Cover page

Name of the workshop: *2002 International Computer Symposium (ICS2002)*

Title of the paper: *An Efficient Voice over Internet Protocol Technique Combining the Speech Data Encryption and G.729 Error Recovery*

Short abstract: *This paper proposes a Voice over Internet Protocol (VoIP) technique combining the speech data encryption and G.729 error recovery. This technique uses the chaotic data interleaving on inter-frames of voice to make the situation of continuous packet loss becoming an isolated packet loss situation. Then, we propose a Periodical Parameter Re-initialization (PPR) recovery approach to reduce the signal quality degradation in the G.729 decoder due to the lost of state synchronization to the G.729 encoder. Besides, the proposed VoIP technique also uses the idea of chaotic data encryption on intra-frames of speech to scramble the data sequence within a speech frame. The simulation results show that using the proposed VoIP technique can have averagely 20 db gain as compared with the original G.729 CODEC with respective to different packet loss rates ranging from 5% to 30%.*

Names of the authors: *Jiun-In Guo, Chiun-Chau Lin, and Sheng-Wei Lin*

Current affiliation: *Department of Computer Science and Information Engineering, National Chung Cheng University*

Postal address: *No. 160 , San-Hsing, Min-Hsiung, Chia-Yi 621, Taiwan, R.O.C.*

Email addresses: *jiguo@cs.ccu.edu.tw, lcch90@cs.ccu.edu.tw, lsw90@cs.ccu.edu.tw*

Fax number: *05-2720859*

Name of the contact author: *Jiun-In Guo*

Keywords: *VoIP, Speech encryption, and G.729 Error resilience*

An Efficient Voice over Internet Protocol Technique Combining the Speech Data Encryption and G.729 Error Recovery

Jiun-In Guo, Chiun-Chau Lin, and Sheng-Wei Lin

Department of Computer Science and Information Engineering,
National Chung Cheng University, Chia-Yi 621, Taiwan, R.O.C.
Email: jiguo@cs.ccu.edu.tw

ABSTRACT

This paper proposes a Voice over Internet Protocol (VoIP) technique combining the speech data encryption and G.729 error recovery. This technique uses the chaotic data interleaving on inter-frames of voice to make the situation of continuous packet loss becoming an isolated packet loss situation. Then, we propose a Periodical Parameter Re-initialization (PPR) recovery approach to reduce the signal quality degradation in the G.729 decoder due to the lost of state synchronization to the G.729 encoder. Besides, the proposed VoIP technique also uses the idea of chaotic data encryption on intra-frames of speech to scramble the data sequence within a speech frame. So, even if the scrambled speech data are stolen in the transmission process, they cannot be recovered back through the original speech decoder. This provides a protection scheme for the transmission of important speech data. In order to demonstrate the performance of the proposed VoIP technique, we have successfully verified and analyzed the proposed technique through software simulation and statistical measures on several test voices. The simulation results show that using the proposed VoIP technique can have averagely 20 db gain as compared with the original G.729 CODEC with respective to different packet loss rates ranging from 5% to 30%.

Keywords: VoIP, Speech encryption, and G.729 Error resilience

1. INTRODUCTION

The emergence of the Internet Protocol (IP) as the standard transport for packet data network has enabled a revolution in communications service and applications. Due to the popularity and high evolution of Internet technologies, Internet has been widely used in the data services like email, ftp, and WWW. New classes of IP-based information appliances are also emerging, which provide interactive audio and video services for peoples. The Voice over Internet Protocol (VoIP) means to transmit voice through the Internet [1-2]. It saves much money to make a long-distance call through the Internet for its low cost. However, the datagram-oriented IP networks typically offer only a "best effort" service for a certain data flow, which does not make any commitment about a required minimum bit-rate or a maximum delay allowed when networks congestion occur. This phenomenon will cause massive continuous packet loss and long packet delay. Consequently, when the networks get congested, real-time packets may arrive too late at the receiver or may be dropped due to buffer overflow at routers. In the case of the transmission of real-time telephone-quality speeches for conferencing applications, packet loss causes signal drops that are very annoying for the listeners. Even we use the speech CODEC with automatically error recovery characteristics like G.723.1 and G.729, the continuous packet loss will cause the speech quality degradation not only in the lost packets, but also in the following packets due to the lost of state synchronization between the speech encoders and decodes. In addition, since the important data is possible to be corrupted or stolen by the hackers in the Internet, it is necessary to perform the data encryption on the data when it is transmitted through the Internet for the sake of ensuring the data security.

In the literatures, there are many researches on the speech encryption using analog scrambling techniques [3-7] and the error recovery of the speech signals based on G.729 [8-13]. In the aspect of speech encryption, most of them use the discrete orthogonal transforms like discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete Hadamard transform (DHT) to perform the speech encryption and decryption in frequency domain. Though this approach has good performance in the speech encryption, it needs large computation load in carrying out the forward and inverse transforms. Thus, a high performance digital signal processor is needed to realize the transformation in the real-time voice applications, which indirectly increases the cost. On the other hand, in the aspect of error recovery of the speech signals based on G.729, the existing approaches [8,9] can be categorized as follows. (1) *Concealment.* The concealment strategy that has already been encapsulated in the decoder can be used by itself. In this way, the decoder can deal with packet loss to some degree without requiring any extra work from either the network or the transmitter. (2) *Retransmission.* The decoder can inform the encoder of errors, and the encoder can retransmit the lost frames. This approach can yield perfect error performance, but with slightly higher bandwidth requirements, and substantially larger delays. (3) *Resynchronization.* The decoder can inform the encoder of errors, but instead of resending the lost frame, the encoder reinitializes its state, and informs the decoder about the re-initialization in the next packet. This approach does not help restore state for lost frames any more than simple concealment. However, it helps restrict the amount of time for which the encoder and decoder states will be different. The utility of this approach depends on the difference between round trip times in the network and the amount of time it takes for encoder and decoder to resynchronize unaided. If the natural resynchronization time is less than the network roundtrip time, this approach is useless. If it is substantially more, it is very helpful. (4) *Forward Error Correction.* This approach is really a class of algorithms. The idea is to send extra information at the encoder that can assist the decoder in recovering lost frames. The possible FEC approaches include redundant encoding [10, 12], channel coding [11], and diversity. (5) *Interleaving.* The data in N consecutive frames can be mixed together before transmission. In this way, loss of a packet destroys only a few bits from each frame. Assuming the coder is more robust to bit errors than frame erasures (which is generally true), this approach may lessen the effect of loss. However, it does so at the expense of the substantial delays. Of course, all of the approaches can be combined to yield any number of a variety of hybrids. Each can also be made adaptive, varying the amount of delay, redundancy, etc. as network losses vary. In general, each of the above approaches is a tradeoff among a number of different variables, which include bandwidth, delay, computational complexity, and speech quality. Which one is best is dependent on the importance placed on these four factors.

For solving the problem analyzed above, we propose a new VoIP technique combining the speech data encryption and G.729 error recovery. This technique uses the chaotic-based Inter-frame data interleaving to make the situation of continuous packet loss being isolated to become an isolated packet loss situation. Then, we try to propose a Periodical Parameter Re-initialization (PPR) recovery approach to reduce the signal quality degradation in the G.729 decoder due to the lost of state synchronization to the G.729 encoder. Besides, the proposed VoIP technique also uses the idea of chaotic Intra-frame data encryption to scramble the data sequence within a speech frame. So, even if the scrambled speech data are stolen in the transmission process, they cannot be recovered back through the original speech decoder. This provides a protection scheme for the transmission of important speech data. In order to efficiently realize the proposed chaotic-based Inter-frame data interleaving and Intra-frame data encryption, we adopt a set of parameters derived from a Chaos system [13,14] to control the operations of data interleaving and data encryption. The features of the proposed VoIP technique include high security, low computational complexity, real-time processing capability, and low extra overhead. These features make the proposed technique suitable for real-time voice over IP applications. The simulation results show that using the proposed VoIP technique can have averagely 20 db gain as compared with the original G.729 CODEC with respective to different packet loss rates ranging from 5% to 30%. In a word, using the proposed VoIP technique can both maintain good voice quality even in the continuous packet loss situation and possess good data security.

2. THE PROPOSED VOIP TECHNIQUE

Fig. 1 shows the simplified block diagram of the proposed VoIP technique combining the speech data encryption and G.729 error recovery. It consists of two major parts. One is the Periodic Parameter Re-initialization (PPR) on the G.729 encoder. The other is the chaotic-based inter-frame data interleaving and intra-frame data encryption controlled by using the parameters generated from a Chaotic system. In the following, we will respectively illustrate the above-mentioned two key methods adopted in the proposed VoIP technique.

2.1 PPR on the G.729 speech encoder

Fig. 2 shows the possible state variables used in the G.729 periodic parameter re-initialization (PPR) scheme. The basic idea in applying the PPR scheme is to minimize the signal quality degradation in the G.729 decoder due to the lost of state synchronization to the G.729 encoder. The reason for this prolonged period of error is that the states of the encoder and decoder are no longer synchronized after a period of packet loss. What is meant by state in this paper is all memory that resides in both encoder and decoder that is never exchanged between them but rather constantly updated to ensure they stay the same at both sides. The coder generates a set of parameters to model the speech based on its present states. These parameters are then transmitted to the decoder that reproduces the speech from these parameters and the locally generated states. Under normal operations, both sets of states are updated to remain the same so that the original speech is reproduced from the companion set of transmitted parameters. During a packet loss, the states in the encoder and decoder are no longer synchronized because of the receiver’s inability to properly update its states. Thus, even when no more packets are lost, the decoder is using erroneous states to reproduce speech from the correctly received packets. This clearly results in a significant distortion in the speech.

Therefore, we propose an idea of periodic parameter re-initialization (PPR) scheme that periodically re-initialize the important parameters shown in Fig. 2 in the G.729 voice CODEC to minimize the signal quality degradation in the G.729 decoder due to the lost of state synchronization to the G.729 encoder under an IP network with packet loss. The period of the re-initialization scheme is related to the packet loss rates in the network, which has different effects on reducing the speech degradation. This period can be obtained from the simulation results.

2.2 Chaotic-based inter-frame data interleaving and intra-frame data encryption

Fig. 3 shows the operations of the proposed chaotic-based inter-frame data interleaving and the intra-frame data encryption in the proposed VoIP technique. These operations are controlled by a chaotic bit-string generated from the chaotic bit-string generator (CBSG) based on the chaotic system of 1-D logistic map [13]. Using the chaotic bit-string in data interleaving and data encryption possesses the features of low computational complexity, high security, and no distortion. These features have been illustrated in our previous works [15-17] in the image encryption applications. For the chaotic-based inter-frame data interleaving, we adopt frame re-ordering techniques controlled by the permutation control unit (PCU) based on the chaotic bit-string. In addition, for the chaotic-based intra-frame data encryption, we adopt the pixel value transformation technique including bit swapping and XOR operations, controlled by the value transformation control unit (VTCU) based on the chaotic bit-string to translate the original pixel value into an encrypted one. Using these two techniques can both achieve high data security and avoid the severe quality degradation caused by the continuous packet loss.

2.2.1 Notations and definitions

For clearer illustration, we first define the notations and definitions as follows.

g the voice of size N sample points,

$g(n)$ the one-byte value of g at n, $0 \le n \le N-1$,

g' the result of g after applying the inter-frame data interleaving,

g'' the result of g' after applying the intra-frame data encryption,

framesize the size of a frame,

frameno the total frame number $\lfloor N / framesize \rfloor$ in g, where $\lfloor y \rfloor$ denotes the smallest integer that is smaller than or equal to y,

frame(i) the i^{th} frame in g , $0 \le i \le frameno-1$,

interleaveno the number of frames to be regarded as a group. In each group of g, the inter-frame data interleaving is applied locally.

Definition 1: The operation *SwapFrame*(p, q) is defined as to swap *frame*(p) and *frame*(q).

Definition 2: The operation $SwapBit_w(d_r, d_s)$ is defined as to swap bit d_r and bit d_s if w is equal to 1 or preserve their original values if w is equal to 0.

2.2.2 Data interleaving and data encryption

Step 1: Set the parameters N, *framesize*, and *interleaveno*. Compute $L = \lfloor \log_2(interleaveno) \rfloor$, $diff = (interleaveno - 2^L)$, and $frameno = \lfloor N / framesize \rfloor$.

Step 2: Determine the parameter μ and initial point $x(0)$ of the 1-D logistic map [10], $f_\mu(x) = \mu x(1-x)$, where μ should be selected as the values that can result in chaos and $0 < x(0) < 1$. Evolve successive states from the map by $x(n+1) = \mu x(n)(1-x(n))$, and the preceding sixteen bits below the decimal point of the binary representation of $x(n)$, n = 1, 2, …, are extracted to constitute the chaotic bit-string sequence $b(0)$, $b(1)$, $b(2)$, ….

```
Step 3: For $ii = 0$ To $\lfloor frameno / interleaveno \rfloor - 1$
    For $jj = 0$ To $interleaveno - 1$
        $offset = (ii \times interleaveno + jj) \times L;$                                  (1)
        $no = \sum_{i=0}^{L-1} 2^i \times b(offset + i) + \sum_{j=L}^{diff+L-1} b(offset + j)$;          (2)
        $SwapFrame(ii \times interleaveno + jj, ii \times interleaveno + no)$;              (3)
     End
     End
Step 4: For $kk = 0$ To $N - 1$
        Let $g'(kk) = \sum_{i=0}^{7} d_i \times 2^i$;
        For $ll = 0$ To 3
            $SwapBit_{b(4\times kk+ll)}(d_{ll}, d_{ll+4})$;                                  (4)
        End
        For $mm = 1$ To 7 Step 2
            $d_{mm}' = d_{mm}$ XOR $b(4\times kk+mm)$;                                    (5)
        End
        $g''(kk) = \sum_{i=0}^{7} d_i' \times 2^i$;

    End
```

Step 5: The encryption result *g*’’ is obtained and the algorithm is terminated.

3. SIMULATION RESULTS

In verifying the proposed VoIP technique, we have performed simulation results on it under the simulation environment like that shown in Fig. 1. In this environment, we would like to demonstrate the performance of the proposed VoIP technique combining the G.729 PPR and chaotic-based data interleaving. In the simulation, we have made the following assumptions:

(a) The speech CODEC used in the simulation is G.729.

(b) The packet loss is simulated by randomly dropping the desired number of packets according to the packet loss rate.

(c) The packet loss distribution includes single packet loss, continuously two-packet loss, three-packet loss, and four-packet loss.

(d) The period (*interleaveno*) used in the proposed chaotic data interleaving technique is assumed to be 10-frame, 15 frames, and 20 frames.

(e) The frame period of the G.729 is assumed to be 10ms.

In the following, we will illustrate the simulation results of the proposed VoIP technique in several aspects.

3.1 The simulation results of the proposed G.729 PPR

Fig. 4 shows the simulation results of the proposed G.729 PPR. Fig. 4(a) shows the mean SNR of different re-initialization periods with respect to different packet loss rates from 5% to 30%. From this result, we can select a more suitable re-initialization period N in the proposed G.729 PPR scheme. From the experiment results, we find that N=20 is good for different packet loss rates. Therefore, we use N=20 in the proposed G.729 PPR scheme in the following simulation. Fig. 4(b) shows the mean SNR distribution with and without the G.729 PPR scheme with respect to different packet loss rate in a single packet loss situation. From the simulation results, we conclude that using the G.729 PPR scheme can have about 4db gain in the single packet loss situation. Fig. 4(c) shows the mean SNR distribution with and without the G.729 PPR scheme with respect to different packet loss rate in a continuously two-packet loss situation. From the simulation results, we find that using the proposed G.729 PPR scheme can have about 2db gain. Fig. 4(d) shows the mean SNR distribution with and without the G.729 PPR scheme with respect to different packet loss rate in a continuously three-packet loss situation. From the simulation results, we find that the proposed G.729 PPR is only good for lower packet loss rates, i.e. 5% and 10%, but it is not good for the higher packet loss rates, i.e. 15%, 20%, 25%, and 30%. From the simulation results shown in Fig. 4, we summarize the results in the following.

(a) The proposed G.729 PPR scheme performs better in a single packet loss situation than continuously multiple-packet loss situations. This fact gives us a motivation to use chaotic data interleaving to scramble the normal speech frames to reduce the continuously multiple-packet loss situation and translate them to isolated packet loss situation.

(b) The re-initialization period N in the proposed G.729 PPR scheme is suggested to be about 20 under different packet loss rates.

3.2 The simulation results of the proposed chaotic data interleaving

Using chaotic data interleaving is to make the continuously multiple-packet loss situation to become isolated packet loss situation by scrambling the speech packets after encoding process. When the decoder receives the scrambled speech packets, it will perform the data de-interleaving process to covert the speech packets to be in normal order before decoding process. For verifying the performance of the chaotic data interleaving operation, we perform software simulation on a sequence of test voice with predefined packet loss rates of 10%, 20%, 30%, 40%, and 50%. In the simulation, we set the parameters N = 1600000, *framesize* = 80, and *frameno* = 20000. We assume one voice packet contains one voice frame. We randomly select certain percentages of frames in the packet loss, and simulate the cases of maximally continuous 10, 15, and 20 packets lost in the original voice sequence. In each case, we perform the simulation with *interleaveno*=10 and *interleaveno*=15 under 10 %, 20%, 30%, 40%, and 50% packet loss rates. We summarize the simulation results in Fig. 5. Fig.5(a), Fig.5(b), and Fig.5(c) respectively show the simulation results with maximally continuous 10, 15, and 20 packets lost. In each figure, there are five groups of results that respectively show the simulation under different packet loss rates, i.e. 10%, 20%, 30%, 40%, and 50%. In each group of results, there are three data bars. The first one indicates the packet loss distribution of the original voice sequence. The second one and the third one respectively show the packet loss distribution of the voice sequence when the proposed chaotic data interleaving scheme with *interleaveno*=10 and *interleaveno*=15 is applied. Besides, Fig.5(d) shows the comparison of the packet loss distribution in different lengths of maximally continuous lost packets under 30% packet loss rate. From the results in Fig. 5, we summarize the observations in the following.

(a) When the proposed chaotic data interleaving is applied, the occurrence of single or continuously two packets loss situations increase dramatically under different values of *interleaveno*, and different lengths of maximally continuous packet loss. The proposed chaotic data interleaving has the capability to reduce the long continuously packet loss situations and translate them into isolated packet loss situations, which is beneficial to the proposed G.729 PPR scheme.

(b) Larger values of *interleaveno* as compared with the lengths of maximally continuous packet loss have better results to reduce the long continuously packet loss situations. However, larger values of *interleaveno* introduce longer round-trip delay in real-time voice processing. We have to consider this factor in choosing the values of *interleaveno.* In the following simulation, we use the *interleaveno*=10 for reducing the round-trip delay as most as possible.

3.3 The simulation results of the proposed VoIP technique combining the G.729 PPR and chaotic data interleaving

From the simulation results shown in Fig. 4 and Fig. 5, we find that the proposed G.729 PPR scheme performs better in the single packet loss situation. And the proposed chaotic data interleaving can effectively translate the continuously multiple-packet loss situation to single packet loss situation. So, by combining the proposed G.729 PPR scheme and the chaotic data interleaving, the proposed VoIP technique will have better performance than we use them individually. Fig. 6 shows the simulation results of the proposed VoIP technique combining the G.729 PPR scheme and the chaotic data interleaving. We find that using the proposed VoIP technique can have much better results than using G.729 PPR only. The main reason is that we can use chaotic data interleaving first to translate the continuously multiple-packet loss situation into isolated packet loss situation such that the G.729 PPR can perform well in this situation. The simulation results show that using the proposed VoIP technique can have averagely 20 db gain as compared with the original G.729 CODEC under different packet loss rates from 5% to 30%.

3.4 The simulation results of the chaotic data encryption

In order to demonstrate the encryption performance of the proposed VoIP technique, we perform the statistical measures on six voice test sequences. As a representative, only the voice “Sound1” is shown in Fig. 7(a). All the test voices are encrypted by the chaotic intra-frame data encryption without considering the packet loss problem. The encrypted result of “Sound1” is shown in Fig. 7(b). Then, we compute the autocorrelation function and fractal dimension of each voice and its encrypted one. The short-time autocorrelation function is defined as follow:

$$R_n(k) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)x(m+k)w(n-m-k), \tag{6}$$

where w is a window sequence positioned at a time corresponding to sample index n [18]. This function is usually used to estimate the pitch period of voiced speech. The short-time autocorrelation functions of the 10th frame of the original and encrypted voices are shown in Fig. 8(a) and 8(b), respectively. Here, rectangular window with window width 400 is used. In fact, the quality $R_n(0)$ is the energy of the windowed voice. However, we normalize the short-time autocorrelation function so that $R_n(0)=1$. From the waveforms shown in Fig. 7, the encrypted voice is randomly distributed. So, it is difficult to recognize the contents of voice after encryption. This can also be verified by the quantitative measures of the autocorrelation function and fractal dimension. For the autocorrelation function of the original voice, shown in Fig 8(a), the local peaks are at about multiples of 43 samples, indicating the average pitch period of about 5.4 ms. For the encrypted voice, shown in Fig 8(b), all values are small except $R_n(0)$. There are no strong periodicity peaks. Based on our experimental results, we found that most of the encrypted voices have similar results. Obviously, the encrypted voice has less correlation among voice samples than that of the original voice.

4. CONCLUSION

We have proposed a Voice over Internet Protocol (VoIP) technique combining the speech data encryption and G.729 error recovery. This technique uses the Inter-frame data interleaving to make the situation of continuous packet loss becoming an isolated packet loss situation. Then, we propose a Periodical Parameter Re-initialization (PPR) recovery approach to reduce the signal quality degradation in the G.729 decoder due to the lost of state synchronization to the encoder. Besides, the proposed VoIP technique also uses the idea of Intra-frame data encryption to scramble the data sequence within a speech frame. The simulation results show that using the proposed VoIP technique can have averagely 20 db gain as compared with the original G.729 CODEC with respective to different packet loss rates ranging from 5% to 30%.

5. ACKNOWLEDGEMENTS

This work was supported by the National Science Council under grant: NSC90-2213-E-194-059.

6. REFERENCES

[1] Lenore V, T.: 'Voice over IP: turning up the volume', Telecommunications, March 1998, pp.28.

[2] Robin, G.: 'Voice over IP services: the sound decision', Data Communications, March 1998, pp.75.

[3] Sridharan, S., Dawson, E., and Goldburg, B.: 'Speech encryption in transform domain', Electronics Letters, 1990, 26(10), pp.655-657.

[4] Matsunaga A., Koga, K., and Ohkawa, M.: 'An analog speech scrambling system using the FFT technique with high level security', IEEE J. Sel. Areas Commun., 1989, SAC-7, pp.540-547.

[5] Sakurai, K., Koga, T., and Muratani, T.: 'A speech scrambler using fast Fourier transform techniques', IEEE J. Sel. Areas Commun., 1984, SAC-2, pp.434-442.

[6] Milosevic, V., Delic, V., and Senk, V.: 'Hadamard transform application in speech scrambling', 13th Internal Conference on Digital Signal Processing, 1997, pp.361-364.

[7] Sridharan, S., Dawson, E., and Goldburg, B.: 'Fast Fourier transform based speech encryption system', IEE Proceedings-I, 1991, 138 (3), pp.215-223.

[8] J. C. Bolot, "End-to-end packet delay and loss behavior in the Internet," Proc. ACM Sigcomm'93, pp.289-298, CA, Sept. 1993.

[9] D. Sanghi, A. Arawala, O. Gudmundsson, B. Jain, "Experimental Assessment of End to end behavior in the Internet," IEEE Infocom 1993, CA, Mar. 30-April 1, pp.867-874, 1993.

[10] J. C. Bolot, A. Vega Garcia," Control mechanisms for packet audio in the Internet," IEEE Infocom'96, CA, March 1996.

[11] N. Shacham, P. McKenny," Packet recovery in high speed networks using coding and buffer management," Proc. infocom'90, pp.124-131.

[12] V. Hardman, M. Sasse, M. Handley, and A. Watson, "Reliable audio for use over the Internet," Proc. INET'95.

[13] Parker, T. S. and Chua, L. O.: 'Chaos - A Tutorial for Engineers', Proceedings of The IEEE, 1987, 75 (8), pp. 982-1008.

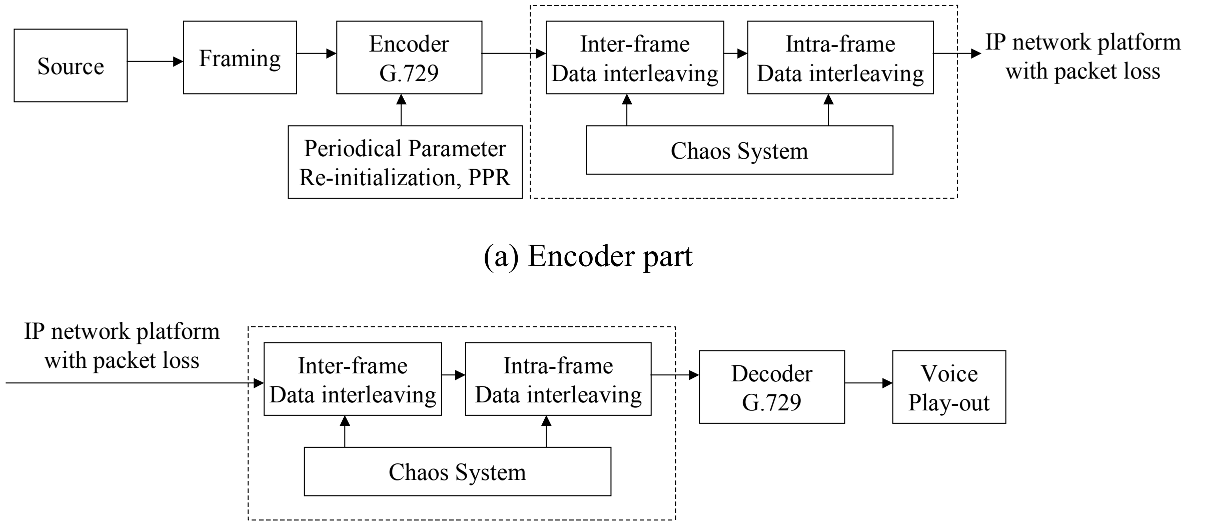
[14] Wu, C. W. and Rulkov, N. F.: 'Studying Chaos via 1-D Maps - A Tutoria, IEEE Transactions on Circuits and Systems I-Fundamental Theory and Applications, 1993, 40 (10), pp. 707-721.

[15] Yen, J. C. and Guo, J. I.: 'A New Hierarchical Chaotic Image Encryption Algorithm and its VLSI Realization', IEE Proceedings, Vision, Image and Signal Processing, 2000, 147 (2), pp. 167-175.

[16] Yen, J. C. and Guo, J. I.: 'A new MPEG/encryption system and its VLSI architecture', 1999 International Symposium on Communications, Nov. 1999, Kaohsiung, Taiwan, pp.215-219.

[17] Yen, J. C. and Guo, J. I.: 'A New Chaotic Key-Based Design for Image Encryption and Decryption', 2000 IEEE International Symposium on Circuits and Systems, May 2000, GENEVA, SWITZERLAND, pp.IV-49~IV-52.

[18] Rabiner, L. R. and Schafer, R. W.: 'Digital Processing of Speech Signals', Prectice-Hall Inc.



(a) Encoder part

(b) Decoder part

Fig.1: The simplified block diagram of the proposed VoIP technique combining the speech data encryption and G.729 error recovery.

State #	Size (bits)	Description
1	640	Memory of 4th order Moving Average (MA) predictor for LPC
2	2464	Past overall excitation values needed to build adaptive codebook vector
3	160	Memory for LP synthesis filter
4	64	Memory of 4th order MA predictor for fixed codebook gain
5	16	Past adaptive codebook gain used by fixed codebook pre-filter

Fig.2: The possible states used in the G.729 periodic parameter re-initialization (PPR) scheme.

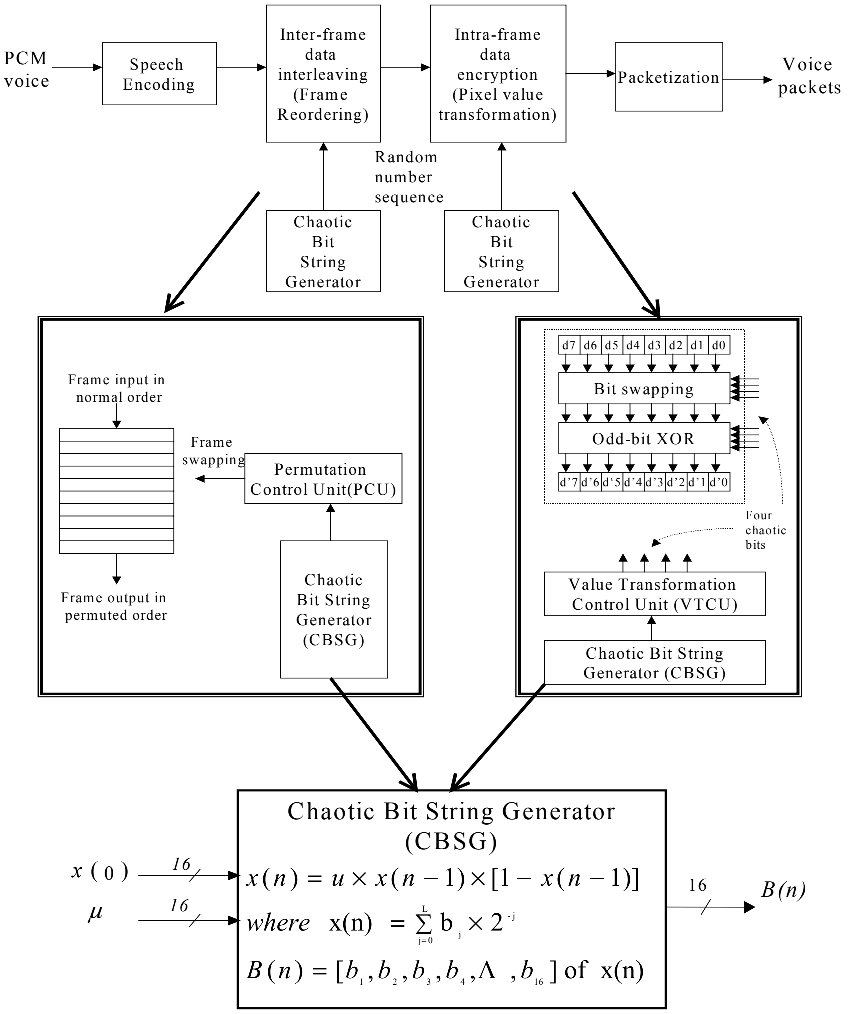
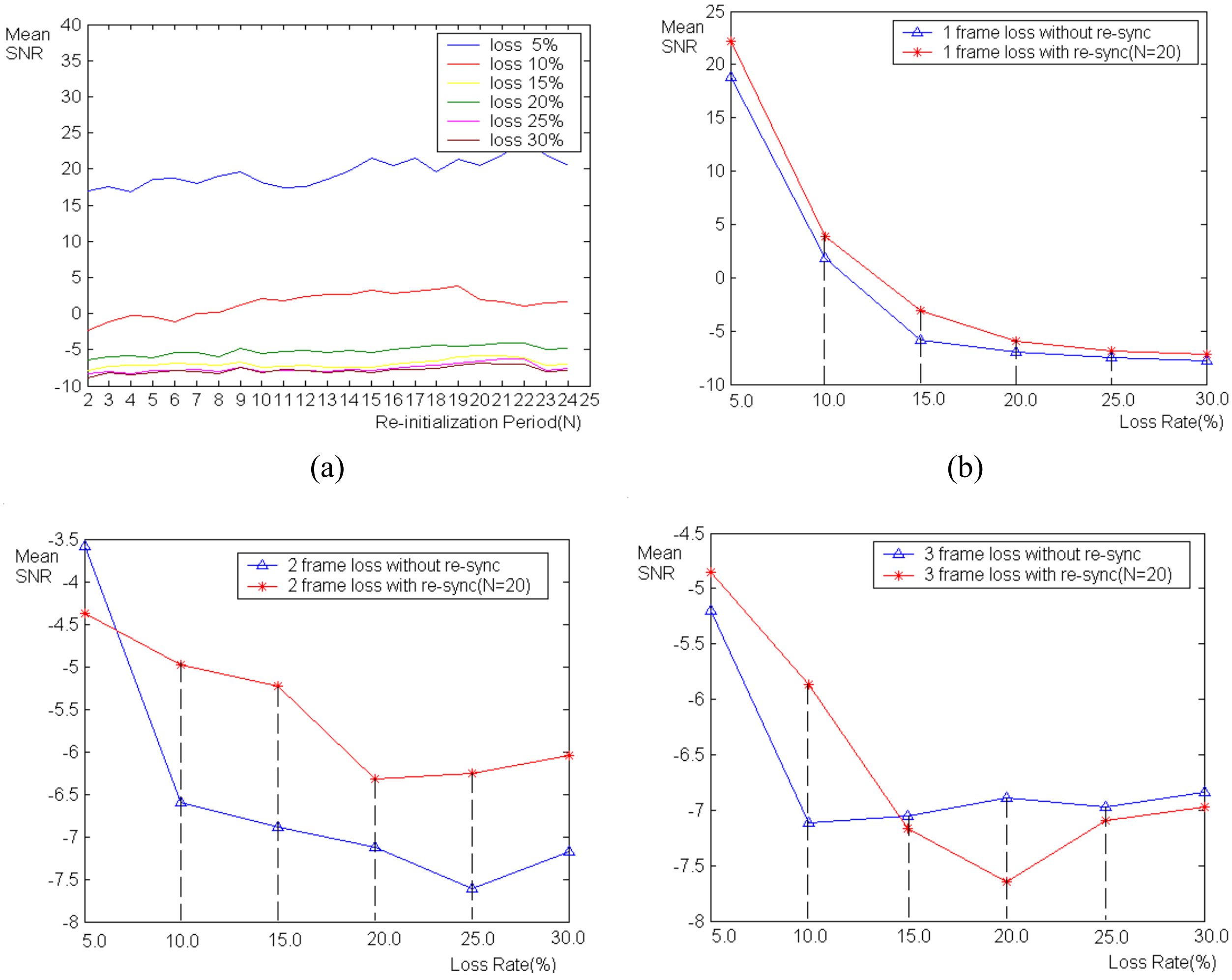

PCM voice
Speech Encoding
Inter-frame data interleaving (Frame Reordering)
Intra-frame data encryption (Pixel value transformation)
Packetization
Voice packets
Random number sequence
Chaotic Bit String Generator
Chaotic Bit String Generator
Frame input in normal order
Frame swapping
Permutation Control Unit(PCU)
Frame output in permuted order
Chaotic Bit String Generator (CBSG)
d7 d6 d5 d4 d3 d2 d1 d0
Bit swapping
Odd-bit XOR
d'7 d'6 d'5 d'4 d'3 d'2 d'1 d'0
Four chaotic bits
Value Transformation Control Unit (VTCU)
Chaotic Bit String Generator (CBSG)
Chaotic Bit String Generator (CBSG)
$x(0)$
μ
16
16
$x(n) = u \times x(n-1) \times [1 - x(n-1)]$
where $x(n) = \sum_{j=0}^{L} b_j \times 2^{-j}$
$B(n) = [b_1, b_2, b_3, b_4, \Lambda, b_{16}]$ of x(n)
16
B(n)

Fig. 3. The operations of the inter-frame data interleaving and intra-frame data encryption in the proposed VoIP technique.



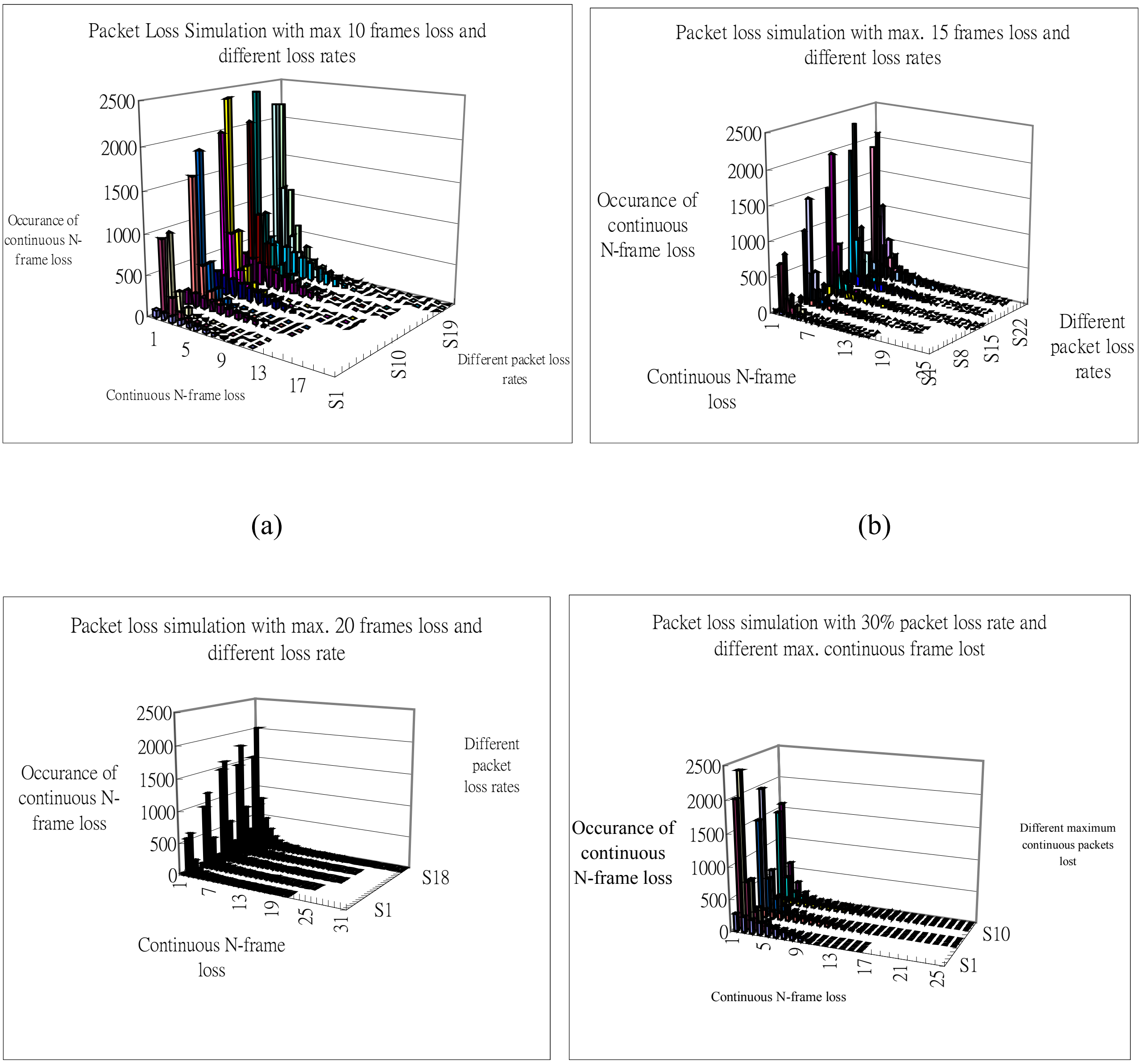
(a)

(b)

(c)

(d)

Fig. 4. The simulation results of the G.729 PPR scheme; (a) The mean SNR of different re-initialization periods with respect to different packet loss rates from 5% to 30%; (b) The mean SNR distribution with/without the G.729 PPR scheme with respect to different packet loss rate in a single packet loss situation; (c) The mean SNR distribution with/without the G.729 PPR scheme with respect to different packet loss rate in a continuously two-packet loss situation; (d) The mean SNR distribution with/without the G.729 PPR scheme with respect to different packet loss rate in a continuously three-packet loss situation.



Packet Loss Simulation with max 10 frames loss and different loss rates

(a)

Packet loss simulation with max. 15 frames loss and different loss rates

(b)

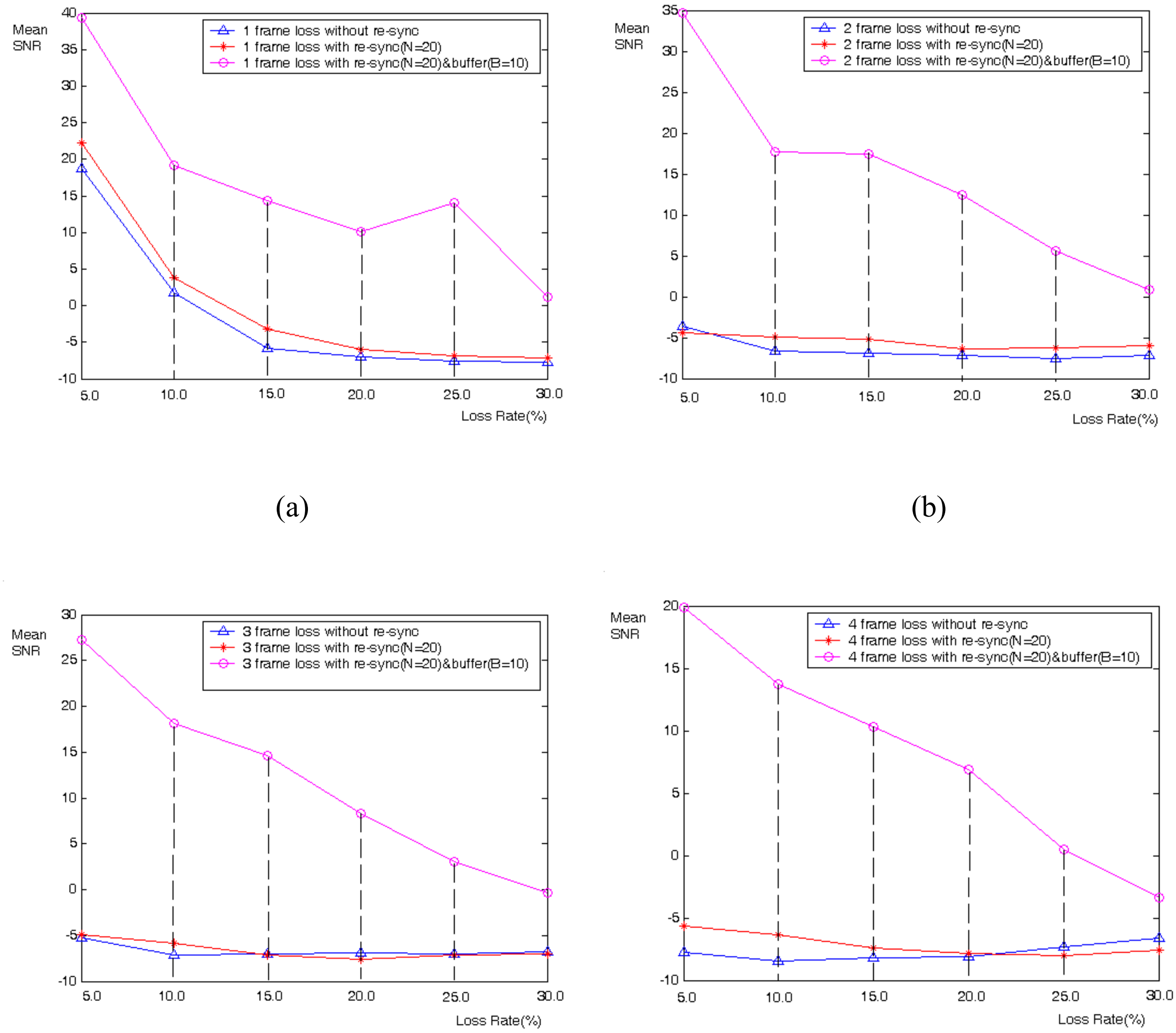
Packet loss simulation with max. 20 frames loss and different loss rate

(c)

Packet loss simulation with 30% packet loss rate and different max. continuous frame lost

(d)

Fig. 5. The simulation results of the packet loss distribution with different packet loss rates (10%, 20%, 30%, 40% and 50%) when the proposed inter-frame data interleaving technique with *interleaveno* equal to 10 and 15 are applied. (a) the simulation result of maximally 10 continuous packets are lost; (b) the simulation result of maximally 15 continuous packets are lost; (c) the simulation result of maximally 20 continuous packets are lost;(d) the simulation result of different maximally continuous lost packets (10, 15, and 20) under 30 % packet loss rate.



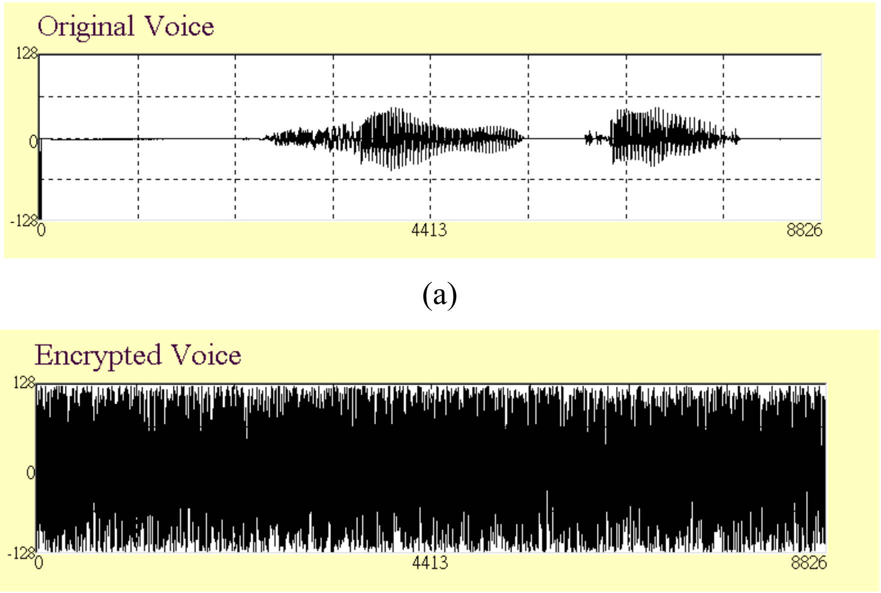
(a)

(b)

(c)

(d)

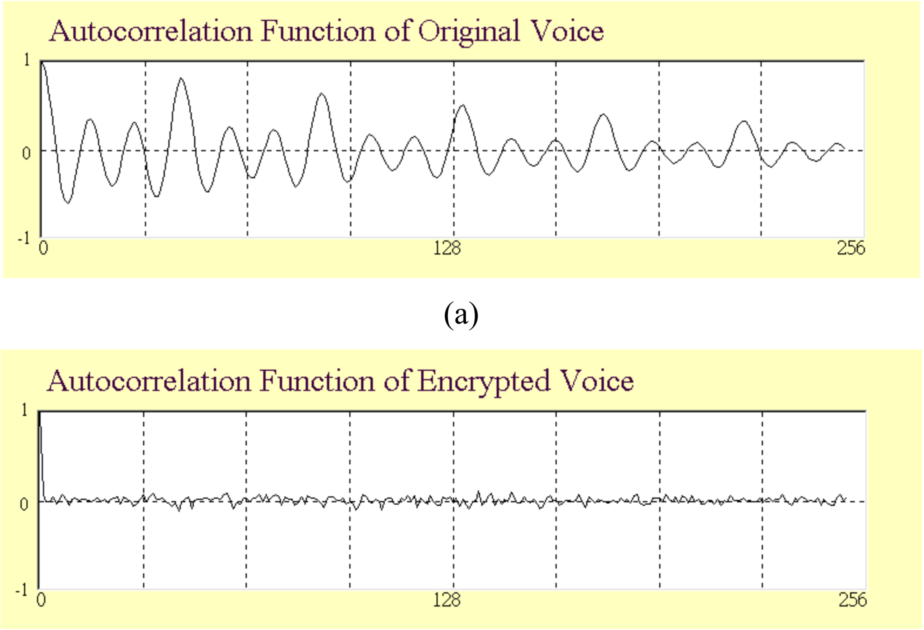
Fig. 6. The simulation results of the proposed VoIP technique combining G.729 PPR scheme and chaotic data interleaving with buffer size B (*interleaveno*)=10 frames. (a) The mean SNR distribution of the proposed VoIP technique/with G.729 PPR only/without G.729 PPR with respect to different packet loss rate in a single packet loss situation; (b) The mean SNR distribution of the proposed VoIP technique/with G.729 PPR only/without G.729 PPR with respect to different packet loss rate in a continuously two-packet loss situation; (c) The mean SNR distribution of the proposed VoIP technique/with G.729 PPR only/without G.729 PPR with respect to different packet loss rate in a continuously three-packet loss situation; (d) The mean SNR distribution of the proposed VoIP technique/with G.729 PPR only/without G.729 PPR with respect to different packet loss rate in a continuously four-packet loss situation.



(a)

(b)

Fig. 7. The waveforms of voice (a) the original one; (b) the encrypted one when the proposed intra-frame data encryption scheme is applied.



(a)

(b)

Fig. 8. The autocorrelation functions of the 10th frame of (a) the original voice; (b) the encrypted voice.