SIMILAR DATA RETRIEVAL FROM ENORMOUS DATASETS ON ELF/VLF WAVE SPECTRUM OBSERVED BY AKEBONO

Y Kasahara*, A Hirano, and Y Takata

Kanazawa University, Kakuma-machi, Kanazawa-shi, Ishikawa, 920-1192, Japan
**Email: kasahara@is.t.kanazawa-u.ac.jp*

ABSTRACT

As the total amount of data measured by scientific spacecraft is drastically increasing, it is necessary for researchers to develop new computation methods for efficient analysis of these enormous datasets. In the present study, we propose a new algorithm for similar data retrieval. We first discuss key descriptors that represent characteristics of the VLF/ELF waves observed by the Akebono spacecraft. Second, an algorithm for similar data retrieval is introduced. Finally, we demonstrate that the developed algorithm works well for the retrieval of the VLF spectrum with a small amount of CPU load.

Keywords: Similar data retrieval, Database, VLF spectrum, Akebono, Event finding system

1 INTRODUCTION

Many rockets and spacecraft have been launched into various regions of the Earth's ionosphere and magnetosphere. In the field of geophysics, most research is performed by the following methods: (1) a scientist discovers an observation result as evidence of new theories or (2) a scientist tries to explain a very unique observation result theoretically. In recent years, the total amount of data measured by scientific spacecraft has drastically increased as the number of scientific spacecraft increases and the resolution of each instrument becomes higher. As the amount of data increases, it is necessary for researchers to develop new computation methods for the efficient analysis of the enormous datasets because it is almost impossible to survey all datasets manually.

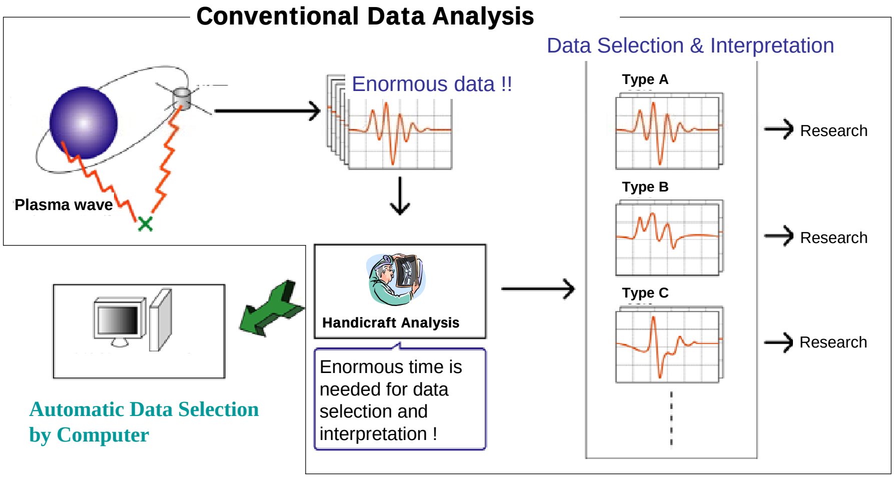

Conventional Data Analysis
Data Selection & Interpretation
Enormous data !!
Type A
Research
Plasma wave
Type B
Research
Type C
Research
Handicraft Analysis
Enormous time is needed for data selection and interpretation !
Automatic Data Selection by Computer

Figure 1. Development of an intelligent database system

Our aim is to develop a new computational technique to discover interesting and/or epoch-making datasets from enormous datasets (Figure 1). Many types of computational methods have already been devised and tested for achieving this. Some of these were developed based on learning algorithms such as neural network and pattern recognition. These algorithms work very powerfully to retrieve the required datasets from enormous databases if the characteristics of the required datasets are well known. However, it is also important for scientists in geophysics and space physics to discover new and unusual phenomena. Furthermore, applications of unsupervised learning algorithms are also being studied. Representative algorithms based on unsupervised learning are categorized into several algorithms, such as self-organizing maps (SOM), cluster analysis, principal component analysis, reinforcement learning, and so on.

In the field of geophysics and space science, trials of applying intelligent computational techniques to complex and enormous datasets have been undertaken. Takano et al. (1999) proposed a learning algorithm based on learned wavelet filters for the detection of the occurrence time of geomagnetic sudden commencement (SC). Higuchi et al. (2000) and Nagao et al. (2002) have applied a procedure based on the concept of spline fitting for automatic detection of large-scale field-aligned currents (LSFACs) and geomagnetic jerks, respectively. Kasahara et al. (2002) introduced a computational method for extracting physical parameters from datasets obtained by particle detectors onboard spacecraft. These applications are, however, basically categorized into learning algorithms. On the other hand, Higuchi et al. (2002) proposed a new method for Pi2 onset time determination with information criterion. Ueno et al. (2002) demonstrated that the Expectation-Maximization (EM) algorithm is applicable to three dimensional particle datasets for separation of multi-component particle distribution. It is noted that these techniques are applicable to any non-stationary datasets.

In the present study, we introduce a new algorithm for similar data retrieval. This method has evolved from an automatic classification algorithm proposed by Akimoto et al. (2003). In the previous work done by Akimoto et al. (2003), cluster analysis was introduced in order to classify the spectrum data observed by the Akebono spacecraft into various kinds of ELF/VLF wave phenomena. It was demonstrated that this algorithm worked quite well for classification of wave phenomena. However, problems have remained: (1) it is quite difficult to determine the appropriate number of wave phenomena to be classified and (2) classification boundaries are sometimes unclear because it is not out of the ordinary for several wave phenomena to be observed simultaneously in the same time and frequency domain. In order to solve these problems, in the present study we propose a more flexible algorithm. In Section 2, a concept of the proposed algorithm for similar data retrieval is introduced. In Sections 3 and 4, we describe an overview of the newly developed “event finding system” and its algorithm, respectively. The performance of our prototype system is demonstrated in Section 5, and we summarize our study in Section 6.

2 CONCEPT OF SIMILAR DATA RETRIEVAL

2.1 DATASETS ON ELF/VLF WAVES OBSERVED BY AKEBONO

Akebono is a Japanese scientific spacecraft that was launched in February, 1989, for observations of the Earth’s magnetosphere. The VLF instruments onboard Akebono are designed to investigate plasma waves from a few Hz to 20 kHz (Kimura et al., 1990). Many kinds of plasma waves have been observed in the Earth's magnetosphere by the VLF instruments. Some of these are artificial waves propagating from the ground while the others are natural waves generated in the space plasma. As the spectrum of each wave is attributed to various generation mechanisms and propagation modes, it is quite useful to observe these waves and clarify their characteristics in deriving the plasma environment as well as reconstructing the spatial structure of the Earth’s magnetosphere.

Akebono has been in operation for almost 20 years, and the mission data obtained from it runs to ~1.7 Tbytes in total in digital format. Akebono carries eight scientific instruments, and the data assigned to the VLF instruments account for one third of them. The VLF instruments consist of five subsystems. One of the subsystems is the MCA (multi-channel analyzer), which measures the wave spectrum in electric and magnetic fields from 3.18 Hz to 17.8 kHz with 16 channels of band-pass filters and a time resolution of 0.5-second (Hashimoto et al., 1997). Another VLF subsystem is the wide-band receiver (WBA), which produces analogue telemetry data. These analogue data are recorded on magnetic tapes in the amount of ~20,000. Their digitization process is now in progress, and we will finally obtain ~18 Tbytes of digital data from the original analogue data.

The data product amounts are summarized in Table 1. We also show the amount of data to be produced by the KAGUYA spacecraft. KAGUYA (SELENE) is a Japanese moon orbiter launched in September, 2007. The WFC (waveform capture) is a subsystem of the LRS (Lunar Radar Sounder) onboard KAGUYA (Ono et al., 2008; Kasahara et al., 2008). The WFC was also developed to measure the ELF/VLF wave, measuring electric wave signals from 100Hz to 1MHz. As shown in Table 1, the total amount of data produced by KAGUYA is much larger than that from Akebono.

Table 1. Examples of amounts of data products by Akebono and KAGUYA spacecraft

Akebono	Total amount of digital telemetry	240 MB/day	~1.7 TB (since 1989)
	VLF/MCA (a part of digital telemetry)	2 MB/day	~14 GB (since 1989)
	VLF/WBA (analogue telemetry)	2.5 GB/day	~18 TB (since 1989)
KAGUYA	Total amount of telemetry	10 GB/day	~3.6 TB/year
	LRS/WFC	1 GB/day	~360 GB/year

In the present study, we develop a prototype of a similar data retrieval system using spectrum data obtained by the MCA onboard Akebono. In this retrieval system, a user can retrieve wave spectra similar to that which he/she has selected as reference data. We demonstrate that the developed algorithm works well for the retrieval of the VLF spectrum obtained by the MCA. We also evaluate the computation time for data retrieval, taking this into account for a future application to much larger datasets such as Akebono-WBA and KAGUYA WFC data.

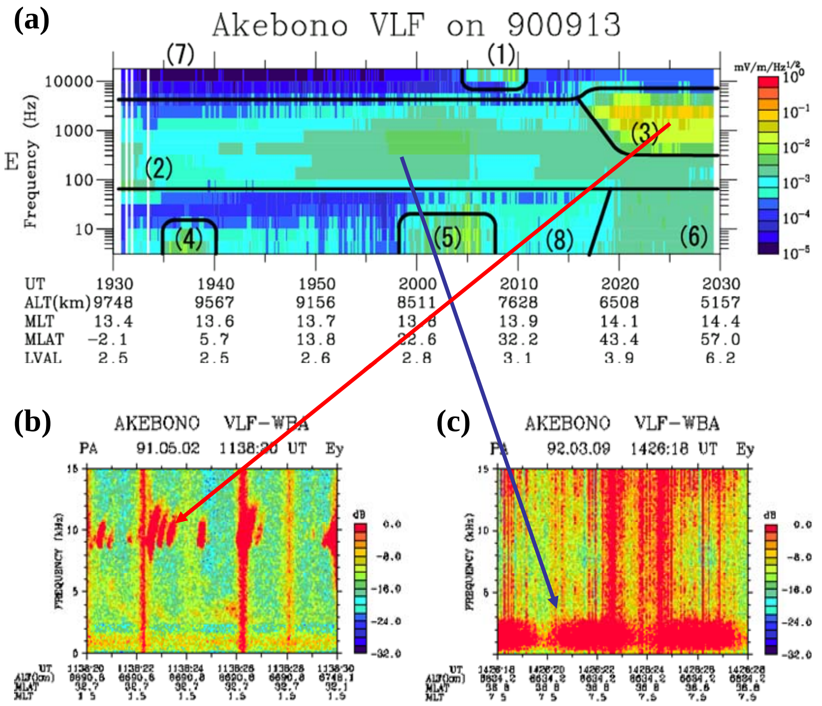


Figure 2. An example of wave spectrum observed by Akebono: (a) A one-hour plot of wave spectrogram for the electric field measured by the MCA, (b) a ten-second plot of the wave spectrogram of the WBA when chorus emissions were observed, and (c) a ten-second plot of the wave spectrogram of the WBA when magnetospheric hiss were observed.

Figure 2 shows an example of the wave spectrogram observed by Akebono. Figure 2(a) is a one-hour plot of the electric wave spectrum measured by the MCA on September 13, 1990. The horizontal and vertical axes represent time and frequency, respectively. The logarithmic color scales are given on the right, with the most intense wave intensities indicated in red. Several kinds of wave phenomena were observed during this period. We have carefully checked the full-precision data with a time resolution of 0.5 second and manually separated the spectrogram into eight cells as shown in Figure 2(a). The cells labeled from (1) to (6) in the figure are natural wave phenomena detected during this period. On the other hand, the cells labeled (7) and (8) are a domain where no wave phenomenon was detected although some faint wave activity was recognized. The wave phenomena observed in the cells labeled (2) and (3) are known to be “plasmaspheric hiss” and “chorus emission,” respectively, based on their structural features of wave spectra. In order to show the difference between wave characteristics, examples of the magnified spectrogram of “chorus emission” and “plasmaspheric hiss," which were obtained on May 2, 1991, and March 9, 1992, are shown in Figures 2(b) and 2(c), respectively. These are ten-second plots of the wave spectrogram measured by the WBA. We find several spectrum elements that are likely to be rubbed with a brush around 8-13 kHz in Figure 2(b). These spectrum elements are called "chorus emission." On the other hand, there are broadly extended spectra around 0.5-3 kHz in Figure 2(c). This is called "plasmaspheric hiss." As shown in Figure 2(c), plasmaspheric hiss is characterized by a moderate variation of wave intensity both in the time and frequency domain. As shown in Figures 2(b) and 2(c), discrimination of wave phenomena is easy if we carefully examine wave spectra drawn in full-precision resolution. However, it is almost impossible for us to examine whole datasets given by full-precision drawings by hand crafted analysis. In particular, if we unexpectedly discover an unusual spectrum, it is absolutely impossible to find manually similar events from the other datasets.

2.2 KEY DESCRIPTORS FOR PLASMA WAVE DATA

In this section we present key descriptors that are effective in distinguishing the wave phenomena measured by the MCA. As was shown in Figures 2(b) and 2(c), wave phenomena observed in the Earth's magnetosphere are basically classified according to the frequency range and structural features of wave spectra, such as narrowband or broadband in the frequency domain and fine or moderate variation in the time domain. Therefore, if we could define these features quantitatively, similar data retrieval using a computational method can be realized.

In the present study, we introduce eight kinds of key descriptors to describe features of wave phenomena observed by the MCA. These key descriptors are determined based on criteria used for discrimination of wave phenomena and are defined from a physical point of view. The eight descriptors are shown in Figure 3. The original data used in Figure 3 are completely identical with those shown in Figure 2(a). The time resolution of each key descriptor is one minute. The frequency resolution is the same as for the original MCA data (16 frequency points). Because it is not our main purpose in the present paper to describe detailed discrimination criteria, we mention the meanings of these key descriptors only briefly.

Figure 3(a) is a contour map of "averaged wave intensity" in the electric field measured by the MCA. This panel is almost the same as Figure 2(a) except that the data below noise level from the MCA receiver are eliminated. The second key descriptors are "time variation of wave intensity." These parameters are calculated by applying the Fourier transformation (FFT) to the MCA data for a 64 second duration at each frequency point, that is, a 128-point FFT is performed on the time series of the MCA data as the time resolution of the MCA is 0.5 second. We can obtain 64 parameters at different time scales by performing the FFT. We average these into four categories (0-0.2Hz, 0.3-0.45Hz, 0.55-0.7Hz, and 0.8-0.95Hz) and adopt them as key descriptors as shown in Figures 3(b), (c), (d), and (e), respectively. These descriptors become large for "chorus emission" because chorus emissions consist of discrete elements as shown in Figure 2(b) while they become small for "plasmaspheric hiss" because the time variation of wave intensity is quite moderate (see Figure 2(c)). It is noted that we eliminate the parameters and harmonics from these key descriptors at a time scale around 0.25Hz. This is because 0.25Hz corresponds to the half of spin cycle (The spin period of Akebono is 8 seconds), and it is an important parameter in the physical point of view. Therefore, we separately define it as a key descriptor for "spin modulation" as shown in Figure 3(f). As wave intensity is basically modulated by a spin period because the antennas mounted on the spacecraft rotate with a spin period, the degree of intensity modulation at 0.25Hz is closely related to the propagation direction of the wave. Figure 3(g) shows a key descriptor for the ratio of wave intensity between electric and magnetic fields. This descriptor closely corresponds to the refractive index of the wave phenomenon. Figure 3(h) is a key descriptor for the dispersion of wave intensity. This descriptor characterizes the dynamic range of wave intensity.

By using these key descriptors, we can characterize wave phenomena quantitatively. We can define any type of key descriptor as shown in Figure 3, but needless to say, it is very important to introduce parameters that are meaningful for the discrimination of wave phenomena.

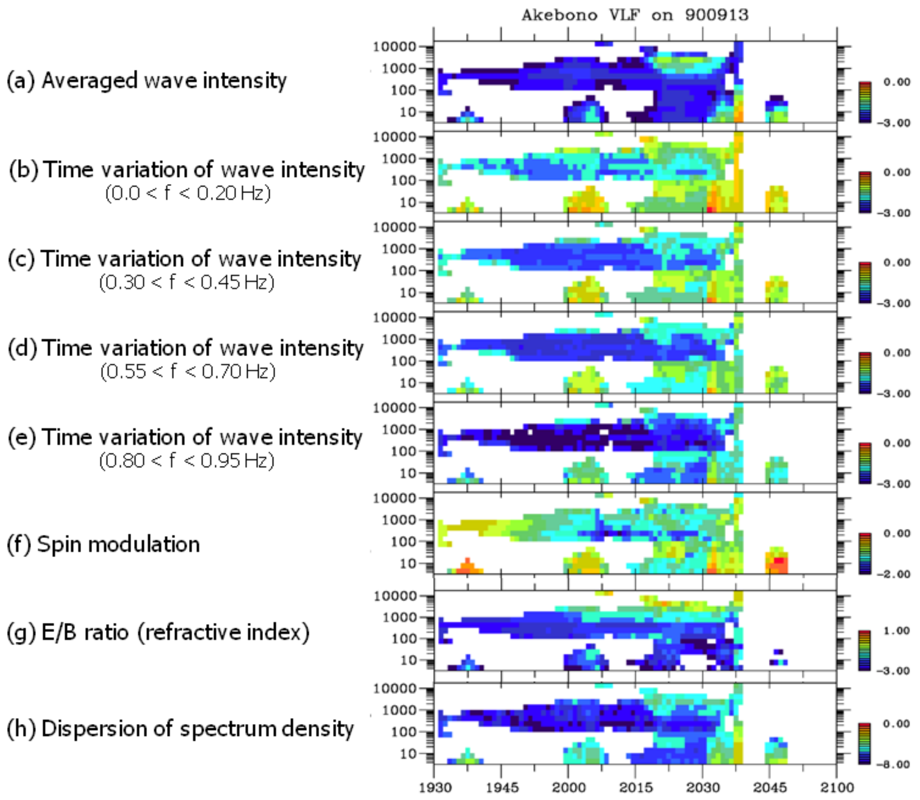


Figure 3. The key descriptors used for the proposed similar data retrieval system

3 OVERVIEW OF THE EVENT FINDING SYSTEM

The purpose of the present study is to develop a computational technique to retrieve datasets similar to the one the user has selected as a reference datum. We call this data retrieval system an “event finding system.” In the design of the event finding system, we took the generality of the system into account in order to apply it not only to the MCA data but also to other kinds of datasets in future work. A block diagram of the event finding system is shown in Figure 4. The left part of the block diagram in Figure 4 is designed for data registration. In this part, we calculate several kinds of key descriptors from original observational data. In our prototype system using the MCA datasets, we introduce eight kinds of key descriptors as shown in Figure 3 with a time resolution of one minute (See Section 2.2). As was discussed in Section 2.2, it is very important to choose parameters of key descriptors that are meaningful for dataset discrimination. Because key descriptors strongly depend on datasets, appropriate key descriptors should be examined independently, adjusting the characteristics of the datasets.

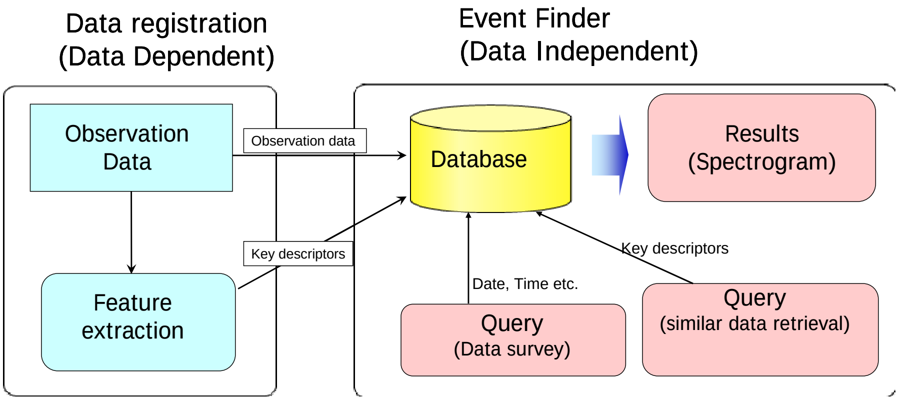


Figure 4. A block diagram of the developed event finding system

The right part of the block diagram in Figure 4 plays the role of event finder, which takes a general account of the system. Original datasets and their key descriptors are registered in the database, but the functions for database management are designed to work independently of the dataset characteristics. There are two kinds of queries for data retrieval: one extracts reference data from original datasets using date and time search keys and the other uses key descriptors as search parameters for similar data retrieval. Applying these queries to the database, search results are provided in the form of spectrograms.

Figure 5 shows a user interface and applications used in the event finding system. A web browser accepts requests from the user and returns the requested data. The event finder system was developed on the Linux operation system (OS) in which Apache and PostgreSQL were installed as web server and database management system (DBMS), respectively. Tomcat and Java servlet are used for the interface between the web server and the DBMS.

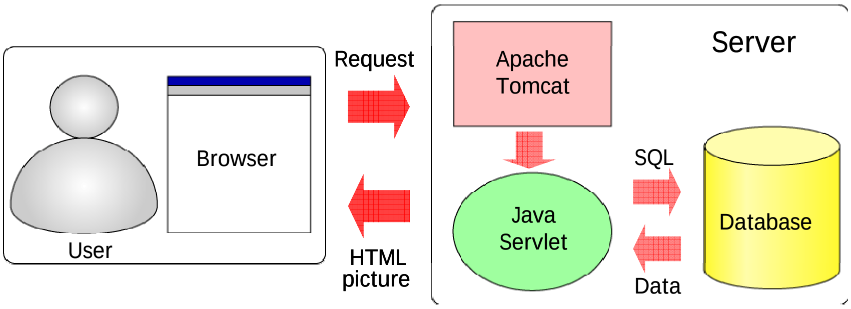
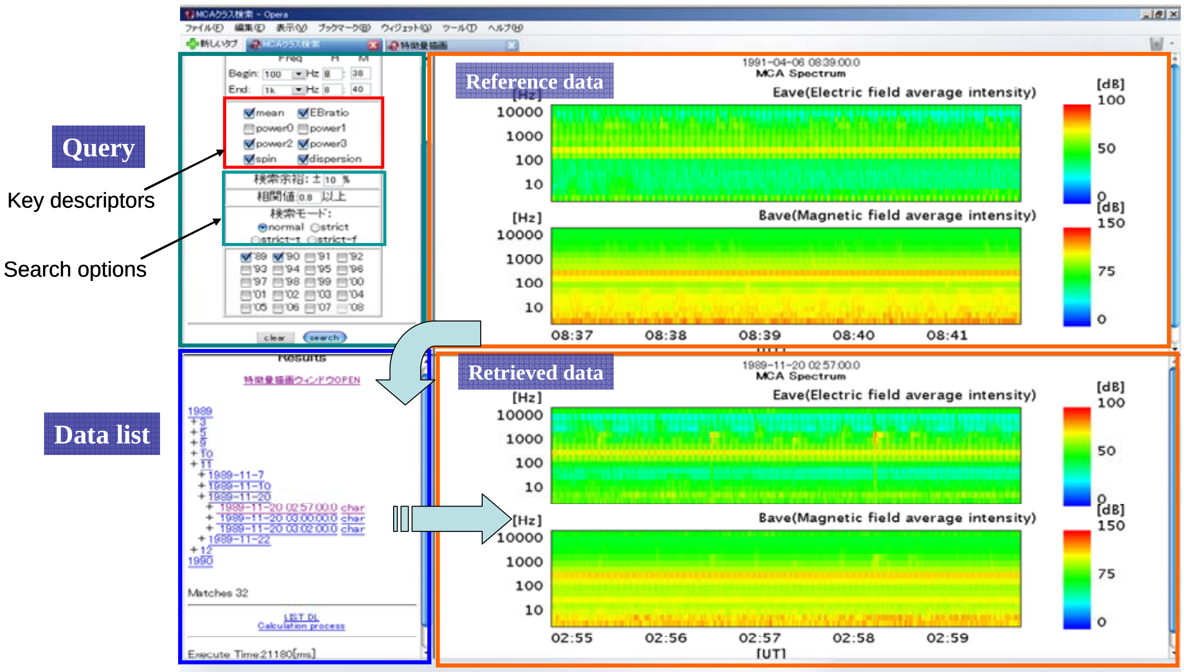


Figure 5. A user interface and applications in the event finding system



Freq H M
Begin: 100 Hz 8 : 38
End: 1k Hz 8 : 40

☑mean ☑EBratio
☐power0 ☐power1
☑power2 ☑power3
☑spin ☑dispersion

検索余裕: ± 10 %
相関値 0.8 以上
検索モード:
● normal ○ strict
○ strict-t ○ strict-f

☑'89 ☑'90 ☐'91 ☐'92
☐'93 ☐'94 ☐'95 ☐'96
☐'97 ☐'98 ☐'99 ☐'00
☐'01 ☐'02 ☐'03 ☐'04
☐'05 ☐'06 ☐'07 ☐'08

1989
+3
+5
+9
+10
+11
+1989-11-7
+1989-11-10
+1989-11-20
+1989-11-20 02:57:00.0 char
+1989-11-20 03:00:00.0 char
+1989-11-20 03:02:00.0 char
+1989-11-22
+12
1990

Matches 32

Execute Time21180[ms]

Figure 6. A snapshot of the event finding system

Figure 6 shows a snapshot of the event finding system, an overview of which is described in this section. As was introduced in the previous section, observation data can be characterized by several key descriptors. In the event finding system, key descriptors with original datasets obtained by the MCA are registered in a database system. First, a user designates the date and time in which he wishes to browse; then a corresponding spectrogram of the MCA is provided in the upper right panel as a reference spectrogram (see Figure 6). Second, when a user finds an interesting spectrum in the reference spectrogram, he designates a rectangular subregion to include the interesting spectrum by inputting the lowest and highest frequencies and the start and end times in the upper left panel. He can also choose key descriptors to be used for a similar data retrieval process (shown in the red rectangle in the upper left panel in Figure 6) and several options (shown in the green rectangle in the upper left panel in Figure 6; see Section 4 for further description). When a query is executed using these parameters, the system replies with a list of the dates and times of datasets that are similar to the requested reference spectrum in the lower left panel. Finally, the retrieved spectrogram is given in the lower right panel when one of datasets listed in the lower left panel is selected.

4 ALGORITHM OF SIMILAR DATA RETRIEVAL

In this section, a detailed algorithm for similar data retrieval applied to the system is described. In this paper, we describe a prototype system using spectrum data obtained by the MCA. As shown in Table 1, the total number of MCA datasets is very small when compared with other datasets such as WBA and WFC. However, the number of total records registered in the database amounts to ~4 x 10^7 when we register key descriptors at each "grid" on the MCA spectrogram with a time resolution of one minute in the horizontal axis and at 16 frequency points in the vertical axis for all the datasets measured since 1989. In our system, a query for similar data retrieval is given by a rectangular subregion specified with the coordinates of its four corners: the lowest and highest frequency and the start and end times. In the following, we define "reference subregion" as reference data specified by user and "test subregion" as the test data to be compared with the "reference subregion."

It is obvious that it takes tremendous computation time to compare one by one all test subregions with a reference subregion. In order to search appropriate datasets similar to the reference data within a finite computation time, we adopted the following two step search algorithm, a schematic diagram of which is shown in Figure 7. The pink rectangle in the left bottom panel of Figure 7 is an example of a "reference subregion" while the pink rectangle in the right bottom panel of Figure 7 is a "test subregion" to be compared with the "reference subregion." In the first step (Step 1), candidates for similar datasets are selected by matching a set of key descriptors at the central point (central grid) of a "test subregion" with the set of key descriptors at the central point of the "reference subregion." The user can provide a margin of values to be used for matching. In the second step (Step 2), the correlation coefficient between each candidate for the test subregion and the reference subregion is calculated one by one, and a list of candidates whose correlation coefficients are higher than a given threshold is produced. Finally, a list of datasets, which include spectra similar to the requested reference spectrum, is presented to the user. The threshold used in the second step can also be changed according to the user's research purpose. Using the proposed two step search algorithm, computation time is drastically reduced when compared with full pattern matching of whole datasets.

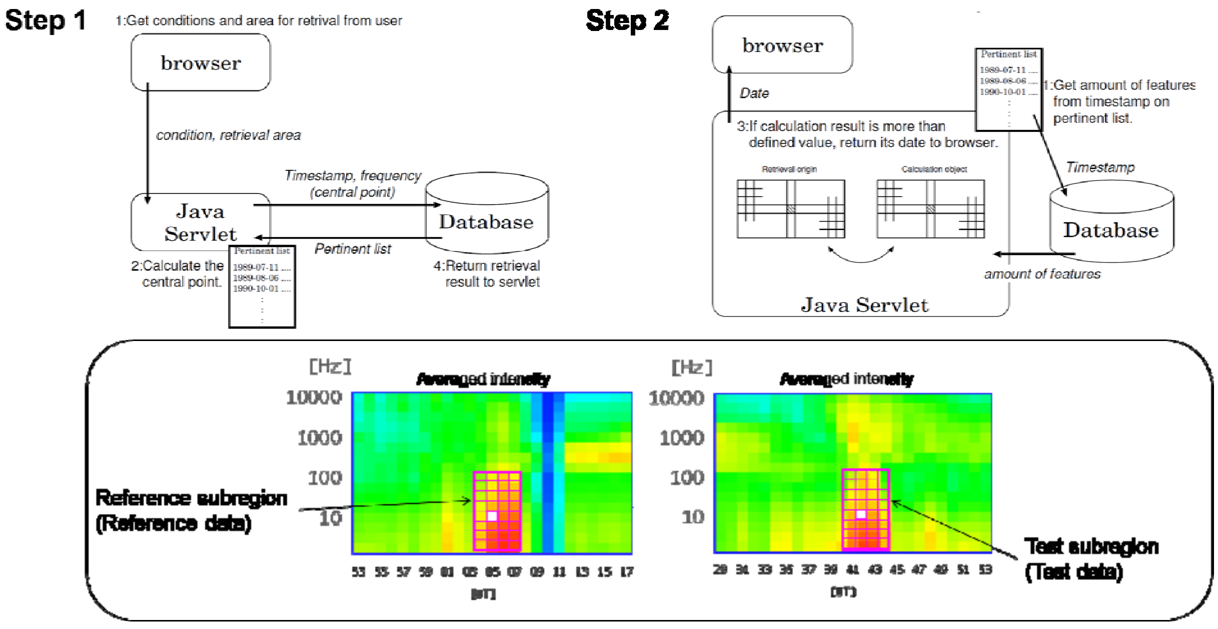


Figure 7. A conceptual diagram of the two step search algorithm used in the similar data retrieval process

5 EVALUATION OF THE SYSTEM

In this section, we demonstrate the performance of our prototype similar data retrieval system. Figure 8 shows examples of retrieval results. The upper panels show the reference spectrogram for chorus emissions observed on April 6, 1991 (on the left side) and the plasmaspheric hiss observed on the same day (on the right side), respectively. The lower panels are the spectrograms retrieved by the event finding system. As shown in Figure 8, similar spectrograms are successfully retrieved in both cases. It is noted that any type of spectrum including a non-stationary spectrum is applicable to the proposed algorithm as reference data because this algorithm does not require any preliminary knowledge of the wave phenomenon. Thus, we can even apply this algorithm to a newly discovered spectrum in order to search similar datasets among the whole datasets.

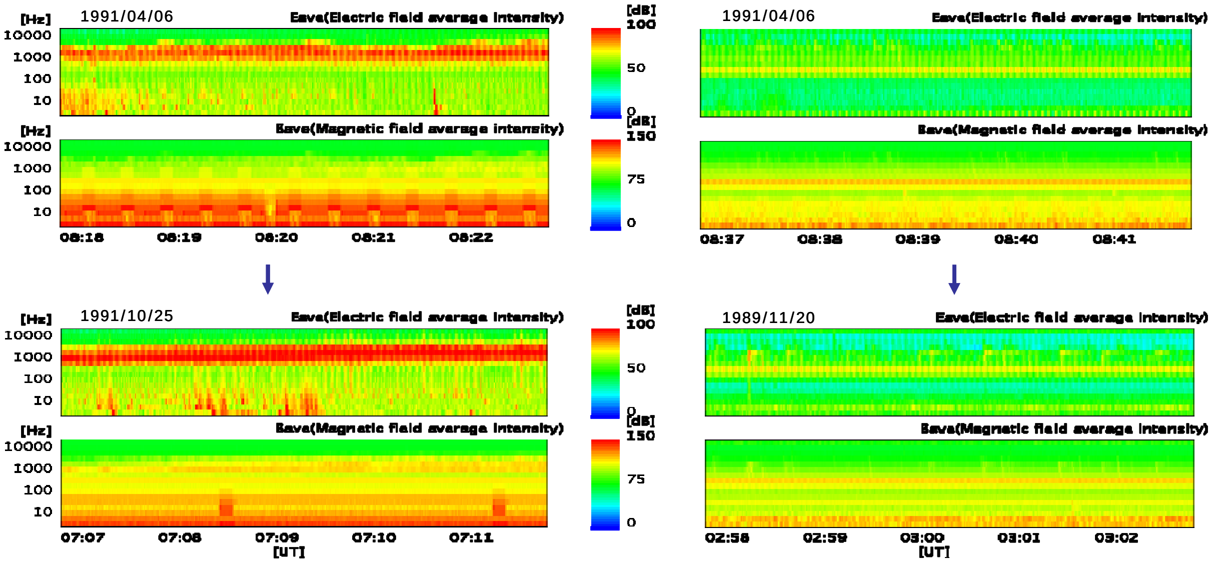


Figure 8. Examples of similar data retrieval for chorus (left) and plasmaspheric hiss (right). The upper panels show reference data and the lower panels show the retrieved results.

Finally we evaluate the computation time of our prototype system. The configurations of the developed system are shown in Table 2. As shown in Table 2, we developed our prototype system using open source products on a general-purpose personal computer.

As was mentioned in Section 4, we registered key descriptors with a time resolution of 1 minute for all the MCA datasets measured from 1989 to 2008. The number of records registered in the database is shown on the right vertical axis in Figure 9. The horizontal axis in Figure 9 shows the elapsed years since 1989, when Akebono was launched. We found that the total number of registered records amounts to ~4 x 10^7. Computation times for retrieval are also indicated on the left vertical axis in Figure 9. In the evaluation, four types of typical wave phenomena, including chorus emission and plasmaspheric hiss, were examined. The error range shown in Figure 9 indicates the standard deviation of the retrieval times. As the elapsed year becomes greater, the error range becomes larger because the number of candidates selected in “Step 1” drastically changes depending on the reference data. However, it was found that we can retrieve similar datasets for ~3 minutes even though we survey all records. This computation time is small enough for practical use.

Table 2. Specifications of the similar data retrieval system developed in the present study

Hardware Specification		Software Specification	
Operating System	Vine Linux 3.7	Web Server	Apache 2.2.6
CPU	Intel Pentium 4 (2.80 GHz)	Servlet Container	Tomcat 5.5.25
Cache Size	1024 kbytes	DBMS	PostgreSQL 8.2.5
Memory	1024 Mbytes	Java Development Kit	JDK 1.5.0.13

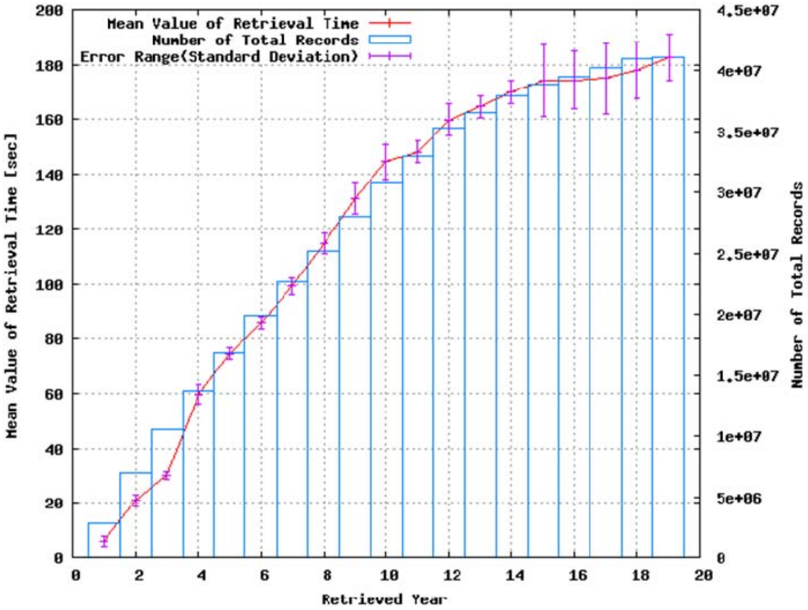


Figure 9. Time spent for similar data retrieval (left vertical axis) and total number of data records registered in the system (right vertical axis) versus elapsed year (horizontal axis) since 1989, when Akebono was launched

6 CONCLUSION

In the present study, we propose a new algorithm for similar data retrieval applicable to the ELF/VLF spectrum measured by Akebono. In order to manage these enormous datasets effectively, we first constructed a database system on VLF/ELF waves obtained by the MCA onboard Akebono. Second, we introduced several kinds of key descriptors, such as wave intensity, time variation of wave spectrum, and ratio between electric and magnetic wave components, in order to describe the distinctive wave spectrum features. We applied a proposed algorithm for similar data retrieval by storing these key descriptors in the database. Finally, we evaluated the performance of our system and concluded that the developed algorithm works well for the purpose of similar data retrieval, and its computation time is also small enough for practical use. It is especially notable that the generality of the system is taken into account so that the proposed method is applicable not only to MCA data but also to other kinds of datasets. In the future, an evaluation of the accuracy of the retrieval result and computation time using enormous datasets such as WBA onboard Akebono and WFC onboard KAGUYA should be undertaken.

7 ACKNOWLEDGEMENTS

This research was supported by the Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science (20510006).

8 REFERENCES

Akimoto, Y., Y. Goto, Y. Kasahara, & T. Sato, (2003), Automatic Classification of Electromagnetic Waves from Database Obtained by the Akebono Satellite (in Japanese), *Trans. IEICE Japan*, *J86-D2*(5), 598-607.

Higuchi, T., & S. Ohtani, (2000), Automatic Identification of a Large-scale Field-aligned Current Structures, *J. Geophys. Res.*, *105(A11)*, 25305-25315.

Higuchi, T., S.-I. Ohtani, T. Uozumi, & K. Yumoto, (2002), Pi2 onset time determination with information criterion, *J. Geophys. Res.*, *107(A7)*, doi:10.1029/2001JA003505.

Hashimoto, K., I. Nagano, M. Yamamoto, T. Okada, I. Kimura, H. Matsumoto, & H. Oki, (1997), Exos-D (Akebono) very low frequency plasma wave instruments (VLF), *IEEE Trans. Geoelectr. Remote Sens.*, *35*(2), 278-286.

Kasahara, Y., R. Niitsu, & T. Sato, (2002), Computational Analysis of Plasma Waves and Particles in the Auroral Region Observed by Scientific Satellite, *Progress in Discovery Science*, *Lecture Notes in Computer Science*, 2281, Springer, 426-437.

Kasahara, Y., Y. Goto, K. Hashimoto, T. Imachi, A. Kumamoto, T. Ono, & H. Matsumoto, (2008), Plasma Wave Observation Using Waveform Capture in the Lunar Radar Sounder on board the SELENE Spacecraft, *Earth, Planets and Space*, *60*(4), 341-351.

Kimura, I., K. Hashimoto, I. Nagano, T. Okada, M. Yamamoto, T. Yoshino, H. Matsumoto, M. Ejiri, & K. Hayashi, (1990), VLF observations by the Akebono (Exos-D) satellite, *J. Geomagn. Geoelectr.*, *42*(4), 459-478.

Nagao, H., T. Higuchi, T. Iyemori, & T. Araki, (2002), Automatic detection of geomagnetic jerks by applying a statistical time series model to geomagnetic monthly means, Progresses in Discovery Science, *Lecture Notes in Computer Science, 2281*, Springer, 360-371.

Official page of the Apache Jakarta Project. Retrieved from the World Wide Web, February 4, 2010: http://jakarta.apache.org/

Official page of the Apache Software Foundation. Retrieved from the World Wide Web, February 4, 2010: http://www.apache.org/

Official page of the Apache Tomcat, Retrieved from the World Wide Web, February 4, 2010: http://www.apache.org/

Official page of the PostgreSQL, Retrieved from the World Wide Web, February 4, 2010: http://www.postgresql.org/

Ono, T., A. Kumamoto, Y. Yamaguchi, A. Yamaji, T. Kobayashi, Y. Kasahara, & H. Oya, (2008), Instrumentation and Observation Target of the Lunar Radar Sounder (LRS) Experiment on-board the SELENE Spacecraft, *Earth, Planets and Space*, *60*(4), 321-332.

Takano, S., T. Minamoto, H. Arimura, K. Niijima, T. Iyemori, & T. Araki, (1999), Automatic detection of geomagnetic sudden commencement using lifting wavelet filters, *Discovery Science (Proceedings of the Second International Conference on Discovery Science), Lecture Note in Artificial Intelligence*, 1721, Springer, 242-251.

Ueno, G., N. Nakamura, T. Higuchi, T. Tsuchiya, S. Machida, & T. Araki, (2002), Application of multivariate Maxwellian mixture model to plasma velocity distribution, Progresses in Discovery Science, *Lecture Notes in Computer Science, 2281*, Springer, 372-383.

(Article history: Received 31 January 2009, Accepted 4 February 2010, Available online 12 February 2010)