EFFICIENT CODING AND RESONANCE SPIKE IDENTIFICATION FOR TOPSIDE IONOGRAM PROCESSING

Seiji IGI, Kazuhiro AIKYO, Ryo NISHIZAKI, Takanobu OGATA, Takashi MARUYAMA

Radio Research Laboratories, 2–1, Nukui-Kitamachi 4-chome, Koganei-shi, Tokyo 184

and

Takeo HIRASAWA

National Institute of Polar Research, 9–10, Kaga 1-chome, Itabashi-ku, Tokyo 173

Abstract: This paper describes an effective coding method to eliminate the redundancy contained in digital ionograms and an algorithm to identify the resonance spikes appearing on topside ionograms. This work is a first step toward automatic profile reduction to obtain ionospheric electron density profiles. Topside sounder data recorded by ISIS-2 satellite are digitized and converted to digital ionograms. A quantitative comparison of data compression techniques, based on the run-length and predictive coding method, is made and leads the conclusion that the modified run-length coding method is most effective and useful from the practical view point. This simulation experiment results in self-consistent determination of characteristic frequencies with good accuracy except for the ionograms with obscure resonance spikes.

1. Introduction

With the advent of the topside sounder satellites, a vast amount of information about the topside of the ionosphere has been collected and stored in the form of ionograms. Although the reduction of topside ionograms to electron density profiles makes it easy to study detailed morphological features of the topside of the ionosphere, only a limited amount of ionograms have yet been analyzed, due to the complexity of N(h) analysis.

An on-line ionogram processor (ONLIP) system developed by the Radio Research Laboratories (RRL) in 1974 is a computer-aided N(h) reduction system with the ability to scale echo traces and, if necessary, to correct them using a graphic input unit (MATUURA *et al.*, 1978). Although some ionogram reduction work using the ONLIP system has been performed and published, the further development of the data processing techniques permits us to improve this system. This paper describes a first step toward the automatic reduction of N(h) profiles from topside ionograms without the need of manual scaling and echo pattern recognition by an operator. The following five problems relating to computer algorithms of automatic reduction are left unsolved: (1) to compress data contained in digital ionograms, (2) to identify resonance spikes, (3) to search for critical frequencies, (4) to scale echo traces, and (5)

to evaluate N(h) profiles, with the aid of a computer. This paper is mainly concerned with the problems (1) and (2), that is, the effective coding method to eliminate the redundancy contained in digital ionograms, and the algorithm to scale the characteristic frequencies through the resonance spike identification.

In the following, Section 2 describes the outline of automatic reduction system. A procedure for conversion to digital ionograms is given in Section 3. Section 4 deals with the redundancy reduction of digital ionograms. The resonance spike identification method is described in Section 5.

2. Automatic Reduction of N(h) Profiles

A flow-chart describing the process for automatic reduction of a topside ionogram

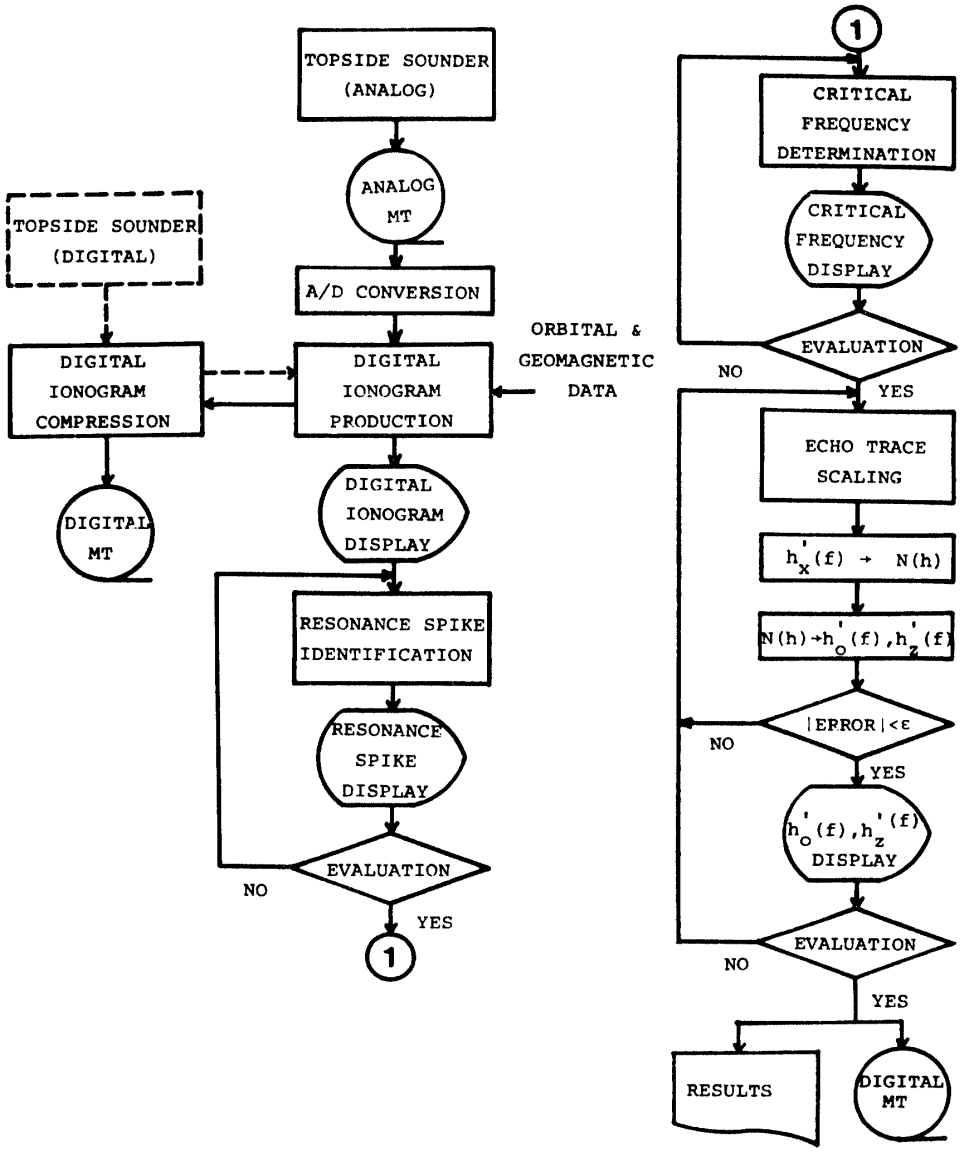


Fig. 1. Flow-chart describing the process for automatic reduction of topside ionograms to electron density profiles.

to an electron density profile is shown in Fig. 1. First, digital ionograms must be reproduced for computer analysis from the telemetered analog video signal. The ionogram observed by the topside sounder onboard the ISIS spacecraft consists of several hundreds of individual analog A-scan lines. To deal with the analog ionograms using a computer, the video signals have to be digitized and sounding frequencies determined for each line.

The second problem is the compression of the data contained in digital ionograms. In recent years the data obtained by scientific satellites are transmitted in machine readable form, that is PCM form, because of high noise immunity of PCM. Thus the scientific satellite data contain a large amount of binary sequences, which usually exceeds the channel capacity provided by the satellite communication link. For example, each ionogram shown in Fig. 6 has about four megabits of binary digits, so that a high level of ground storage capacity is needed to store such an amount of data. To solve the problem, an effective method to eliminate redundancy contained in digital ionograms has to be developed.

The final goal is to derive the N(h) function from a knowledge of the apparent range $h'(f)$ as a function of frequency f. A lamination method for derivation of N(h) function was developed by HOJO (1969) and HOJO and NISHIZAKI (1971) in RRL. Our present objective is the replacement of the processes depending on the abilities of an operator by pattern recognition and decision with the aid of a computer. The first step of the N(h) analysis is to determine the plasma frequency f_N at the satellite, which is directly related to the local electron density. This obliges us to establish the algorithm to identify the characteristic frequencies listed below in a self-consistent manner.

f_{NS} : plasma resonance frequency (same as the cutoff frequency of the O-mode echo) at the satellite
f_{XS} : cutoff frequency of the X-mode echo at the satellite
f_{ZS} : cutoff frequency of the Z-mode echo at the satellite
f_{HS} : electron gyrofrequency at the satellite
f_{TS} : upper hybrid frequency at the satellite

For the next step automatic tracking of echo traces becomes the problem of utmost importance, since this procedure determines the $h'(f)$ function and critical frequencies in the F-layer directly related to the maximum electron density.

Conventionally most of the topside N(h) profiles have been derived from analysis of the X-mode echoes, however, a calculation method using all of O-, X-, and Z-mode traces should be developed ultimately.

3. Production of Digital Ionograms

First of all for computer analysis, digital ionograms have to be made from the analog video signals telemetered from ISIS satellites.

Figure 2 illustrates the schematic diagram for the production of digital ionograms. A normal ISIS-2 video waveform including echoes is exemplified on the left hand side of Fig. 2. The apparent range from the satellite is sampled at 5 km steps after detecting the line synchronization pulse. Amplitude is quantized (8 bits; 256 levels

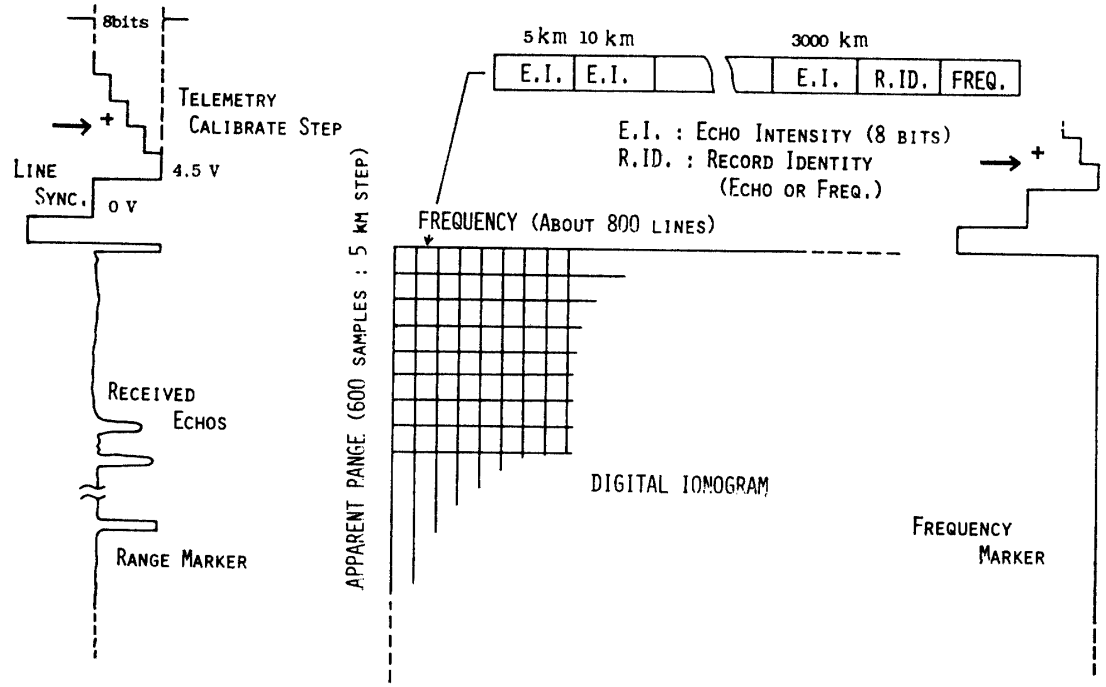


Fig. 2. Schematic illustration for the production of digital ionograms. A normal ISIS-2 video waveform including echoes is exemplified on the left, and video waveform of the frequency marker is shown on the right of the figure. The sample data of one frequency line is shown at the top of this figure.

of gray scale) between 0 V and 4.5 V in telemetry calibration steps. On the top of the figure is shown the sampled data of one frequency line, which are stored in the magnetic disc memory sequentially. The video waveform of the frequency marker is shown on the right hand side of Fig. 2. The frequency of the normal video signal is interpolated from the frequency markers. The telemetry calibration steps, the line synchronization pulses, and the frequency markers are automatically detected with the aid of the computer. The 31 and 40 frames of digital ionograms, which are recorded at Syowa (Antarctica) and at Kashima Stations (Japan), respectively, are used as typical examples for the simulation experiment.

4. Redundancy Reduction of Digital Ionograms

The method of elimination of the redundancy contained in the digital ionograms can be regarded as one particular branch of picture coding methods. Many picture coding methods have been suggested (for review, see NETRAVALI and LIMB, 1980), and applied to satellite data (TUNER, 1973; BENELLI *et al.*, 1980). Various bilevel coding methods were reviewed by HUANG (1977), including the run-length and the predictive coding methods, and YOSHIDA (private communication, 1981) applied the run-length method to coding of bottomside ionograms. In this paper we use run-length methods RL-1, RL-2 and RL-3 and the predictive coding method for the data compression of the topside ionograms. Figure 3 gives brief descriptions of these four methods.

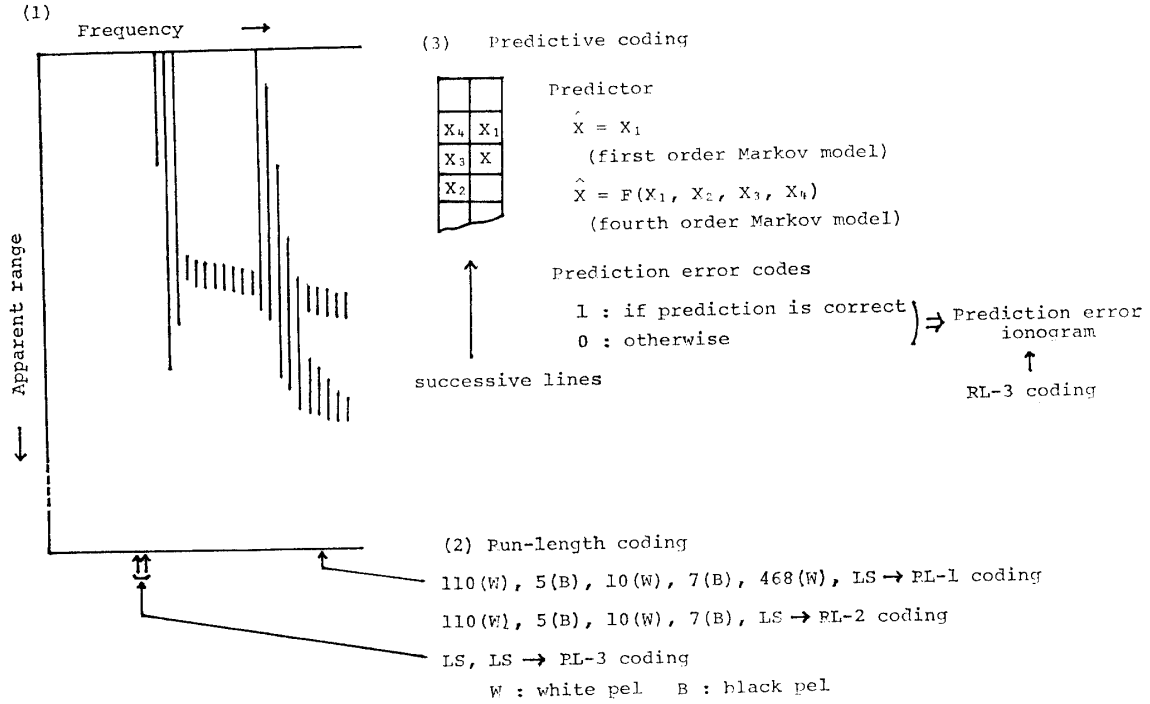


Fig. 3. Brief descriptions of the run-length and predictive coding methods.
(1) Rough diagram of a digital ionogram.
(2) Illustration of the three run-length (RL) coding methods. The figures and notation give constituents of the digital A-scan.
(3) Illustration of the predictive coding method. The figure shows an example of pels in the successive two scan lines.

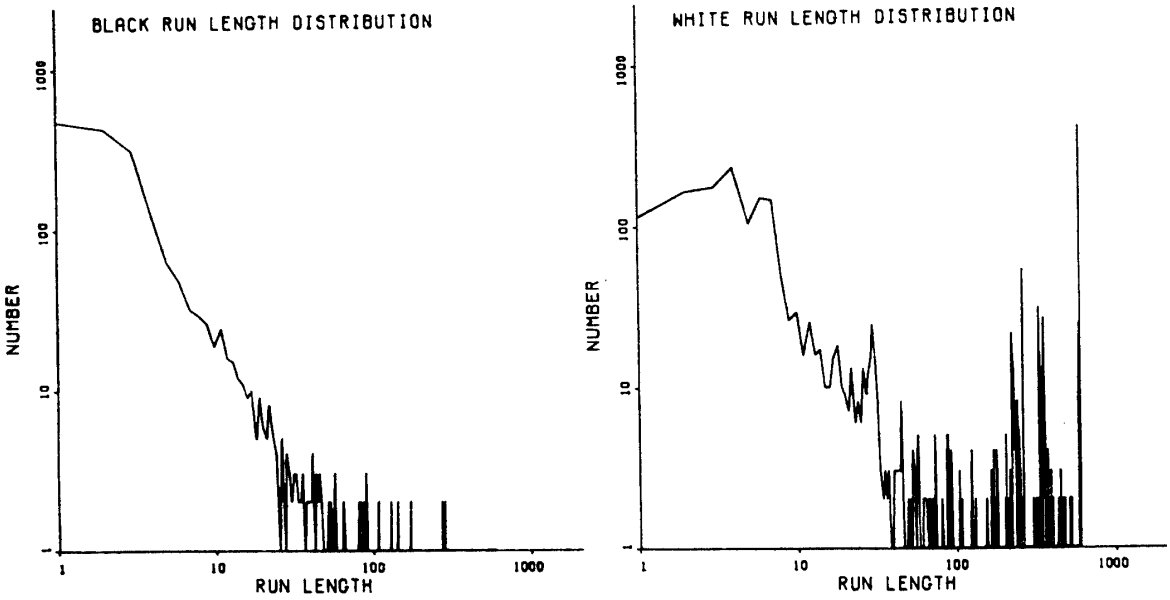


Fig. 4. Black and white run-length distribution of sample 1 in Fig. 6.

a. Run-length coding method 1 (RL-1)

The schematic illustration of run-length coding methods RL-1, RL-2, and RL-3 is shown in the bottom part of Fig. 3. Assume that an A-scan line constitutes 110 white pels (picture elements) followed by 5 black pels, 10 white pels, 7 black pels, and then 468 white pels. Then in the run-length coding method the messages we transmit are the "run-lengths" 110, 5, 10, 7, 468 and the line synchronization (LS) word. Each run-length tells us the relative location of the next boundary points along the A-scan line. Since white and black runs always alternate, the color of each run needs not be transmitted if the initial color is known. The problem of assigning code words to each run in an optimal way can be solved by using the Huffman code, which minimizes the average code word length for a given probability distribution of the run-length. Since the probability distribution is different for white and black runs (for an example see Fig. 4), individual Huffman code word sets must be provided for coding the white and black runs respectively. However, since constructing the Huffman code is a complex task, each run is encoded by the modified Huffman code (MUSMANN and PREUSS, 1977) shown in Table 1.

Table 1. Code word assignment for the modified Huffman code.

Run length	Code words for white runs	Code words for black runs
0	00110101	0000110111
1	000111	010
2	0111	11
3	1000	00
4	1011	011
.	.	.
.	.	.
.	.	.
63	00110100	000001100111
64	11011	0000001111
128	10010	000011001000
192	010111	000011001001
.	.	.
.	.	.
.	.	.
1728	010011011	0000001100101
Line synchronizing word	00000000001	

b. Run-length coding method 2 (RL-2)

As the total number of samples in one scan line is known beforehand, we need not transmit a run to predetermine it. The last runs are not transmitted in this simulation, because they will probably be very long runs.

c. Run-length coding method 3 (RL-3)

No white line needs to be transmitted, because all white lines are recognizable by the subsequence of line synchronization (LS) words.

d. Predictive coding method

The effective elimination of redundancy can be expected due to the close correlation of successive scan lines. The most popular method used is the predictive coding method illustrated in the top part of Fig. 3. The first step is to make a prediction of the present picture element X from the previously transmitted samples $X_1, X_2, \cdots, X_n$. Although two types of predictors; the first order Markov model, and the fourth order Markov model, are given in Fig. 3, the latter case will be explained in detail.

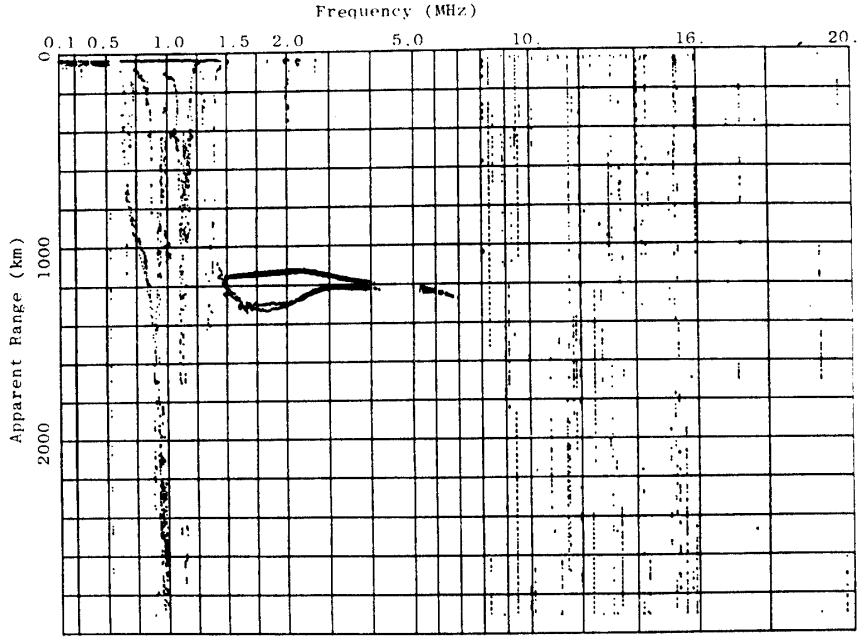
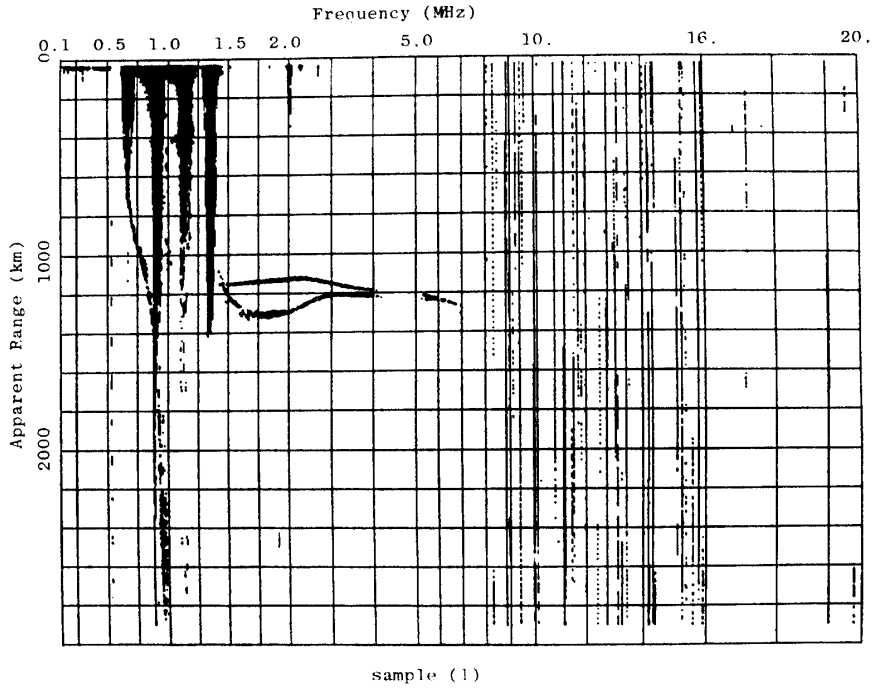


Fig. 5. Ionogram consisting of the prediction error codes for sample 1 in Fig. 6.

Fig. 6a.



Conditional probabilities $P(X|X_1, X_2, X_3, X_4)$ are calculated from the representative ionogram, and the predictor for X, $\hat{X}=F(X_1, X_2, X_3, X_4)$ is determined from the table of conditional probabilities P (for further detail see PREUSS (1975)). The predictive error coded ionograms (as shown in Fig. 5) are made by comparing the predicted values $\hat{X}$ with the real values X and encoded by run-length coding method 3. In this experiment first order Markov model, that is, $\hat{X}=X_1$, is used as the predictor.

The following four samples of ionograms shown in Fig. 6 are used to decide which method is the most effective.

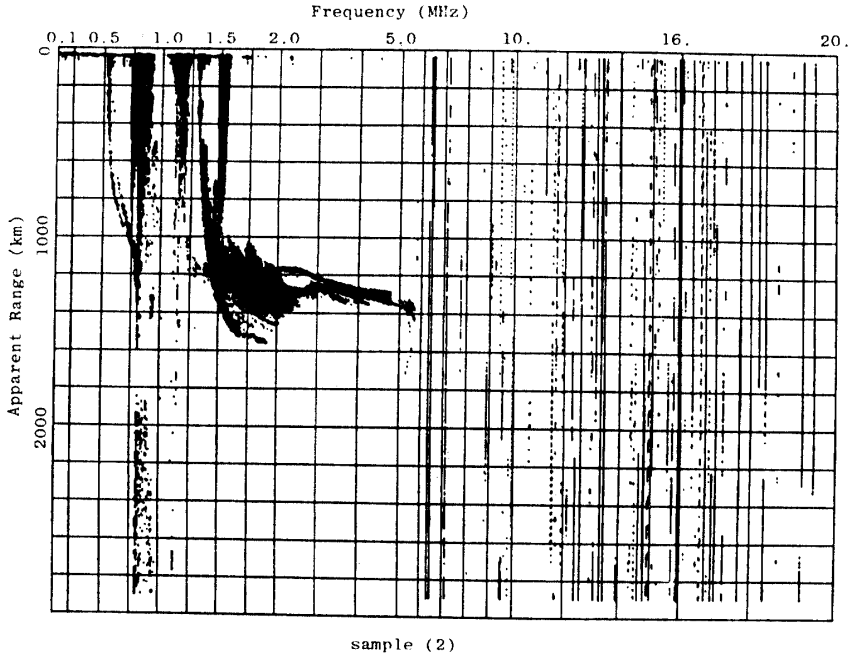


Fig. 6b.

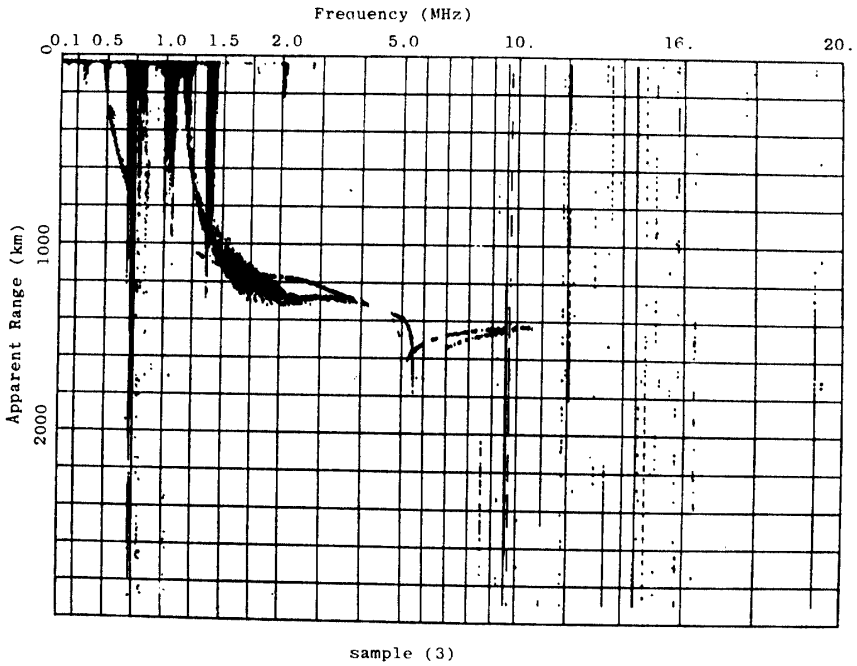


Fig. 6c.

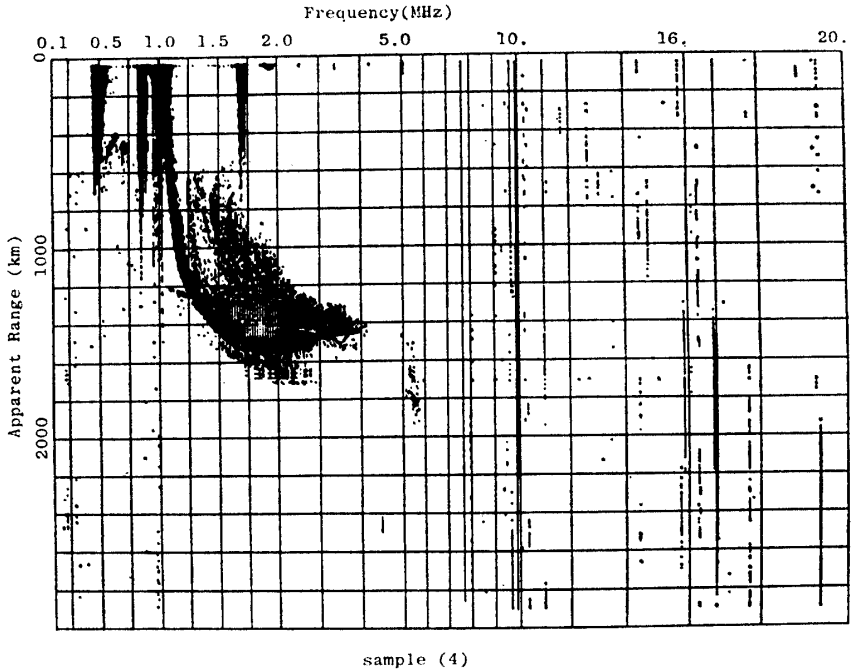


Fig. 6d.

Fig. 6. Sample ionograms used in the simulation of the data compression.

Sample 1: normal traces recorded at Kashima Station
Sample 2: complicate traces recorded at Kashima Station
Sample 3: normal traces recorded at Syowa Station
Sample 4: complicate traces recorded at Syowa Station

Simulation results for the four samples are given in Table 2. Generally the message set contains up to run-length N, where N is the number of pels in a scan line. If the probabilities of run-lengths, 1, 2, $\cdots$, N, are $p_1, p_2, p_3, \cdots, p_N$, the average length in bits per run, L, satisfies the following inequality according to Shannon's first theorem on the noiseless coding of information,

$$H \leq L \,, \tag{1}$$

where the run-length entropy, H, is defined as,

$$H = -\sum_{i=1}^{N} p_i \log_2 p_i \,. \tag{2}$$

The following condition can be met for the average length, l, in bits per pel,

$$\frac{H}{V} \leq l \,, \tag{3}$$

where V is the average run-length in pels as given by,

$$V = \sum_{i=1}^{N} i p_i \,. \tag{4}$$

In other words, the entropy per pel, $h = H/V$, gives a lower bound for the average length in bits per pel. The entropy (black and white) per pel also gives a lower bound under the condition that white and black runs are coded independently by different

Table 2. Simulation results of data compression for four samples

Method	Sample			
	(1)	(2)	(3)	(4)
Entropy	0.049	0.060	0.057	0.076
Entropy (black and white)	0.046	0.057	0.054	0.073
Run-length coding 1	0.083	0.093	0.089	0.109
Run-length coding 2	0.071	0.082	0.079	0.098
Run-length coding 3	0.059	0.070	0.065	0.086
Predictive coding	0.076	0.090	0.086	0.115

(unit; bits/pel)

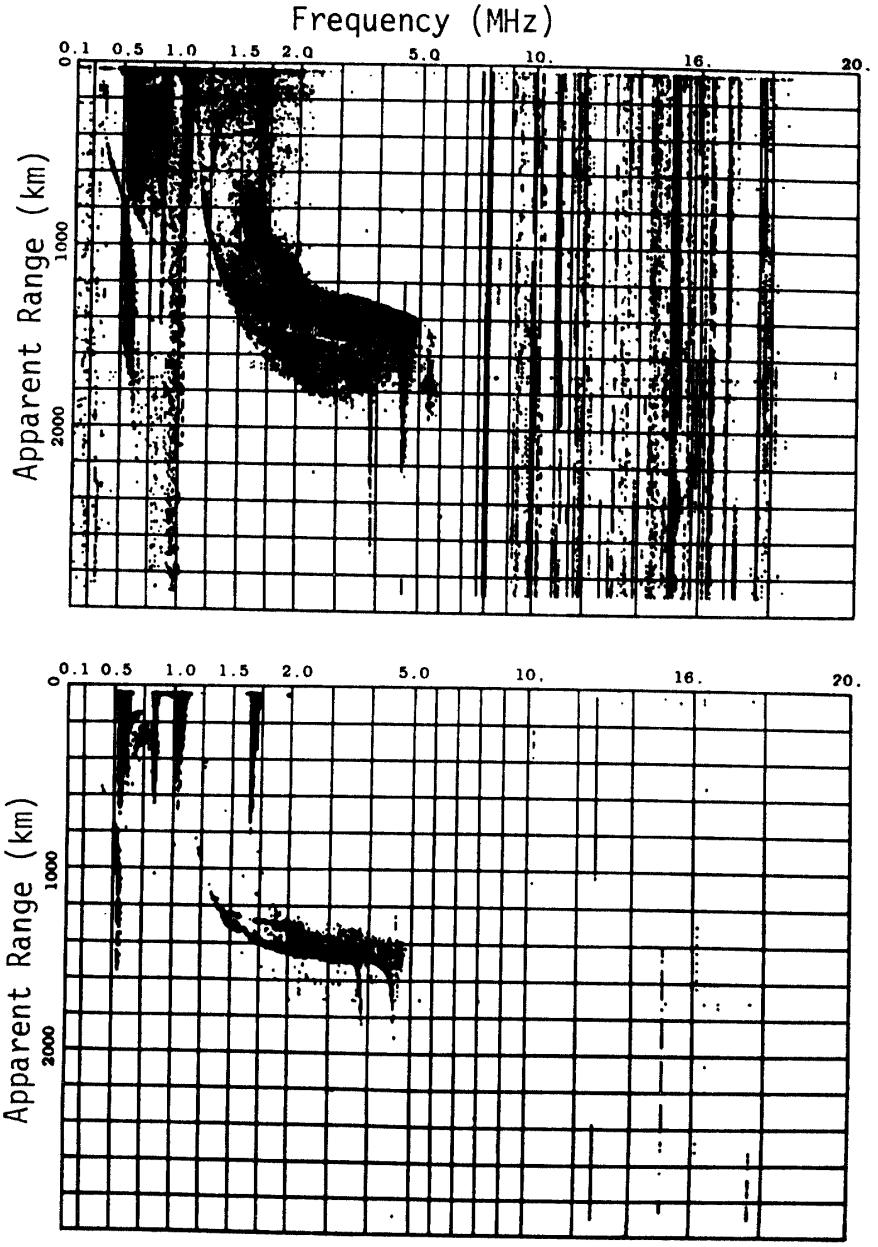


Fig. 7. Two examples of ionograms with different threshold selections from the same video signal.

codes. Therefore, a most effective data compression is achieved by coding with an average length as near as possible to the entropy per pel.

Simulation results for the four samples are given in Table 2. It can be concluded that the run-length coding 3 method is most effective as far as proximity to the entropy is concerned. In this paper we deal only with the data compression problem, assuming noiseless channel. In general, however, the compressed data are more sensitive to channel noise. For example, in the run-length coding 1 method we can easily detect the occurrence of a single bit error, whereas in the run-length coding 2 method the occurrence cannot be detected. Therefore some compensation measures, channel coding for example, must be taken in the transformation of compressed data on a noisy channel.

Figure 7 shows examples of two ionograms with different threshold levels from the same video signal. These figures indicate that the optimal threshold level to detect the main echoes is different for each sounding frequency. This fact suggests that multilevel coding methods for digital ionograms will have to be developed in future.

5. Identification of Resonance Spikes

The first step toward the automatic N(h) analysis needs a resonance spike identification algorithm, as stated in Section 2. The flow chart of this algorithm is illustrated in Fig. 8.

This algorithm involves the following steps:

Step 1. The multilevel ionogram is converted to a bilevel one.

Step 2. Bilevel echoes are integrated along the apparent ranges.

Step 3. The echo histogram is smoothed by a running average method with three terms.

Steps 4–5. A differential histogram derived from smoothed histogram enables detection of peak and valley frequencies. Afterward these frequencies are tabulated, to make it convenient to proceed to subsequent processing steps.

The illustration associated with the first four steps is given in Fig. 9. The next steps are as follows.

Step 6. To calculate the electron gyrofrequency f_{HC} from the IGRF model.

Step 7. To look for f_H nearest to f_{HC} in the table of peak frequencies.

To look for $2f_H$ for middle and high latitude ionograms, and f_H for low latitude ionograms, because these resonance spikes tend to appear far from other resonance spikes.

The procedures from step 8 to step 17 constitute the loop for identification of the other resonance spikes.

Step 8. To assume a plasma frequency by considering following properties.

(1) The value of f_N is assumed to be within a frequency range of two megahertz centered on the plasma frequency already determined in preceding ionogram, on the premise that an abrupt change in the electron density does not occur between two successive ionograms.

(2) In each loop, the assumption of f_N is made in the order of the peak heights within the frequency range mentioned above.

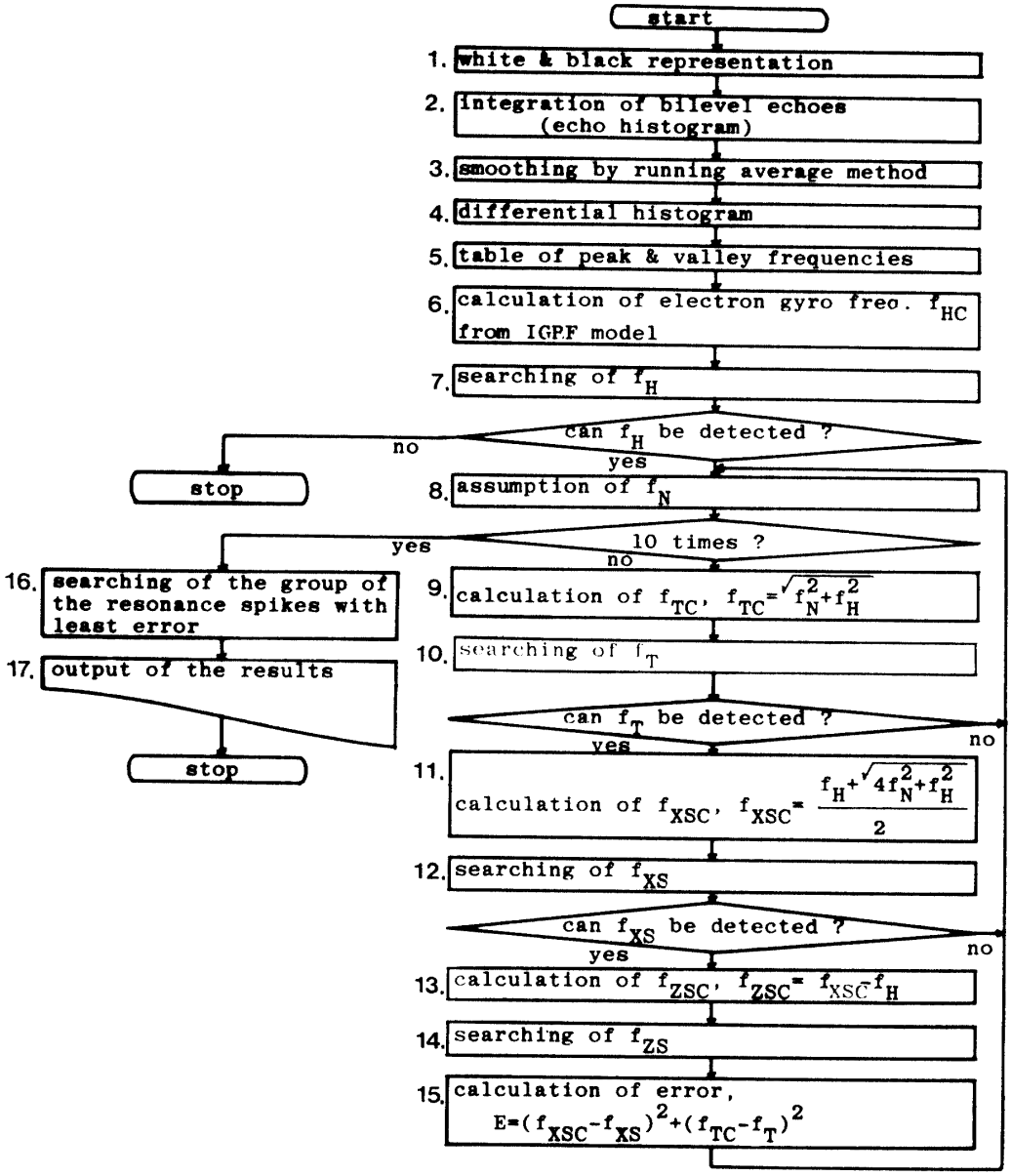


Fig. 8. Flow-chart of the resonance spike identification algorithm.

(3) The degree of the sharpness of the many peaks appearing in the smoothed histogram is taken into consideration in the assumption of the plasma frequency.

Steps 9–15. The frequencies f_{TC}, f_{XSC}, and f_{ZSC} are calculated from the assumed plasma and gyrofrequency by following relations.

$$f_{TC}=\sqrt{f_N^2+f_H^2}\,, \tag{5}$$

$$f_{XSC}=\frac{f_H+\sqrt{4f_N^2+f_H^2}}{2}\,, \tag{6}$$

$$f_{ZSC}=f_{XSC}-f_H\,. \tag{7}$$

Then, peaks nearest to f_{TC}, f_{XSC}, and f_{ZSC} in the table are looked for next. If the

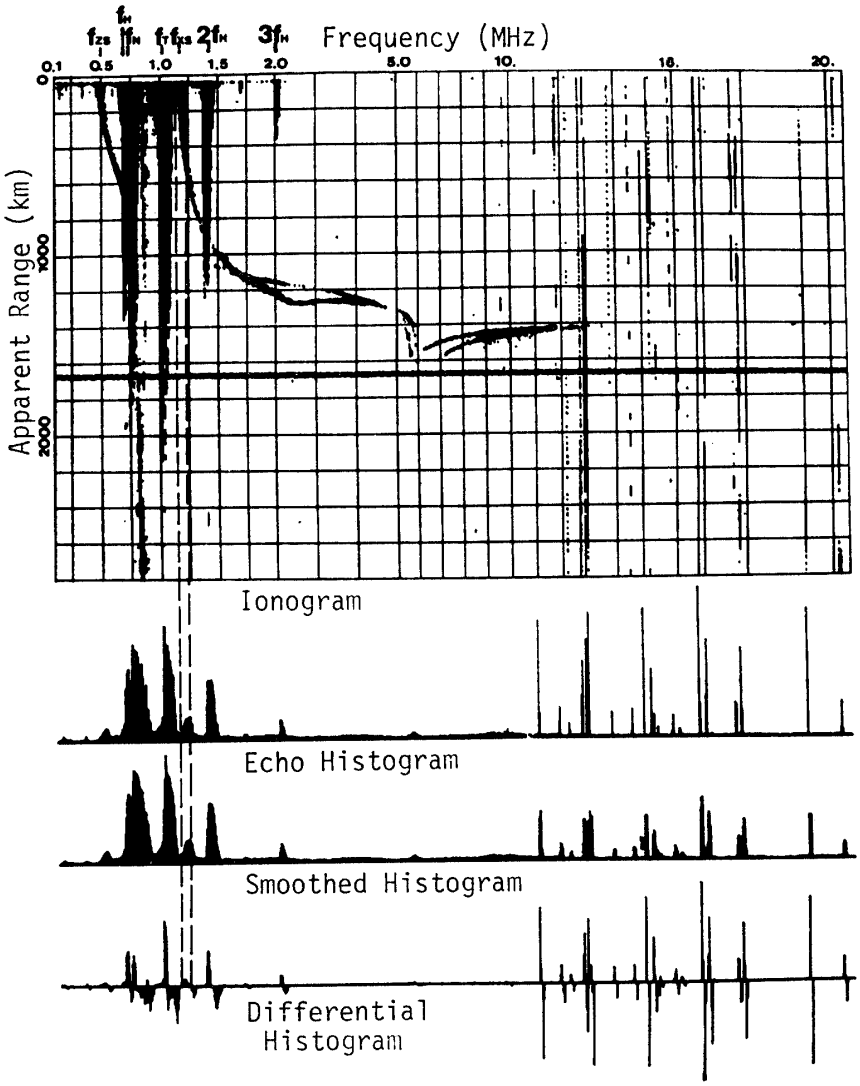


Fig. 9. Illustration of the resonance spike identification algorithm. The top panel shows a typical ionogram recorded at Syowa Station, Antarctica. The second, third, and last panels show the integrated histogram of bilevel echoes along the apparent range, the smoothed histogram by running average method with three terms, and the differential histogram for the detection of peak and valley frequencies, respectively.

frequency f_{XS} has no spiky structure in the smoothed histogram, it can be scaled at the mid point between the peak and the valley (see Fig. 9).

In case of no detection of f_T or f_{XS} the procedure is started again at step 8. The square error, E, is calculated as follows, where the cutoff frequency of the Z-mode, f_{ZS}, is not taken into account because of frequent absence of the Z-mode echo.

$$E=(f_{XSC}-f_{XS})^2+(f_{TS}-f_T)^2 . \tag{8}$$

In case of $f_N>f_H$, the procedures from step 8 to 15 are iterated ten times, and five times if $f_N<f_H$, because the resonance duration of f_N is known to be long for $f_N>f_H$ compared with the opposite case (MULDREW, 1972).

Steps 16–17. Resultant output.

A group of resonance spikes with least square error given by eq. (7), E, is finally selected as the best result.

Simulation of the data recorded at Syowa and at Kashima Stations are shown in Fig. 10 and Fig. 11, respectively. The algorithm was not able to identify groups of resonance frequencies on 3 out of 31 frames of ionograms acquired at Syowa Station and 1 out of 40 frames of ionograms at Kashima Station. This lack of identification can be considered to be the result either of masking of the plasma resonance spike by Z-mode spread echoes or of the short duration of the plasma resonance spike.

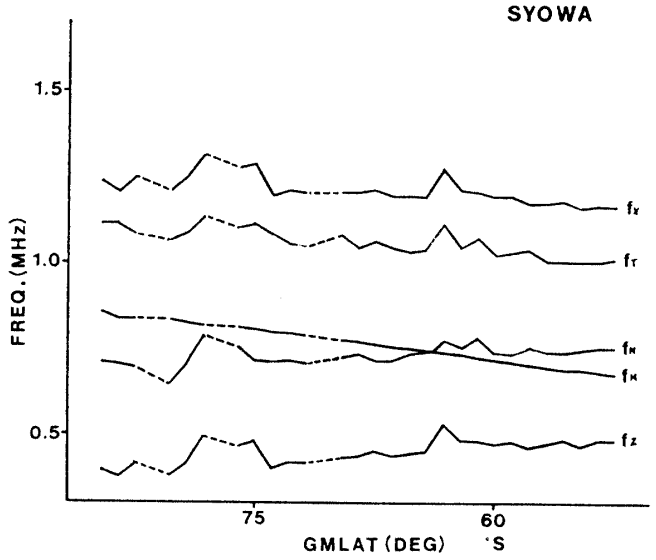


Fig. 10. Simulation results of the resonance spike identification algorithm applied to the ionograms recorded at Syowa Station. The broken line indicates the unidentified points.

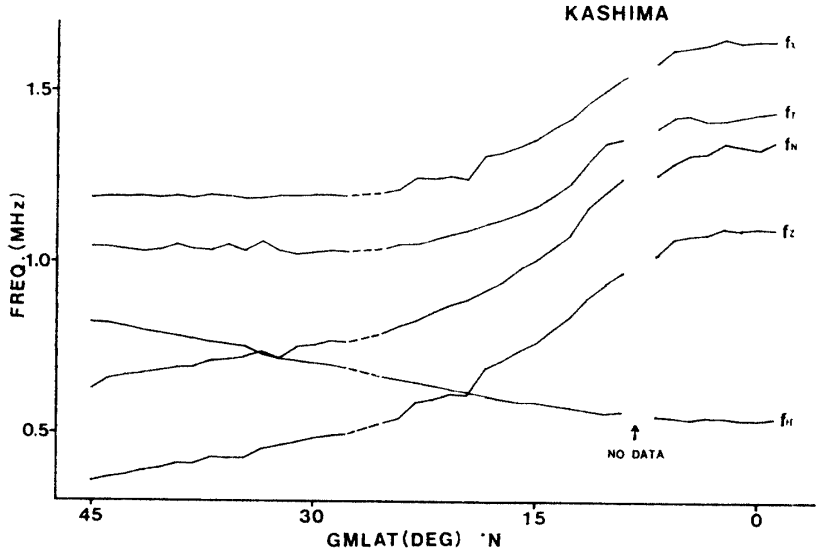


Fig. 11. Simulation results of the resonance spike identification algorithm applied to ionograms recorded at Kashima Station. The broken line indicates the unidentified points.

6. Conclusion

This paper describes the procedure for automatic reduction of N(h) profiles from topside ionograms. As a first step towards the development of the automatic reduction system, the production of digital ionograms, the compression of the data contained in digital ionograms, and the identification method of resonance spikes are described and some results of computer simulation are given.

The production of digital ionograms occupies the most important part of the reduction process. On the bilevel data compression problem, it is concluded that the run-length coding method 3 is most effective among the four coding methods. Considerable information, however, is removed by the bilevel coding of ionograms, so that the multilevel coding method has to be developed as a next step. The algorithm to identify resonance spikes appears to be useful and practical, and the detection of the characteristic frequencies are successfully made by appling this algorithm to some examples of ISIS data acquired at Syowa and Kashima Stations.

The determination of critical frequencies, scaling of echo traces, and evaluation of N(h) profiles with the aid of computer are left for a future study.

References

BENELLI, G., CAPPELLINI, V. and LOTTI, F. (1980): Data compression techniques and applications. Radio Electron. Eng., **50**, 29–53.

HOJO, H. (1969): Toppusaido ionoguramu kaiseki (Reduction of topside ionograms). Denpa Kenkyusho Kihô (Rev. Radio Res. Lab.), **15**, 292–310.

HOJO, H. and NISHIZAKI, R. (1971): Reduction of topside ionograms for field-aligned propagation paths. Nature, Phys. Sci., **233**, 121–123.

HUANG, T. S. (1977): Coding of two-tone images. IEEE Trans. Commun. Technol., COM-**25**, 1406–1425.

MATUURA, N., OGATA, T., NISHIZAKI, R., NAGAYAMA, M., IGI, S., MARUYAMA, T., YAMANISHI, M., NISHIYAMA, I. and IDE, T. (1978): Denrisô kansoku eisei (ISS) kenkyû unyô sisutemu (Satellite operation planning and data analysis system for ISS (Ionosphere Sounding Satellite)). Denpa Kenkyusho Kiho (Rev. Radio Res. Lab.), **24**, 121–146.

MULDREW, D. B. (1972): Electron resonances observed with topside sounders. Radio Sci., **7**, 779–789.

MUSMANN, H. G. and PREUSS, D. (1977): Comparison of redundancy reducing codes for facsimile transmission of documents. IEEE Trans. Commun., COM-**25**, 1406–1424.

NETRAVALI, A. N. and LIMB, J. O. (1980): Picture coding: a review. Proc. IEEE, **8**, 366–406.

PREUSS, D. (1975): Two dimensional facsimile source encoding based on Markov model. Nachrichtentech. Z., **28**, 358–363.

TUNER, L. F. (1973): Data compression techniques as a means of reducing the storage requirement for satellite data: a quantitative comparison. Radio Electron. Eng., **43**, 599–608.

(Received November 30, 1981; Revised manuscript received January 22, 1982)