DSA verification. Because the verification criterion is one of the secure ElGamal type signature schemes as proposed in [4], i.e. with signature equa-

tion $rmx = k+s \bmod \emptyset(p)$ and signature verification $y^{rm} = r\alpha^s \bmod p$), only the legitimate signer with knowledge of x can generate the signatures to satisfy the verification. Although the attacker can generate bogus signatures in the signature collection protocol, these signatures cannot satisfy the batch verification criterion. There are some other secure ElGamal type signature schemes as proposed in [4] that can also be used to design similar DSA type secure interactive batch verification protocols. We list those schemes in Table 1.

Table 1: Secure ElGamal type signature schemes

	Signature equation	Signature verification
(1)	$mx=rk+s \bmod \emptyset(p)$	$y^m=r^r\alpha^s \bmod \emptyset p$
(2)	$sx=rk+m \bmod \emptyset(p)$	$y^s=r^r\alpha^m \bmod \emptyset p$
(3)	$sx=k+mr \bmod \emptyset(p)$	$y^s=r\alpha^{mr} \bmod \emptyset p$
(4)	$(r+m)x=k+s \bmod \emptyset(p)$	$y^{r+m}=r\alpha^s \bmod \emptyset p$
(5)	$sx=k+(m+r) \bmod \emptyset(p)$	$y^s=r\alpha^{m+r} \bmod \emptyset p$

Conclusion: Instead of using an insecure batch verification criterion as proposed by the Naccache *et al.* in Eurocrypt '94, we propose several secure batch verification criteria in this Letter. By using the interactive batch verification protocol, the signer follows the signature collection protocol to generate n signatures through interactions with the verifier and the verifier validates all these n signatures at once through the batch verification criterion.

6 December 1994
Electronics Letters Online No: 19950203

L. Harn (*Computer Science Telecommunications Program, University of Missouri – Kansas City, MO 64110, USA*)

References

1 NACCACHE, D., M'RAIHI, D., RAPHEALI, D., and VAUDENAY, S.: 'Can DSA be improved: Complexity trade-offs with the digital signature standard'. Pre-proc. Eurocrypt'94, 1994, pp. 85–94
2 NIST: 'Digital signature standard' (FIPS PUB XX, 1993)
3 LIM, C.H., and LEE, P.J.: 'Security of interactive DSA batch verification', *Electron. Lett.*, 1994, **30**, (19), pp. 1592–1593
4 HARN, L., and XU, Y.: 'Design of generalised ElGamal type digital signature schemes based on the discrete logarithm', *Electron Lett.*, 1994, **30**, (24), pp. 2025–2026

Speaker-independent Mandarin plosive recognition with dynamic features and multilayer perceptrons

W.-Y. Chen and S.-H. Chen

Indexing terms: Neural networks, Speech recognition

A new method for recognising plosives in isolated Mandarin syllables is discussed in the Letter. After automatically detecting the plosive segment of the input utterance, some dynamic features are extracted from its spectral parameter contours using orthonormal polynomial transforms. Next, an MLP trained with an algorithm based on a minimum error criterion is employed to distinguish plosives using these features. A promising recognition rate of 73.6% is achieved in a speaker-independent test using a database containing utterances of 110 syllables uttered by 100 speakers.

Introduction: Each character in Mandarin speech is pronounced as a syllable. An isolated Mandarin syllable can be decomposed phonetically into initial and final subsyllable units. Only 22 initials (including a dummy one) and 39 finals are available in Mandarin speech. Six of these 22 initial subsyllables are plosives, /b, d, g, p, t, k/. Similar to the vocabulary in the English E-set, Mandarin syllables with the same final subsyllable and different plosive initial subsyllables form a confusing set. Identifying these six plosives still remains the most challenging task in Mandarin speech recognition. Previous related investigations have emphasised the selection of effective features for recognition [1–3]. Wang [1, 2] *et al.* suggested extracting some features from the burst spectrum, the format transition and the voice-onset time (VOT). Next, an MLP was employed to recognise the three unaspirated Mandarin plosives /p, t, k/. Although a high recognition rate was obtained, the method was tested only on a small database containing utterances of a small vocabulary (nine syllables with /p,t,k/ followed by /i,a,u/) as generated by seven speakers. Besides, recognition features were not automatically extracted. Manual preprocessing should be required to detect the VOT.

In this study, a new method for recognising the six plosives in isolated Mandarin syllables is discussed. In this method, the plosive part of the input testing utterance is first detected automatically. For the plosive segment, the percepture linear predictive (PLP) [4] features are extracted. In the PLP analysis technique, some properties of hearing, e.g. the critical-band spectral resolution, the equal-loudness pre-emphasis, and the intensity loudness power law, are simulated to obtain an auditory-like spectrum. The PLP features are the cepstral coefficients of an autoregressive all-pole model of the auditory-like spectrum of speech. Next, these PLP features are transformed into another feature set by using orthonormal polynomial transforms to represent the dynamics of spectral parameter contours. An MLP trained with an algorithm based on the minimum error criterion [5] is then employed to recognise the plosive. Notably, the number of recognition features is fixed and independent of the length of the testing utterance. Therefore, the time alignment between the testing plosive segment and the MLP recogniser is not required. Moreover, the performance of the proposed method is tested in a speaker-independent recognition mode using a database containing 100 repetitions of utterances of 110 syllables which are all possible combinations of the six plosives and 39 final subsyllables.

Orthonormal polynomial transform: The speech signal is a dynamic signal in nature. As an utterance is divided into segments, each segment should be treated as a dynamic rather than a static signal. This phenomenon is especially true for plosives to recognise in this study. Therefore, instead of representing each parameter contour of a plosive segment by a constant curve, approximating it by a smooth curve is preferred. Distortion can thereby be reduced owing to the better curve fitting. In this study, the smooth curve is a reconstructed version of the original parameter contour obtained by orthonormal polynomial expansion using several low-order coefficients. Specifically, a parameter contour of a plosive segment with length $N + 1$ frames is allowed to be denoted by $f(n/N)$, $n = 0, \ldots, N$. The smooth curve used to approximate it can then be expressed by

$$\hat{f}\left(\frac{n}{N}\right) = \sum_{j=0}^{r} \alpha_j \phi_j\left(\frac{n}{N}\right) \quad 0 \le n \le N \tag{1}$$

where

$$\alpha_j = \frac{1}{N+1} \sum_{n=0}^{N} f\left(\frac{n}{N}\right) \phi_j\left(\frac{n}{N}\right)$$

and r is the order of the orthonormal polynomial expansion. As $r = 2$, the first three basis functions of the orthonormal polynomial transform can be expressed as [6]

$$\phi_0\left(\frac{n}{N}\right) = 1 \tag{2}$$

$$\phi_1\left(\frac{n}{N}\right) = \left[\frac{12N}{N+2}\right]^{\frac{1}{2}} \left[\frac{n}{n} - \frac{1}{2}\right]$$

$$\phi_2\left(\frac{n}{N}\right) = \left[\frac{180N^3}{(N-1)(N+2)(N+3)}\right]^{\frac{1}{2}} \left[\left(\frac{n}{N}\right)^2 - \frac{n}{N} + \frac{N-1}{6N}\right]$$

for $0 \le n \le N$ and $N \ge 3$. Notably, all these three basis functions are normalised, in length, to [0,1]. After performing orthonormal polynomial transforms, coefficients of all parameter contours are accumulated and fed into the following MLP recogniser for plosive discrimination. If there are p spectral parameter contours, $p(r + 1)$ recognition features are obtained.