PAPER NAME

Turnitin Period Informaton Quantization.pdf

AUTHOR

florentinus Setiawan

WORD COUNT

1934 Words

CHARACTER COUNT

12612 Characters

PAGE COUNT

7 Pages

FILE SIZE

4.0MB

SUBMISSION DATE

Oct 6, 2022 3:54 PM GMT+7

REPORT DATE

Oct 6, 2022 3:54 PM GMT+7

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 14% Internet database
- 1% Publications database

Excluded from Similarity Report

- Crossref database
- Crossref Posted Content database
- Submitted Works database
- Manually excluded sources
- Manually excluded text blocks

PERIOD INFORMATION QUANTIZATION CODEBOOK OF THE SEGMENTAL SINUSOIDAL MODEL

Florentinus Budi Setiawan
Fakultas Teknologi Industri, Universitas Katolik Soegijapranata

ABSTRACT

Segmental sinusoidal model is an approximation method based on sinusoidal model for speech signal, especially for periodic part. The periodic signal can be decomposed by infinite sinusoidal signal with combination of amplitude, frequency and phase. Quantization is a method to code or compress a speech signal. The proposed quantization method in this paper is sampling signal at minimum and maximum part over one certain block. Parameters of speech signal are detected at its peaks, both positive and negative peaks. The useful parameters are peaks and period between consecutive peaks. To obtain the lower size of signal parameter, a look-up table is needed. Signal parameters are quantized into vector quantization form, so that codebook has to be generated in order to indexing of the vector quantization. In this paper, we show the experimental results, consist of codebook design and its performance based on the reconstruction of signal. Reconstruction process is realized based on periods and peaks information. The experimental results show that this is a correlation between the size of codebook and performance of reconstructed signal.

INTRODUCTION

Communication between one to the other would be performed if level of audibility is kept. Level of audibility could be maintained if level of periodicity is kept. So, the periodicity of speech signal is important for human oral communication. Periodic signal can be decomposed into sinusoidal components using Fourier series. Speech signal has the near periodic components. Human speech signal is consist of the voiced signal and unvoiced signal. The voiced signal has the certain period which mentioned as pitch. In the frequency domain, the lowest formant shows the pitch period. In the higher frequency in the speech channel, there are the second, third and fourth formant in vary frequency. In the other hand, there are some formant in the range frequency above 4 kHz. Unvoiced signal has the frequency spectra from zero until the indefinite Hertz. This signal is more difficult to analyze than the voiced signal, because of similarity characteristics with noise. Voiced signal contains redundant component. One pitch period of voiced signal has a high correlation with the one pitch period of the adjacent segment. With the high correlation, we can compress a train of signal, referred by one period of voiced signal. On the other hand, speech signal can be coded or compressed into the lower rate based on the models like linear prediction. Based on speech signal characteristics, we propose a model referred with sinusoidal form. In speech coding, there is a method referred with sinusoidal model called sinusoidal transformation coding (STC). The other method is STN (Sines+Transients+Noise). By using the sinusoidal model, compression can be realized in order to reduce the rate of speech signal data when it will be transmitted. The new method for the signal compression is the segmental sinusoidal model. By using the segmental sinusoidal model, extraction of the periods and the peaks parameter of a block of the speech signal can be implemented. In this paper, we proposed the new method to arrange a period quantization codebook of the speech signal. Encoder sends the index of this codebook to the decoder. Then on the decoder, the train of the period signal is founded based on the sent of the codebooks index.

SEGMENTAL SINUSIODAL MODEL

Sinusoidal model proposed by Almeida et.al and McAulay. Almeida's approximation is implemented by finding correlation of harmonic phase between the consecutive frame of signal. In the other hand, McAulay uses mixed-voicing methods, that phase of voiced signal is picked up from spectral envelope, under minimum phase assumption. In this condition, unvoiced phase is random. Spectral envelope is defined with linear prediction coefficients. Speech signal can be represented by the following formula

$$s(n) \approx \tilde{s}(n) = \sum_{k=1}^{K} A_k \cos(\varpi_k (n)n + \Phi_k(n)$$

$A_k(n)$ is representing amplitude, then $\varpi_k(n)$ is frequency and $\Phi_k(n)$ is representing the phase at the k-th of sinusoidal components. Signal in this model can be represent as the k sinusoidal of signal. Then the length of signal is infinite. If this signal will be quantized into sinusoidal components, k is infinite the large amount of sinusoidal components can be reduced with showing the significant components. The more components of sinusoidal signal are showed the higher quality of speech signal.

Speech signal maybe decomposed into modulated components. Speech signal is also modeled into amplitude and frequency modulation. System analysis and synthesis modeling based on overlap-add sinusoidal model combination for synthesis and speech quality enhancement is proposed by George, 1997.

Human speech signal can be approximated by segmental sinusoidal model. A segment of speech signal from a maximum peak to the minimum consecutive peak can be approximate as a cosine signal from 0 to . Then, from a minimum peak to the maximum peak can be approximate as a cosine signal from to 2. The following figure shows a segment of speech signal. The 765-th to

778-th sample is consist of 14 samples is approached by a half period of two cosine signals. The 765-th to the 774-th sample is approximated as negative cosine function, then the 774-th to 778-th sample. In this figure, the original signal is marked as (+) and the sinusoidal approach of signal is marked as (o). Part of signal from minimum to maximum is reconstructed with a half period of negative cosine function, then the part of signal from maximum to minimum is reconstructed with a half period of cosine function.

The sinusoidal signal approximation is obtained by finding the maximum and minimum peaks on the observation frame. Maximum *i*-th peak is denoted by *p(i)* and the minimum *i*-th peak is denoted by *v(i)*. The *p(i)* is the maximum peak is located before the minimum peak *v(i)*, so that the reconstructed signal from the maximum peak to the minimum peak can be formulated as [11]:

$$s_{pv}(n) = a_0 + \sum_{i=1}^{k} a_1 \cos\left(\frac{(n - n_{p(i)})\pi}{n_{v(i)} - n_{p(i)}}\right)$$

Where :

$$a_0 = \frac{p(i)+v(i)}{2} \quad \text{and} \quad a_1 = \frac{p(i)-v(i)}{2}$$

The a_0 and a_1 are the Fourier coefficients for the DC components and the first harmonic. Then the minimum peak to the maximum peak can be reconstructed by the following formula

$$s_{vp}(n) = a_0 - \sum_{i=1}^{k} a_1 \cos\left(\frac{(n - n_{v(i)})\pi}{n_{p(i+1)} - n_{v(i)}}\right)$$

Where :

$$a_0 = \frac{v(i)+p(i+1)}{2} \quad \text{and} \quad a_1 = \frac{p(i+1)-v(i)}{2} \qquad (5)$$

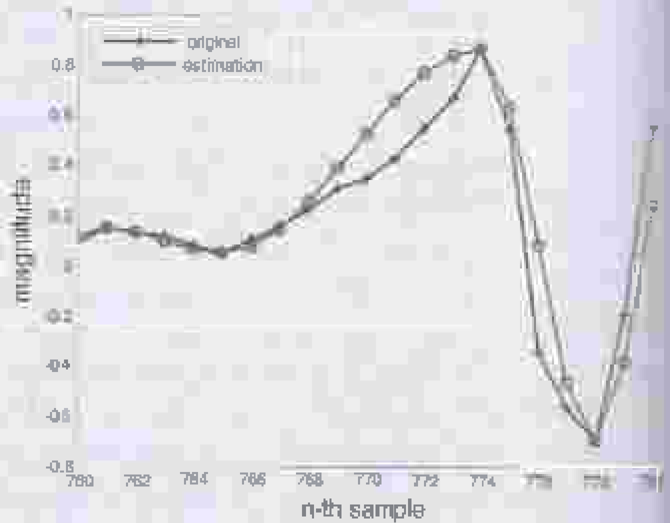

Fig 1. Segmental sinusoidal signal modeling

In a frame is consists of *2k* segments of reconstructed signal contain of *k* cosine signals and *k* negative cosine signals. The Eq.(2) and Eq(4) are the clipped Fourier series. The higher order of the Fourier series is reduced into zero in order to simplify the coding process for the lower rate. These equations means that the higher frequency is reduced, so that only the DC-offset and the first Fourier coefficient is passed into decoder.

QUANTIZATION OF THE PERIOD INFORMATION

Speech signal is quantized based on peaks value and distance between two consecutive peaks with segmental sinusoidal model. Period length quantization can be reduced in size by application of codebook or look-up table

Codebook is design with respect on the statistical characteristics of the coded signal. The method to finding the code vectors is training the large number of signal vector.

In period length quantization, codebook is design based on the quantization of period length value of the large number of speech signal, in order to obtain the accurate code vector to minimize the distortion of period length. In this paper, we explain the design of codebook model based on the quantization of period lengthSpeech signal is fetched every 30 ms (240 samples) before separated into voiced and unvoiced. This signal will be coded into segmental sinusoidal model with bit allocation is 120 bits for 30 ms in order to obtain the rate of 4000 bit per second. Speech signal can be coded into segmental sinusoidal model for every 120 samples. If the length of pitch period is more than 80 samples, it is applied the 2:1 decimation process. So the maximum number of maximum and minimum peaks is no more than 25Based on the experimental results, the period length of quantization is vary from one to 45. 2 Distribution of the quantization of the period is decreased for increasing of the length. More than 90 % the value of the quantization is less than 10.

The unvoiced signal contributes less than five of the period length quantization. The voiced signal contributes the length of period quantization more than five. The quantization of the length of period into codebook is reduced the quality of reconstructed signal.

The number of period length quantization for a block of signal with 30 ms length is varied. The coding process is more optimal if the quantization is coded into blocks with variation of the number index of codebook number of period length quantization that less than five is dominant. Variation the number of period length quantization is vary from one to 45, so that it needs 6 bit for each value.

kp0	kp1	kp2	...	kp24

0 6 7 13 14 20 21 160 161 167 bit

Fig. 2 Bit results of period length quantization

The length of period is quantized into six bit (64 probability of quantization value) based on the higher value which is 45. the value of quantization is coded into codebook with vary in length because of variation of *kp* for a block with 30 ms length.

CODEBOOK FOR PERIOD QUANTIZATION

Speech signal is fetched every 30 ms due to obtain more than probability of one pitch length of signal. Quantization of the period length for a 30 ms of signal is divided into five blocks. Each of blocks contain five value of period length quantization. For each blocks will be coded into codebooks with vary in theirs index.

The period legth quantization is coded into codebooks that vary in length due to theirs probability number of *kp* in a 20 ms of speech signal. The five blocks of signal that quantized in period length has coded into codebooks with length of 11 bits until five bits. The beginning of block is allocated the more number of bits because of its high of probability to come, especially for voiced signal. The following diagram shows the codebooks arrangement.

Table 1. Bit allocation for discrete bin

bin	block	Period codebook	bit
36-40	8	7	56
31-35	7	8	56
26-30	6	9	54
21-25	5	10	50
16-20	4	11	55
11-15	3	11	33
6-10	3	11	33
1-5	3	11	33

Some codebooks will use a large number of the memory place, in order to reduce coded data signal rate. The wider codebook has 2048 different codes, when the smallest has 32 ones. Each of codes consists of five period length quantization to include five dimension code vector. The number of memory that is allocated to the highest codebook is 5 x 2048 = 10240 places for integer. Each of the integers needs 8 bit in memory, so that memory allocation for 11 bit codebook is 8 bits x 10240 = 81920 bits, or 10240 Bytes. Total memory allocation for four codebooks with variation in size is $(2^{11}+2^{10}+2^{9}+2^{7}+2^{5})\times5\times8\,bit$ = 149760 bits = 18720 Bytes. The total bait is needed to allocate for period length quantization is (11+10+9+7+5) = 42 bits.

Then the rest 38 bit is allocated for peak quantization, formant and signal sign.

Table 2. Bit allocation for probability bin appear

bin	Block	configuration	Period CB	bit
40	8	5555 5555	7	56
39	8	5555 5554	7	56
38	8	5555 5544	7	56
37	8	5555 5444	7	56
36	8	5555 4444	7	56
35	7	5555 555	7	49
34	7	5555 554	7	49
33	7	5555 544	7	49
32	7	5555 444	7	49
31	7	5554 444	7	49
30	6	555 555	8	54
....				
....				
5	2	5	11	22
4	1	4	11	11
3	1	3	11	11
2	1	2	11	11
1	1	1	11	11

CONCLUSION

Human speech signal is consist of the voiced signal and unvoiced signal. The voiced signal has the certain period which mentioned as pitch. Unvoiced signal has the frequency spectra from zero until the indefinite Hertz. One pitch period of voiced signal has a high correlation with the one pitch period of the adjacent segment. With the high correlation, we can compress a train of signal, referred by one period of voiced signal.

The periodic signal can be decomposed by infinite sinusoidal signal with combination of amplitude, frequency and phase. Quantization is a methods to code or compress a speech signal. Parameters of speech signal are detected at its peaks, both positive and negative peaks. The useful parameters are peaks and period between consecutive peaks. In order to obtain the lower size of signal parameter, a look-up table is needed. Signal parameters are quantized into vector quantization form, so that codebook has to be generated in order to indexing of the vector quantization. Reconstruction process is realized based on periods and peaks information. The experimental results show that this is a correlation between the size of codebook and performance of reconstructed signal.

Based on previous explanation and experimental results, we conclude that can be realized a codebook for period length quantization value in order to reduce the number of data to be sent. The more codebook index number, the higher performance of the reconstructed signal.

REFERENCES

L. Rabiner, BH. Juang, *Fundamentals of Speech Recognition*, Prentice Hall international, 1993, New Jersey

S. Furui, *Digital Speech Processing, Synthesis, and Recognition*, Marcel Dekker Incorporation, 1989 New York

JR. Deller, JG. Proakis, JHL. Hansen, *Discrete-Time Processing of Speech*, Macmillan Publishing Company, 1993, New York

BS. Atal, V. Cuperman, A. Gersho, *Advances in Speech Coding*, Kluwer Academic Publishers, 1991, Massachusetts

BS. Atal, V. Cuperman, A. Gersho, *Speech and Audio Coding for Wirelles and Network Applications*, Kluwer Academic Publishers, 1993, Massachusetts

C. Rowden, *Speech Proccesing*, McGraw-Hill Book Co., 1992, London

AM. Kondoz, *Digital Speech : Coding for Low Bit Rate Communications Systems*, John Wiley & Sons Ltd., 1995,West Sussex

RJ.McAulay, TF Quatiery, Speech, *Analysis/Synthesis Based on a SinusoidalRepresentation*, IEEE TASSP, vol. ASSP-34, no. 4, 1986

TF. Quatiery, RJMcAulay, *Speech Trans formations Based on a Sinusoidal Representation*, IEEE TASSP, **vol**. ASSP-34, no. 6, 198S.

A. Rao, R. Kumaresan, *On Decomposing Speech into Modulated Components*, IEEE TSAP vol 8 hal 240-254, 2000

TF. Quatieri, TE. Hanna, GC. O'Leary, *AM-FM Separation Using Auditory-Motivated Filters*, IEEE TSAP, vol 5 hal. 465-480, 1997

EB. George, MJT. Smith, *Speech Analysis/Synthesis and Modification Using an Analysis-by-Synthesis/Overlap-Add Sinusoidal Model*, IEEE TSAP, vol 5, hal 389-406, 1997

FB. Setiawan, *Reconstruction of Quantized Speech Signal by Peak to Peak Segment with Sinusoidal Model*, International Joint Conference TSSA WSSA, Bandung, 2006

FB. Setiawan, *A Codebook for Period Information of Segmental Sinusoidal Model*, ICEEI, Bandung, 2007

14% Overall Similarity

Top sources found in the following databases:

- 14% Internet database
- 1% Publications database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	journal.ui.ac.id Internet	8%
2	icitacee.undip.ac.id Internet	5%
3	citee2015.jteti.ft.ugm.ac.id Internet	<1%
4	blog.unika.ac.id Internet	<1%
5	docplayer.hu Internet	<1%
6	ia800705.us.archive.org Internet	<1%

Excluded from Similarity Report

- Crossref database
- Submitted Works database
- Manually excluded text blocks
- Crossref Posted Content database
- Manually excluded sources

EXCLUDED SOURCES

repository.gunadarma.ac.id:8000 Internet	50%
repo.pens.ac.id Internet	42%

EXCLUDED TEXT BLOCKS

Periodic signal can be decomposed intosinusoidal components using Fourierserie...
repository.gunadarma.ac.id:8000

period which mentioned as pitch. In the
repository.gunadarma.ac.id:8000

frequency domain, the lowest formantshows the pitch period. In the higher
repository.gunadarma.ac.id:8000