論文 / 著書情報
Article / Book Information

Title	Modified Delta Encoding and Its Applications to Speech Signal
Author	Yanna Ma, Hisayori Noda, Izumi Ito, AKINORI NISHIHARA
Journal/Book name	IEEE Region 10 Conference TENCON, , ,
発行日 / Issue date	2010, 11
権利情報 / Copyright	(c)2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Modified Delta Encoding and Its Applications to Speech Signal

Yanna MA, Hisayori NODA, Izumi ITO and Akinori NISHIHARA
Communications and Integrated Systems
Tokyo Institute of Technology
Tokyo, Japan
mayanna@nh.cradle.titech.ac.jp

Abstract—**A modified delta encoding method and its applications to speech signal are proposed. In this method, Hungarian algorithm is applied to find the minimum distance of arbitrary two frames in speech signal and minimum spanning tree is used to find an effective delta encoding path. In simulation, the method is applied to the compression of sinusoidal coding. The results show that the data size after compression is 6% smaller than a usual delta encoding whose path is not suitably permutated. In addition, the proposed method has the potential to apply on data security for practical use because the delta encoding path which can be used as the security key is unordered and long enough.**

Keywords-Delta Encoding; Minimum Spanning Tree; Hungarian Algorithm; Sinusoidal Coding; Data Security

I. INTRODUCTION

Delta encoding is the technique of storing or transmitting data as the difference between successive samples rather than samples themselves. There are many practical applications where the new information received is similar to what is already existed [1]. It can be used in data compression and performs well when values of the encoding data are similar between adjacent data, but loses its efficiency when the encoding data are unsorted. The selection of the delta encoding path is important and can be modeled by an optimization problem. There is no such method that considering delta encoding in terms of graph theory. In our proposed method, the path of encoding is selected by integrating minimum spanning tree (MST) and Hungarian algorithm.

MST is a spanning tree whose total weight is minimized. Hungarian algorithm, on the other hand, is an optimization algorithm for solving assignment problem. It was first proposed by Harold Kuhn in 1955 [2]. The computational time of the original algorithm was $O(n^4)$, and then reduced to $O(n^3)$ by Edmonds and Karp [3]. Sinusoidal coding is a parametric representation that utilizes only sinusoidal components [4].

In the proposed method, a sinusoidal model for the speech waveform is used to develop a new analysis/synthesis technique that is characterized by the amplitudes, frequencies, and phases of the component sinusoids. From the distance between two sinusoids for all frames the weighted graph is determined, which contributes to efficient permutations of delta. In simulation, the proposed method was applied to the compression of sinusoidal coding. The results show that the data size after compression is 6% smaller than normal delta encoding without permutations.

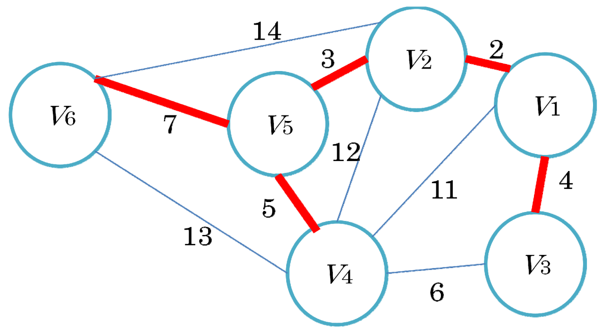

Figure 1. $V1$-$V6$ are vertices and each edge is labeled with its cost. The MST is the sub-graph connected by the red lines

Section II gives some basic knowledge of the proposed method. The modified delta encoding for sinusoidal coding is described in Section III. Section IV discusses the simulations and the results. Finally, Section V includes the conclusions of this paper.

II. PRELIMINARIES

In this section, delta encoding, MST, Hungarian algorithm and sinusoidal coding are described. $\boldsymbol{N}$ denotes the set of natural numbers.

A. Delta encoding

Delta encoding is a way to store data in form of difference between consecutive data rather than data themselves, it can be described as follow:

$$\Delta(V_i, V_{i+1}) = V_i - V_{i+1} \quad (1)$$

where V_i and V_{i+1} represent two successive data.

B. Minimum Spanning Tree (MST)

Given a connected and undirected graph G=(V,E,W), where V is the set of vertices, E is the set of edges between pairs of vertices and $W(e_{ij})$ specifies the weight of each edge e_{ij} between vertex V_i and V_j, $i, j \in \boldsymbol{N}$.

A complete graph is a graph in which every pair of vertices is connected by a unique edge. A spanning tree is a sub-graph without cycle and contains all the vertices in G. MST is a

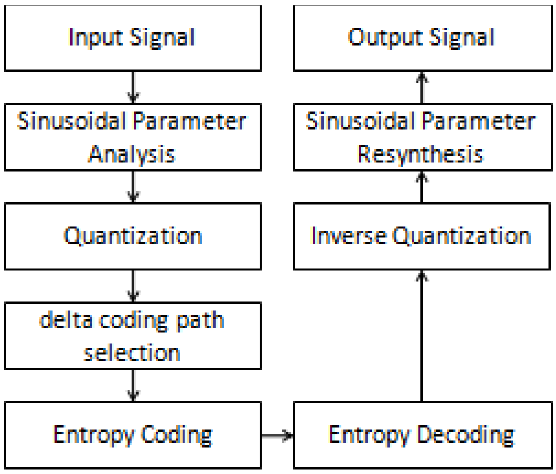

Figure 2. The diagram of sinusoidal coding

spanning tree of a weighted graph with minimum total edge weight. Prim's algorithm is commonly used for finding MST [5]. The algorithm starts from an arbitrary vertex of the complete graph, in each step, the tree grows by merging the smallest edge. Repeat this until all vertices are connected, that is the MST. Fig.1 shows an example of spanning tree and MST, vertices *V1-V6* are connected and labeled with weight, the sub-graph connected by the red lines is an MST.

C. Hungarian algorithm for assignment problem

Assignment problem is a problem aims to find the assignment which minimize (or maximize) the total cost in a weighted graph. Given an assignment problem, convert it to an $n \times n$ matrix $M = (d_{ij})$, $d_{ij} \geq 0$, which expresses the cost of assigning i to j. Let x_{ij} be the *i-j* assignment we want to obtain. The standard form can be expressed as follows:

$$\text{minimize} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij}\, x_{ij}$$
$$\text{subject to}$$
$$\sum_{j=1}^{n} x_{ij} = 1 \quad (i=1,2,\ldots,n) \qquad (2)$$
$$\sum_{i=1}^{n} x_{ij} = 1 \quad (j=1,2,\ldots,n)$$
$$x_{ij} \in \{0,1\}$$

The Hungarian algorithm can be briefly described by the following 4 steps:

1. If necessary, convert the maximum assignment problem to the minimum assignment problem. This conversion can be achieved by replacing each d_{ij} with N-d_{ij}, where $N = max(d_{ij})$.
2. Find out the minimum value of each row and column, subtract it off.
3. Use as few lines as possible to cover all the zeros in the matrix, suppose there are k lines.
   - If $k<n$, find the minimum value of the uncovered elements and subtract it off from all uncovered elements, at the same time, add the minimum value to elements that covers two lines. Go back to the start of 3.
   - If $k=n$, go to 4.
4. Make assignments from the top row, delete that row and column after the assignment is made. If not all assignments can be made and the remaining rows have more than one zero, shift to the column and start from the left to the right until all assignments are made.

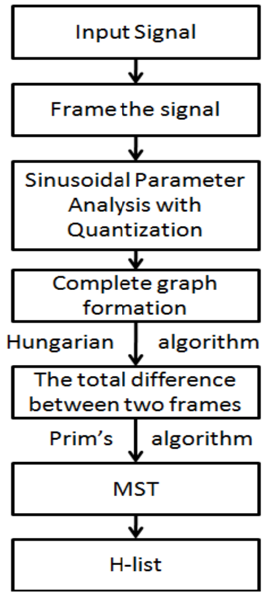

Figure 3. The diagram of the modified delta encoding

D. Sinusoidal coding

The diagram of sinusoidal coding is shown in Fig.2. In the sinusoidal parameter analysis model, the input signal is divided into time frames, and each frame is characterized by the amplitudes, frequencies and phases of the component sinusoids. These parameters are estimated from the short-time Fourier transform using a simple peak-picking algorithm.

In quantization model, the sinusoidal parameters are quantized using auditory scales. The amplitude and frequency are quantized by log scale and mel scale, respectively, while the phase is quantized without auditory scale. In delta coding path selection part, an effective path is selected using the proposed method. In the entropy coding model, the whole data can be encoded by entropy encoding method like Huffman coding [6] and Range Coder [7].

In conventional delta encoding, data are encoded according to their sequence only. Conversely, in our method, the data are rearranged and modeled into a graph.

III. Modified Delta Encoding Method for Sinusoidal Coding

An efficient modified delta encoding method using Hungarian algorithm and MST is proposed in this section.

A. Modified Delta Encoding

Our method achieves more efficient compression ratio by permutating the elements of delta value between signal frames. The permutation is selected by MST, where the Hungarian

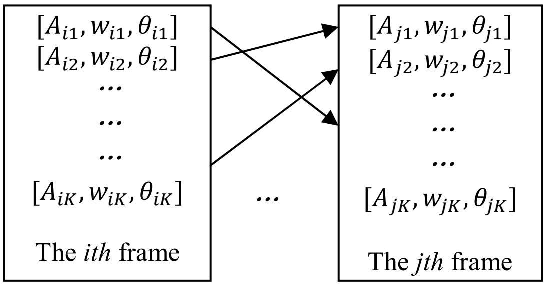

Figure 4. Assignment problem solved by Hungarian algorithm, *k* is the total number of sinusoids in one frame

algorithm is used for calculating the distance between two frames.

Fig.3 gives the detailed model of the modified delta encoding method. The *i*th frame $s_i(n)$ of the input signal can be modeled as the sum of a bank of sinusoidal oscillators [8]:

$$s_i(n) = \sum_{k=1}^{K} A_{ik} \cos(w_{ik} n + \theta_{ik}) \quad (3)$$

where $\{K\}$ is the number of sinusoids in one frame and $\{A_{ik}\}$, $\{w_{ik}\}$ and $\{\theta_{ik}\}$ represent the amplitude, frequency and phase of the *k*th sinusoid in *i*th frame , respectively.

For each frame, it is an unordered data set formed by sinusoids. The distance, d_{ij}, between two sinusoids for all frames is defined as the squared distance of the three extracted parameters.

$$d_{ij} = (A_{ik} - A_{jk})^2 + (w_{ik} - w_{jk})^2 + (\theta_{ik} - \theta_{jk})^2 \quad (4)$$

The distance between two frames is modeled into an assignment problem and obtained by using Hungarian algorithm.

Fig.4 provides an example of the final result of assignment problem, each sinusoid in one frame is assigned to one sinusoid in the other frame. Finally the minimum total distance of two frames is obtained by using Hungarian algorithm.

All frames of the signal form into a complete graph. Fig.5 shows an example of the complete graph and the MST is obtained by applying Prim's algorithm. The adjacent list of the minimum spanning tree is defined as an H-list which is both an efficient delta encoding path and the key for data security.

B. Computational Complexity

The computational time of the distance calculation between two sinusoids is *O*(*d*), where *d* is the number of parameters in one sinusoid. The computational time of Hungarian algorithm is $O(M^3 + M^2 d)$, where M is the number of sinusoids extracted from one frame. The total computational time of the proposed method is $O(N^2 log(N) M^2 (M + d))$, where N is the number of frames.

Although computational complexity of the proposed method is greater than that of the normal coding, the data size after compression is 6% smaller than normal one. In addition, we may generate the secure data that will be described below.

C. Security Consideration

Data security is ensured by transforming the data information from one form into another form by using certain algorithms depending on some keys [9]. The strength of the algorithm is often described by the length of the key, the more bits in the key, the harder it is to decrypt data simply by trying all possible key in an exhaustive search. For data that needs to remain secure for the foreseeable future, the key must be as much as 160 bit [10]. In our proposed method, the key is the H-list obtained from MST, its length depends on the number of speech frames. Suppose there are *N* frames in the input signal, and *A* denotes the number of bits that needed for encoding each frame number, the key length is A^N while the key length is *AN* bits which is long enough to be a security key.

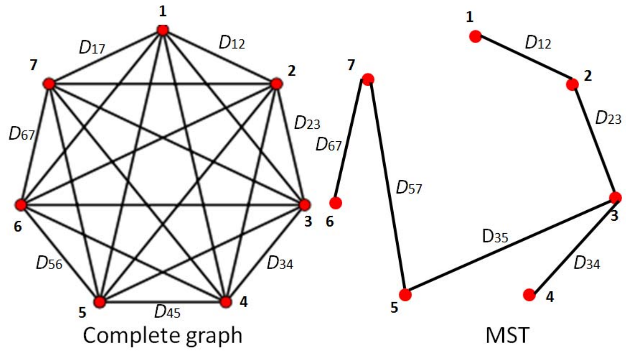

Figure 5. Complete Graph and MST, every vertex and edge represents a frame and the distance between two frames of the input signal, respectivly

TABLE I. THE MACHINE ENVIRONMENT USED IN THE EXPERIMENT

CPU	Intel Core 2 Quad Q6600
MEMORY	4.0GB
OS	Windows 7, 64-bit
IDE	Visual Studio 2010

D. Algorithm Steps

The algorithm steps can be described as follows:

1. Frame the input signal into time segments so that the characteristics can be considered stationary.
2. Extract the sinusoidal parameters, A_{ik}, w_{ik}, and θ_{ik}, from each frame, $s_i(n)$, according to Eq. (3).
3. Calculate all distance, d_{ij}, between two sinusoids for all frames according to Eq. (4).
4. Obtain the *i-j* assignment, x_{ij}, by Hungarian algorithm according to Eq. (2).
5. Calculate the total cost, D_{ij}, from x_{ij} by $$D_{ij} = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij}.$$
6. Find the MST of the complete graph with Prim's algorithm.
7. The connection of MST is stored as an H-list.

IV. SIMULATIONS

The validity and efficiency of the proposed method is evaluated by changing the parameters and the following conditions.

The conditions are normal coding, delta coding, delta coding with MST, delta coding with Hungarian algorithm and delta coding with both MST and Hungarian algorithm. The

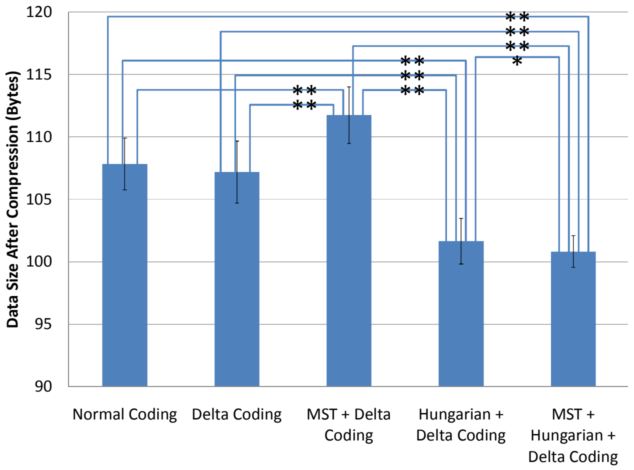

Figure 6. Average compression size of sinusoidal coding. The data size after compression by Delta Coding, MST+Delta Coding, Hungarian+Delta Coding and MST+Hungarian+Delta Coding methods are 0.6%, -3.6%, 5.7% and 6.5% smaller than Normal Coding, respectively. The marks ** and * denote significant differences with 99% and 95% confidence, respectively

machine environment used in the experiment is listed in Table I.

For evaluating the proposed method, 44.1kHz, 1-channel, 16bit, male and female speech signals were used as the input signals. The input signals were segregated into 1024-point length frames. K sinusoids were extracted from each frame. The sinusoidal parameter values were obtained from the short-time Fourier transform using a simple peak-picking algorithm.

Two simulations were done for sinusoidal coding. The first one is the comparison of the data size after compression. Five different coding methods were chosen, they are normal coding, delta coding, delta coding with MST, delta coding with Hungarian algorithm and delta coding with both MST and Hungarian algorithm. Twenty speech samples with 30-second length were chosen to evaluate the method. For each frame, 32 sinusoids were extracted. Fig.6 shows that the data size after the delta encoding with both MST and Hungarian algorithm is smaller than the other four methods. We can confirm that the proposed method is efficient for sinusoidal coding from the aspect of compression size.

The second one is the comparison of the data size after compression between delta coding without permutation, with MST and with both MST and Hungarian algorithm. In the simulation, the number of sinusoids extracted from one frame K={4,6,8,12,16,24,32,48,64} were compared. The results shown in Fig.7 illustrate that the larger the number of sinusoids the larger the data size after compression. When the number of sinusoids is less than 12, the delta coding with MST method performs better than delta coding without permutation. And delta coding with both MST and Hungarian algorithm always performs the best.

V. CONCLUSION

A modified delta encoding method using Hungarian algorithm and Minimum spanning tree has been proposed. In this method, Hungarian algorithm is applied to find the

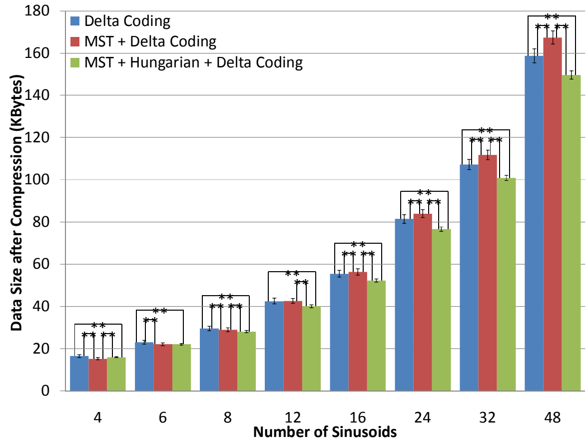

Figure 7. Comparison of the data size after compression between delta coding without permutation, with MST and with both MST and Hungarian algorithm, the marks ** and * denote significant differences with 99% and 95% confidence, respectively

minimum distance between two speech frames while MST is used to find an effective encoding path.

In simulation, the proposed method was applied to sinusoidal coding. The results show that the data size after compression of our method is 6% smaller than usual delta encoding.

The H-list in which the connection of frames is written can be used as a security key in speech or music transmitting and storing. Although it is a private-key for symmetric system, it can ensure the security of the data in some extent because of the long enough key and large enough key space obtained by the proposed method.

REFERENCES

[1] D.Shapira, "Compressed Transitive Delta Encoding," dcc, pp.203-212, 2009 Data Compression Conference, 2009.

[2] H.W.Kuhn, “The Hungarian Method for the Assignment Problem,” Naval Research Logistics Quarterly, Vol.2, pp.83-97, 1955.

[3] J.Edmonds and R.M.Karp, "Theoretical Improvements in Algorithmic Efficiency for Network Flow Problems," Journal of the ACM (JACM), vol.19, no.2, pp.248-264, 1972.

[4] R.McAulay and T.Quatieri, "Speech Analysis/Synthesis Based on a Sinusoidal Representation," IEEE Trans. ASSP, vol.34, no.4, pp.744-754, 1986.

[5] R.C.Prim, "Shortest Connection Networks and Some Generalizations," Bell Systems Technical Journal, pp.1389-1401, 1957.

[6] D.Huffman, "A Method for the Construction of Minimum Redundancy Codes," Proc. IRE, vol.40, no.9, pp.1098-1101, 1952.

[7] M.Schindler, "A Fast Renormalisation for Arithmetic Coding," Proc. Data Compression Conference, vol.0, pp.0572, 1998.

[8] D.G.Rowe, "Techniques for Harmonic Sinusoidal Coding," Ph.D. Thesis, pp.33-40, University of South Australia, 1997.

[9] A.Mousa, "Data Encryption Performance Based on Blowfish," 47th International Symposium ELMAR-2005 forcused on Multimedia Systems and Applications, pp.131-134, Zadar, Croatia, 08-10, 2005.

[10] S. William, Cryptography and Network Security: Principles and Practice, Prentice Hall, Upper Saddle River, New Jersey, 2003.