Spectrum Estimation by Several Interpolation Methods

Manabu Ishihara

Oyama National College of Technology, Oyama-Shi, Tochigi , 323-0806 JAPAN

Summary
In this paper the experimental results on the spectrum estimation of the Japanese speech waveforms /a/, /i/, /u/, /e/, /o/ obtained by three different approaches of interpolation, that is Akima's approximation, the spline function approximation and Lagrange's formula, are reported. In our experiments, the speech waves are at first analyzed by FFT(Fast Fourier Transform) technique and then the FFT spectra are smoothed or fitted by varying the numbers of data points in the frequency region and applying three interpolation methods respectively. These methods are very available to extract the local peaks in the speech spectrum. The data values obtained by the investigation and comparison for the interpolated spectra indicate that Akima's approximation method is fast in calculation of smoothing process and suitable to consider the spectrum envelopes. However spline approximation is useful expect for the case of all data points being used and Lagrange's interpolation formula is not useful for the spectral estimation in our experiments.

Key words:
Spline's function, Akima's approximation, Lagrange's formula, Digital Filter

1. Introduction

A large amount of information is included human voice. As a means of considering the voice information, the voice spectra have been studied.
The original spectra are smoothed by using the 16 bit personal computer. Reviews of theoretical and experimental developments in speech spectra are given by Miura[1].
Moreover, several research papers on speech spectra have recently been reported, for example, such papers discussed on the local peaks in speech spectra[2] and the speech recognition [3],[4], under the assumption that the peaks of speech spectra are important for the perception of voice.

Spline's approximation

The method of estimation using either spline's, Akima's approximation or Lagrange's interpolation formula is basically a curve fitting technique for smoothing a large number of random data points.

In the spline approximation, a polynomial of arbitrary degree is used to process data of speech spectra . The polynomial coefficients are estimated to fit a set of data points between adjacent knots. In this paper, a cubic polynomial passing through all the given data points is utilized to obtain the optimal estimates. The cubic spline's polynomial $S_i(x)$ has to be obtained to satisfy the following relations that at all the data points xi, $S_i(x_i)$ is equal to $f(x_i)$ and the first derivative $S_i'(x_i)$ and the second one $S_i''(x_i)$ are continuous respectively.

$$S_i(x_i) = f_i,$$
$$S_i(x_{i+1}) = f_{i+1},$$
$$S_i'(x_i) = S_{i-1}''(x_i),$$
$$S_i''(x_i) = S_{i-1}''(x_i),$$

Then $S_i(x)$ is represented as
Follows:

$$S_i = \frac{1}{6h_i}[-u_i(x-x_{i+1}) + u_{i+1}(x-x_i)^3 + (6f_{i+1} - u_{i+1}h_i^2)(x-x_i) - (6f_i - u_i h_i^2)(x-x_{i+1})],$$

where

$$u_i = S_i''(x_i),$$
$$h_i = x_{i+1} - x_i,$$

Moreover, u_i can be obtained by solving the following equation

$$h_i u_{i-1} + 2(h_{i-1} + h_i)u_i + h_{i-1}u_{i-1} = 6\left[\frac{f_{i+1}}{h_i} - \left(\frac{1}{h_i} + \frac{1}{h_{i-1}}\right)f_i + \frac{f_{i-1}}{h_{i-1}}\right],$$

Manuscript received February 7, 2006.
Manuscript revised February 28 , 2006.

where

$$S_0''(x_0) = u_0,$$
$$S_{n-1}''(x_n) = u_n$$

Akima's approximation

In the Akima's approximation method, a cubic polynomial is widely utilized for a curve fitting at an interval (x_i, x_{i+1}) between adjacent two points (x_i) and (x_{i+1}). Let $(x_i, f_i), i = 0,1,2\cdots n,$ be a set of $n+1$ data points. Akima's cubic polynomial $P_i(x)$ at an interval (x_i, x_{i+1}) is described as

$$P_i(x) = a_{0i} + a_{1i}(x - x_i) + a_{2i}(x - x_i)^2 + a_{3i}(x - x_i)^3,$$
$$h_i = x_{i+1} - x_i,$$
$$m_i = \frac{(f_{i+1} - f_i)}{h_i},$$

where

$$a_{0i} = f_i,$$
$$a_{1i} = \frac{|m_{i+1} - m_i| m_{i-1} + |m_{i-1} - m_{i-2}| m_i}{|m_{i+1} - m_i| + |m_{i-1} - m_{i+2}|},$$
$$a_{2i} = \frac{(3m_i - 2a_{1i} - a_{1i+1})}{h_i},$$
$$a_{3i} = \frac{(a_{1i} + a_{1i+1} - 2m_i)}{h_i^2},$$

Akima's polynomial is very useful, as well as the spline technique, for the estimation of the optimal curve.

Lagrange's interpolation method

In the Lagrange's interpolation method, the polynomial passing through the given $n+1$ points is in general used. The Lagrange's interpolation polynomial $P(x)$ is represented as follows:

$$P(x) = \sum_{k=0}^{n} L_k(x) f_x$$

where $L_k(x), (k = 0,1,\cdots,n),$

Lagrange's interpolation coefficients, are given by

$$L_k(x) = \frac{(x-x_0)(x-x_1)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_0)(x_k-x_1)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)}$$
$$= \prod_{\substack{i=0 \\ \neq k}}^{n} \frac{(x-x_i)}{(x_k-x_i)}$$

In this paper the experimental results on the spectrum estimation of the Japanese speech waveforms /a/, /i/, /u/, /e/, /o/ obtained by three different approaches of interpolation, that is Akima's approximation, the spline function approximation and Lagrange's formula, are reported.

In our experiments[6], the speech waves are at first analyzed by FFT(Fast Fourier Transform) technique and then the FFT spectra are smoothed or fitted by varying the numbers of data points in the frequency region and applying three interpolation methods respectively. These methods are very available to extract the local peaks in the speech spectrum.

The data values obtained by the investigation and comparison for the interpolated spectra indicate that Akima's approximation method is fast in calculation of smoothing process and suitable to consider the spectrum envelopes. However spline approximation is useful expect for the case of all data points being used and Lagrange's interpolation formula is not useful for the spectral estimation in our experiments. Moreover, the typical example of the digital filter is CIC(Cascaded Integrator and Comb) filter[7]. It is using the computation of the moving average for this technique. This technique can be used for the digital filter.

2. Experimental Results

The legacy computer's mother-board (CPU:80286 (10MHz), memory 640kByte) is utilized for speech signal processing, and the system program is written in both BASIC and machine languages. But the program for calculating the polynomial coefficients is written in BASIC language. The 12 bit A/D converter is used to digitize the analog speech signals, at the sampling frequency of 10 kHz[5].

2.1 Spline function approximation

Cubic spline polynomial has been used to interpolate the data points of voice spectra of Japanese five vowels, obtained by FFT technique. Figure 1 shows a spectral curve obtained by taking data points every other spectral point. It is seen in Figure 1 that almost the same voice spectra as those of the original signals have been obtained.

Time required for calculating this curve fitting has been about 11 sec.

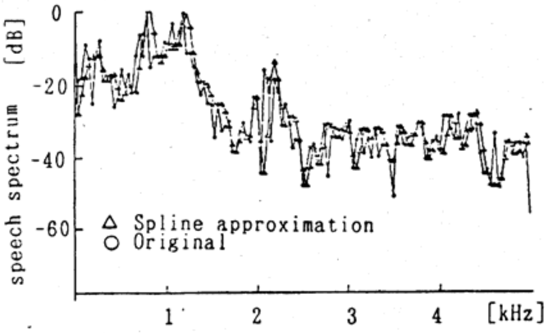

Fig.1 Relation between speech spectrum of vowel of /a/ and spline function. (In case of every other)

Figure 2 shows a spectral curve obtained by taking data points every three spectral points. As is seen in Figure 2 the fine structure of the spectra has been averaged and the smoothed envelope has been obtained.
The calculating time in this case has been about 5 sec.

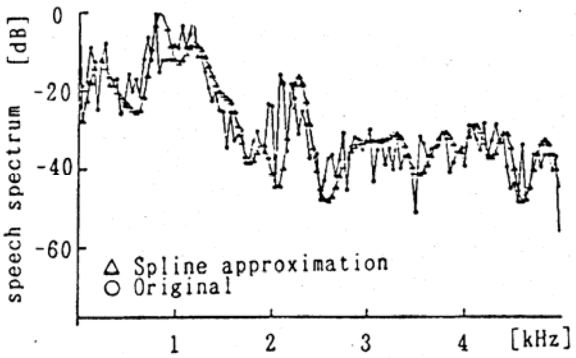

Fig.2 Relation between speech spectrum of vowel of /a/ and spline function. (In case of every three)

In Figure 3 the spectral curve obtained by taking data points every seven spectral points is shown. Fine structure of the original spectra is averaged out, but the locations for the maximal and minimal values are sufficiently distinguished. The calculating time in this case is about 3 sec.

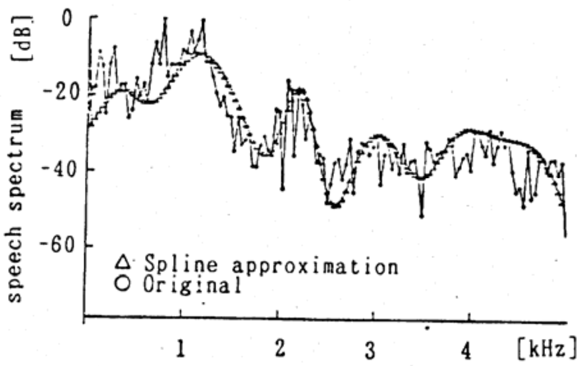

Fig.3 Relation between speech spectrum of vowel of /a/ and spline function. (In case of every seven)

2.2 Akima's approximation

Akima's cubic polynomial has been used to interpolate the data points of voice spectra of Japanese five vowels, obtained by FFT technique. Figure 4 shows a spectral curve obtained by data points. The voice spectra in Figure 4 are the same as those of the original signals. Time required for calculating this curve fitting has been about 11 sec.

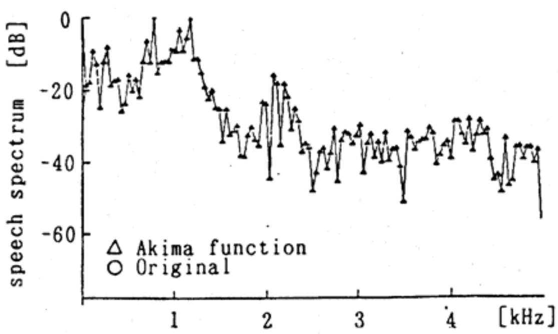

Fig.4 Relation between speech spectrum of vowel of /a/ and Akima function. (In case of same date point)

Figure 5 shows a spectral curve obtained by taking data points every other spectral points. It is seen in Figure 5 that almost the same voice spectra as those of the original signals has been obtained. Time required for calculating this curve fitting has been about 6 sec.
Figure 6 shows a spectral curve obtained by taking data points every three spectral points. As is seen in Figure 6, the fine structure of the spectra has been averaged, and the smoothed envelope has been obtained. The calculating time in this case is about 3 sec.

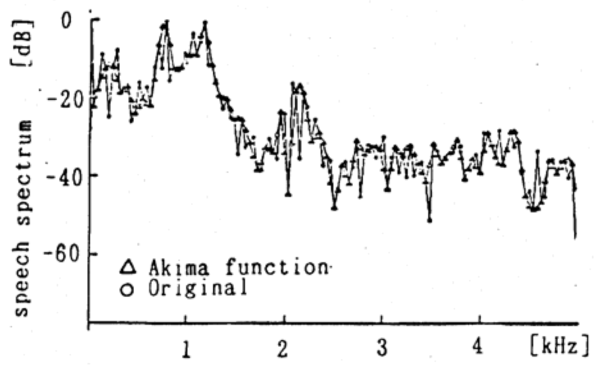

Fig.5 Relation between speech spectrum of vowel of /a/ and Akima function. (In case of every other)

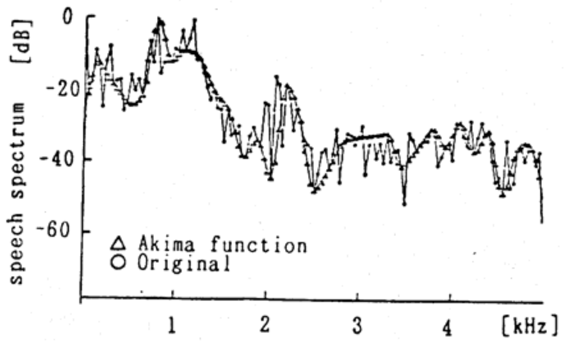

Fig.6 Relation between speech spectrum of vowel of /a/ and Akima function. (In case of every three)

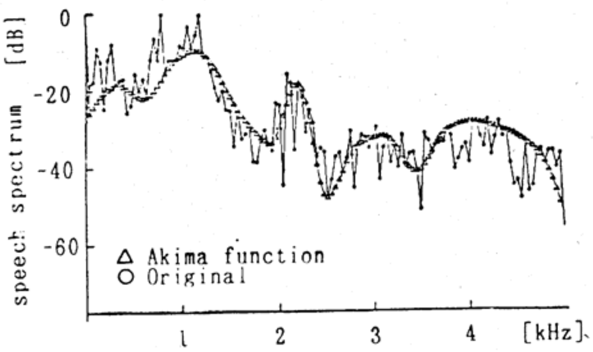

Fig.7 Relation between speech spectrum of vowel of /a/ and Akima function. (In case of every seven)

In Figure 7 the spectral curve obtained by making use of data points every seven spectral ones is shown. The fine structure of the original spectra is averaged out, but the locations for the maximal and minimal values of spectra are sufficiently recognized. The calculating time in this case is about 2 sec.

2.3 Lagrange's formula

Three kinds of interpolation method have been applied to the voice spectra obtained by FFT. The features and temporal variation of the voice spectra have been investigated. As the results, it is shown that the cubic spline polynomial and the Akima's function are the optimal techniques for curve fitting. It is revealed by the simulation that the polynomial of Akima is superior to others in the calculation time. Many attempts for speech recognition using the local peaks of a voice have recently been reported. It is considered that the peaks of the voice spectra could probably be recognized even if the numbers of data points are few.

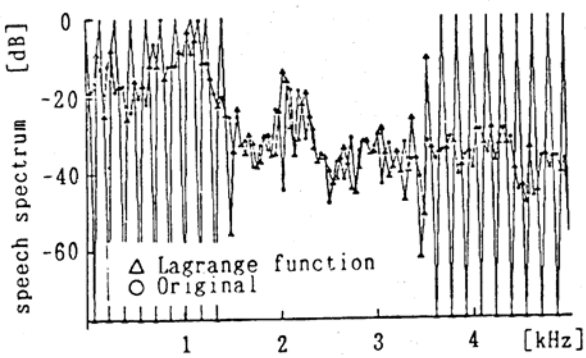

Fig.8 Relation between speech spectrum of vowel of /a/ and Lagrange function. (In case of every other)

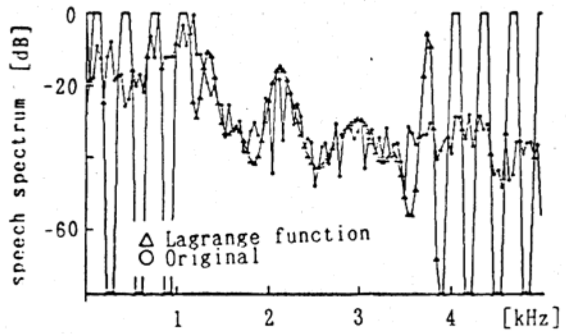

Fig.9 Relation between speech spectrum of vowel of /a/ and Lagrange function. (In case of every three)

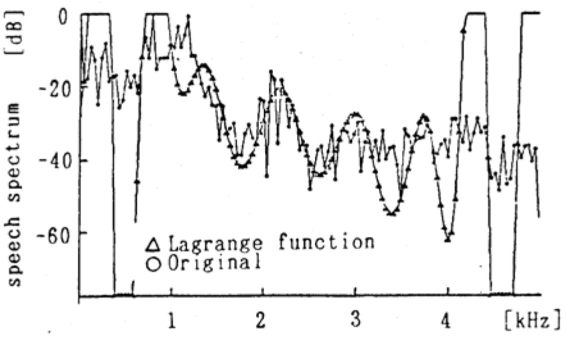

Fig.10 Relation between speech spectrum of vowel of /a/ and Lagrange function. (In case of every seven)

We calculate at high speed if we use the latest computer. We intend hereafter to investigate in detail the relation between the numbers of data points and the envelope of the voice spectra.

2.4 Calculating time

The data values obtained by the investigation and comparison for the interpolated spectra indicate that Akima's approximation method is fast in calculation of smoothing process and suitable to consider the spectrum envelopes. We show the comparison of the computation speed in the figure 11. The experiment on us can be used as the digital filter.

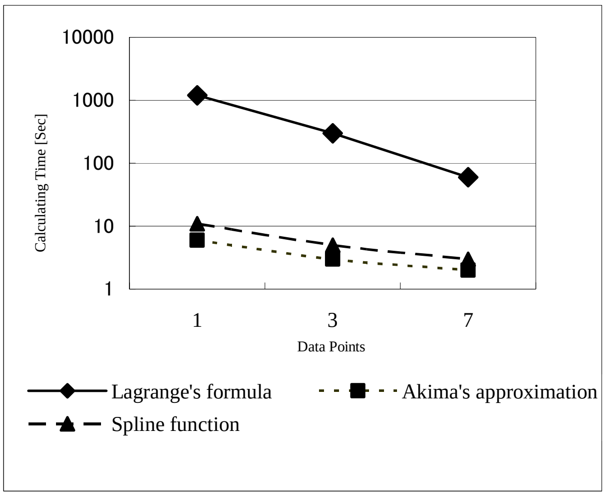

Fig.11 Relation between calculating time and data points

3. Concluding Remarks

Three kinds of interpolation method have been applied to the voice spectra obtained by FFT. The features and temporal variation of the voice spectra have been investigated. As the results, it is shown that the cubic spline polynomial and the Akima's function are the optimal techniques for curve fitting. It is revealed by the simulation that the polynomial of Akima is superior to others in the calculation time. Many attempts for speech recognition using the local peaks of a voice have recently been reported. It is considered that the peaks of the voice spectra could probably be recognized even if the numbers of data points are few.

We calculate at high speed if we use the latest computer. The experiment on us can be used as the digital filter. We intend hereafter to investigate in detail the relation between the numbers of data points and the envelope of the voice spectra.

References

[1] T.Miura etc :"Auditory and Speech," IECE(1983) (in Japanese).
[2] T.Matsuoka and K.Kido : "Investigation on Phonemic Information of Static Properties of Local Peaks in the Speech Spectra," J.Accost. So. Vol.32, No.1, pp.2-23 (1976) (in Japanese).
[3] J.Miwa, Y.Niitsu, S.makino, and K.Kido : "Spoken Word Recognition System Using Gross Features of Speech Spectrum and These Dynamic Properties," J.Accost. Soc. Vol.34, No.3(1978) (in Japanese).
[4] N.Ikeda, T.Aoki, H.Koga, and K.kido : "Recognition of Vowels Using the Local Peaks in FFT Spectrum," J.Accost. Soc. Jap. Vol.41, No.12, pp.886-890 (1985) (in Japanese).
[5] J.Shirataki and M.Ishihara : "Speech Analysis System by Use of Personal Computer," Research Reports of Ikutoku Technical University, part B, No.11, pp.127-131(1987) (in Japanese).
[6] M.Ishihara, J.Shirataki, S.Ieiri and Y.Shikata :" Spectrum Estimation of Speech by Several Interpolation Methods," Proceedings of The First China-Japan International Symposium on Instrumentation, Measurement and Automatic Control, pp.442-448(1989).
[7] Shogo Nakamura :" Digital Signal Processing," Tokyo Denki Pub.,pp.140-154(2004) (in Japanese).

Manabu Ishihara received the B.E. degrees in Electrical Engineering from Ikutoku Technical University, Japan in 1981, respectively. He received the M.E. and Ph.D. degrees in Electrical Engineering from Meisei University, Japan in 1984 and 1987, respectively. Now, He become a Associate Professor in the Department of Electrical and Computer Engineering, Oyama National College of Technology. His research multimedia, computer network, human engineering. He is member of the IEEE, the IEICE, the IEE of Jpn. and ASJ.