

Satellite Image Database Search Engine which Allows Fuzzy Expression of Geophysical Parameters of Queries

Kohei Arai

Faculty of Science and Engineering
Saga University, Saga City, Japan

Abstract—Satellite image database search engine which allows fuzzy expression of geophysical parameters of queries is proposed. A search engine based on knowledge based system which allows fuzzy expression of queries is proposed. A prototype system is created and tested. Whereas conventional search systems had to know in advance the functions of the search system, information about search keys, etc., the search engine proposed here guided the search conditions in a conversational form, by allowing ambiguous expressions (six adverb language hedges) at that time, the user is released from such annoyance. To make this possible, a membership function for each attribute information is defined, and a search condition refinement by fuzzy logic is introduced. The results show that the system accepts a fuzzy expression of query as well as a comprehensive dialogues between users and the system.

Keywords—Search engine; fuzzy expression; knowledge base system; membership function

I. INTRODUCTION

The relational database (relational database) is frequently used for searching the earth observation satellite image database from the viewpoint of its efficiency and system flexibility. A typical example of the database search is a search engine using SQL (Structured Query Language) or QBE (Query by Example) [1], [2]. When these search engines are used, constructing search conditions for the earth observation satellite image database search, it is necessary to input a specific search key in a predetermined format. That is, the search can be performed only after the user sets the search condition according to the determined format. Therefore, it is necessary for the user to know in advance the functions of the search system, information on the search key.

The search system for the Earth Observation Satellite Image Database: EOSID proposed in this paper is based on SQL (since most existing EOSID systems have been developed on the premise of SQL), and allows search requests in natural language. The proposed search engine system translates the search request into a request in the SQL language (agent function [3]) and searches for the desired data. At the search engine is used as EOSID, the system guides the user in a conversational manner (guide system) and also to allow search requests with ambiguous expressions of natural language [4]. Therefore, a membership function for ambiguous expressions is introduced based on the opinions of users in the marine research field (for instance) and enabled

fuzzy search. In order to confirm the effectiveness of the proposed method, a virtual database was constructed in the marine research field. Some experiments are performed with the prototype of the proposed search engine system. As a result, it is found that the search performance is quite good from the viewpoint of easiness of database retrievals.

In the following section, related research works are described. Then, the proposed prediction method and system is described followed by experimental set-up together with experimental results. After that, concluding remarks and some discussions are described.

II. RELATED RESEARCH WORKS

Vague search of earth observation image database based on Fuzzy theory using physical quantities and spatial features is proposed [5] together with earth observation satellite image database system allowing ambiguous search requests [6]. On the other hand, user friendly and efficient catalog information management for earth observation data, is proposed and well reported [7].

Remote sensing satellite image database system allowing image portion retrievals utilizing principal component which consists spectral and spatial features extracted from imagery data is proposed [8]. Meanwhile, data collection and active database for tsunami warning system is proposed [9]. On the other hand, remote sensing satellite image database system allowing image portion retrievals utilizing principal component which consists spectral and spatial features extracted from imagery data is proposed [10]. Also, numerical representation of web sites of remote sensing satellite data providers and its application to knowledge based information retrievals with natural language is proposed [11].

Image retrieval based on color, shape and texture for ornamental leaf with medicinal functionality, meanwhile, is proposed [12]. Also, comparison contour extraction based on layered structure and Fourier descriptor on image retrieval is proposed and evaluated its effectiveness [13]. Also, image retrieval method utilizing texture information derived from Discrete Wavelet Transformation: DWT together with color information is proposed [14].

Metadata definition and retrieval of earth observation satellite data is proposed [15]. Meanwhile, Geographic Information System: GIS based on neural network for

appropriate parameter estimation of geophysical retrieval equations with satellite remote sensing data is proposed [16]. Image retrieval method based on hue information and wavelet description based shape information as well as texture information of the objects extracted with dyadic wavelet transformation is proposed [17]. Wavelet based image retrievals is attempted [18]. Also, image retrieval method based on back-projection is proposed [19].

III. PROPOSED METHOD

A. System Configuration

Fig. 1 shows the configuration of the proposed earth observation satellite image database system. The system allows search requests in natural language. Since the research on natural language understanding has not been completed, a text search method has been introduced here. That is, the text matching between the search request from the user and the search keyword prepared by the system is determined by Z39.504, [20].

A natural language whose expression is ambiguous has a broad meaning and is ambiguous. For example, a search may be requested in a vague image of the user, such as "slightly warm sea area" or "image with extremely small amount of cloud". Existing database search engines cannot handle such cases. In order to deal with this, the following fuzzy search is proposed.

B. Fuzzy Retrieval

A fuzzy search for users in the marine field is shown as an example of an ambiguous search. The search target (object [21] of the constructed virtual database in the marine field is an image ID (Identification No.), and the attribute information (Attributes) serving as a search key is as follows.

- Sea surface temperature
- Chlorophyll a
- Cloud cover
- Observation date
- Sea area
- Image quality
- Sensor name
- Satellite name
- Processing level
- Ground station name
- Pass, row number
- Scene center latitude, longitude

Of these, this time, sea surface temperature, chlorophyll-a and cloudiness were considered to allow search requests with ambiguous expressions.

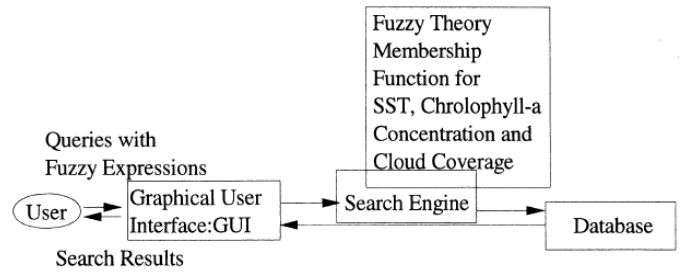


Fig. 1. The Proposed Database System for Earth Observation Satellite Data Retrieval.

C. Creating Membership Functions

A fuzzy set (membership function [22], [23], [24]) was created to express the user's ambiguity regarding the above-mentioned sea surface temperature, chlorophyll-a, and cloudiness. At that time, a questionnaire survey was conducted for researchers in the marine field of the Earth Environment Observation Committee of the Earth Science and Technology Agency, and based on the results (shown in Table I).

Equations (1) to (6) show the membership functions corresponding to low, high, high and low chlorophyll-a concentrations and high and low cloudiness, respectively.

$$\mu_A(x) = \int_0^\infty e^{-0.015x^2/x} \text{ for Low SST} \quad (1)$$

$$\mu_B(x) = \int_0^\infty e^{-0.015(x-30)^2/x} \text{ for High SST} \quad (2)$$

$$\mu_C(x) = \int_0^\infty e^{-0.003(x-35)^2/x} \text{ for High Chlorophyll - a} \quad (3)$$

$$\mu_D(x) = \int_0^\infty e^{-30x^2/x} \text{ for Low Chlorophyll - a} \quad (4)$$

$$\mu_E(x) = \int_0^\infty e^{-0.034(x-17)^2/x} \text{ for High Cloud Coverage} \quad (5)$$

$$\mu_F(x) = \int_0^\infty e^{-3.0x^2/x} \text{ for Low Cloud Coverage} \quad (6)$$

where, the units of the physical quantity (x) are sea surface temperature ($^{\circ}$ C.), chlorophyll-a (mg / m^3), and cloudiness (%), respectively. Next, a linguistic expression (language hedge) that modifies the fuzzy set is defined. As a result, expressions such as "slightly warm" and "very" warm sea surface temperature can be accepted as search requests. The following shows the language hedge used in this prototype system.

- very y
- more or less y
- slightly y
- sort of y
- pretty y
- rather y

where y is a fuzzy set. logical operations using these combinations are also possible, which allows for modification of complex fuzzy sets, and allows search requests using combinations such as "slightly warm, chlorophyll-a sea areas".

TABLE I. THE RESULTS FROM QUESTIONNAIRE FOR DETERMINATION OF MEMBERSHIP FUNCTIONS OF SEA SURFACE TEMPERATURE, CHLOROPHYLL-A CONCENTRATION AND CLOUD COVERAGE

General Item	Physical Item	Attributes	Range	Res1	Res2	Res3	Res4	Res5	Res6
Physical Quantity	SST	Cold	Highest	15	10		10	0	15
		Warm	Lowest	20	20		25	25	25
		Warm Current	Lowest	15	15		20	20	25
			Highest	30	30		28	28	30
	Cold Current	Lowest	0	5			0	5	
		Highest	10	10		10	19	15	
Chlorophyll-a	Denth	Lowest	2		0	0	0	0	
	Sparse	Highest	1			1	1	1	
Area	Japanese Vicinity	North	Lowest	24	25	20	25		25
			Highest	50	50	46	45		54
		East	Lowest	120	120	122	125		120
			Highest	150	150	148	140		150
	Ocean Area	Coastal	Far	20	10		10	50	20
		Offshore	Nearest	360	100		300	100	100
		Beyond	Nearest	360	100		300	100	500
	Kuroshio Current	North	Lowest	10	28		15		25
			Highest	38	38		38		50
		East	Lowest	120	128		130		120
			Highest	150	150		145		145
	Oyashio Current	North	Lowest	35	35		38		30
Highest			50	48		45		50	
East		Lowest	140	142		140		120	
		Highest	160	150		145		145	
Time	Recently		Far		12		12	6	60
	Season	Summer	From	July	June		June	June	July
			To	Sep.	Sep.		Aug.	Aug.	Sep.
Winter	From	Jan.	Nov.		Nov.	Nov.	Dec.		
	To	Mar.	Mar.		Jan.	Feb.	Feb.		
Other	Cloud Coverage	Slightly	Approx.	25	20		0	10	30
		Dense	From	25	50		50	30	60
			To	75	100			100	100

D. Search Procedure

Fig. 2 shows the search procedure. First,

- 1) A search request containing an ambiguous expression occurs,
- 2) Guide the search conditions in a conversational manner,
- 3) Select the membership function using the knowledge corresponding to the search condition,
- 4) Using the membership function in the database and then calculate the degree of attribution of data attributes,
- 5) Display search results.

where, the knowledge corresponding to the search condition is the semantic network shown in Fig. 3. The determination of the degree of belonging is performed as follows. In other words, if the attribute information of the data in the database is between a certain threshold value and the maximum value of the membership function, it is said to belong, and it is determined that the others do not belong. The

determination of the threshold was determined as the grade in the center of gravity (CG) of the fuzzy set shown in equation (7).

$$CG = \frac{\int_x x\mu A(x)dx}{\int_x \mu A(x)dx} \tag{7}$$

where, A is an attribute item, and x is an attribute value (physical quantity).

E. AND, OR Logical Operation of Search Condition

Logical operations (AND, OR) of a plurality of search conditions are also possible. At this time, the degree of membership is given by equation (8).

$$\mu A \cap B = \min [\mu A (x), \mu B (x)] \tag{8}$$

$$\mu A \cup B = \max [\mu A (x), \mu B (x)] \tag{9}$$

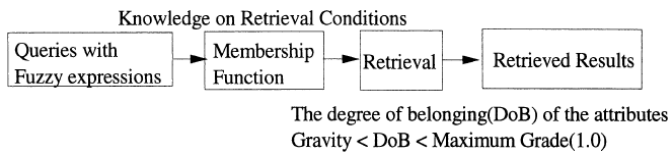


Fig. 2. Search flow of the Proposed Database System which allows Fuzzy Expression of Queries.

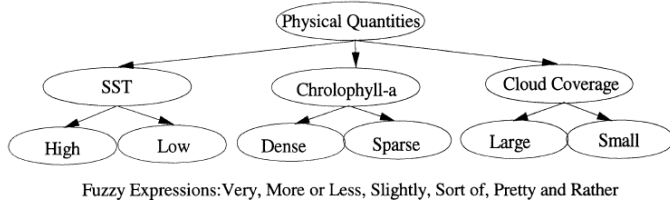


Fig. 3. Semantic Network for the Knowledge of the Proposed Database System.

In other words, when searching for data with a slightly higher sea surface temperature and a very small amount of cloud, use Equation (8) that sets the two search conditions to AND. If any of the conditions is satisfied but data is desired, equation (9) may be used.

IV. EXPERIMENTS

Considering that sea surface temperature, chlorophyll-a, and cloudiness are uniformly distributed in the range of 0 to 40 (°C), 0 to 5 (mg / m³), and 0 to 100 (%), respectively, 1000 pseudo Generated attribute information of earth observation data. Since the purpose here is to evaluate the performance of the ambiguous search, it is assumed that the search has already been completed for attributes other than these three items. Fig. 4 shows the search flow and search results when the following search request is entered.

In Fig. 4, the sentences with bold and underline characters indicates system response. Firstly, the system asks query, then user account and password are asked. After that, SST, chlorophyll-a concentration and cloud coverage are asked by the system. Then the search result with the table which is shown in Fig. 4 is returned from the system. If the user satisfies the search result, then the system asks whether or not the user would like to save the result. After that, the system asks user's willingness on other search request. If the user has no other search request, then the system is closed.

As a result, candidates meeting these search conditions (in this case, seven candidates) were presented, and it was confirmed that the user could further narrow down from these candidates. Experimental result shows the prediction accuracy of retrieve items are 100% of success rate. In terms of functionality of search engine of the current Earth Observation Satellite Image Database Systems, there is no geophysical parameter search based on knowledgebase system with fuzzy search capability. Therefore, the proposed search engine system has the merit on the geophysical research function.

Query: "Slightly Cold, Very Densely Distributed Chlorophyll-a Concentration and Pretty Small Cloud Coverage"

Welcome to Earth Observation Satellite Data (OCTS) Retrieval System

Please key-in your name

: arai

Please key-in your password

:*****

Authentication is completed

What is Sea Surface Temperature do you want to retrieve? (You may use the following words: Very, More or Less, Slightly, Sort of, Pretty and

Rather)

: Slightly cold

What is Chlorophyll-a Concentration do you want to retrieve? (You may use the following words: Very, More or Less, Slightly, Sort of, Pretty and

Rather)

: Very dense

What is Cloud Coverage do you want retrieve? (You may use the following words: Very, More or Less, Slightly, Sort of, Pretty and

Rather):

Pretty small

Now knowledge based retrieval is activated

Just wait for a moment

This is candidate of your retrieved data

Image ID	SST	Chlorophyll-a	Cloud Coverage.
27	8	3	13
174	9	2	11
276	12	3	5
743	15	3	20
835	10	2	18
849	8	2	9
901	14	3	16

Do you satisfy this results? (Yes(y) or No(n))

:y

Do you want save this results? (Yes(y) or No(n))

:n

Do you want to retrieve other data? (Yes(y) or No(n))

:n

Now you may exit this system

Fig. 4. An Example of the Dialogue between users and the Proposed System with the Query of "Slightly Cold, Very Densely Distributed Chlorophyll-a Concentration and Pretty Small Cloud Coverage".

V. CONCLUSION

A search engine based on knowledge based system which allows Fuzzy expression of queries is proposed. A prototype system is created and tested. Whereas conventional search systems had to know in advance the functions of the search system, information about search keys, etc., the search engine proposed here guided the search conditions in a conversational form, by allowing ambiguous expressions (six adverb language hedges) at that time, the user is released from such annoyance.

To make this possible, a membership function for each attribute information is defined, and a search condition refinement by fuzzy logic is introduced. The results show that the system accepts a Fuzzy expression of query as well as a

comprehensive dialogues between users and the system. As a result, candidates meeting these search conditions (in this case, seven candidates) were presented, and it was confirmed that the user could further narrow down from these candidates.

VI. FUTURE RESEARCH WORKS

The proposed system is adopted in the real earth observation satellite database system, and it is a future subject to realize a more usable database system.

ACKNOWLEDGMENT

This research was developed with reference to some of the results obtained by the joint research with the Remote Sensing Technology Center, which was carried out with the support of the National Space Development Agency of Japan. The author thanks Dr. Shinichi Sobue of the National Space Development Agency of Japan, Mitsuhiro Morishita of the Remote Sensing Technology Center, and other concerned parties for their joint research opportunities.

The author, also, would like to thank Professor Dr. Hiroshi Okumura and Professor Dr. Osamu Fukuda for their valuable discussions.

REFERENCES

- [1] Shin-ichi Sobue, Kohei Arai, Bunra Yoshida, Osamu Ochiai, Mina Ogawa, Mineo Sekiguchi, Tomotaka Sekiya, Masao Takagi, User-friendly and efficient catalog information management and provision method for earth observation satellite data, Proceedings of Advanced Database System Symposium'94, Information Proceedings of the Symposium on Processing Society of Japan, Vol.94, No.13, 111-116, 1994.
- [2] Personal Correspondence, gA Global Change Data and Information System (GCDIS), Assited Search for Knowledge (ASKh, Committee on Earth Observation Satellite System, Working Group on Data, Catalog System and network System Subgroup Joint Meeting, 1994.
- [3] Ulrich Geske, Optimization and Simulation of Complex Industrial Systems, Proceedings of the 11th International Conference on Application of Prolog, INAP'98, 1998.
- [4] Masao Fukushima, Introduction to Mathematical Planning, Asakura Shoten, 1996.
- [5] Kohei Arai, Manabu Arakawa, Hirofumi Eto, Vague Search of Earth Observation Image Database Based on Fuzzy Theory Using Physical Quantities and Spatial Features, Journal of the Japan Society of Photogrammetry, Vol. 38, No. 4, pp. 17-25, Aug. 1999.
- [6] Kohei Arai, Hirofumi Eto, Tomoko Nishiyama, Earth Observation Satellite Image Database System Allowing Ambiguous Search Requests, Journal of the Japan Society of Photogrammetry, Vol.38, No.4, pp.47-52, Aug.1999 .
- [7] .Shin-ichi Sobue, Kohei Arai, Fumiyoshi Yoshida, User Friendly and Efficient Catalog Information Management for Earth Observation Data, Information Processing Society of Japan, Advanced Database Workshop, pp. 111-116, December 1994, 1994.
- [8] Kohei Arai, Remote sensing satellite image database system allowing image portion retrievals utilizing principal component which consists spectral and spatial features extracted from imagery data, International Journal of Advanced Research in Artificial Intelligence, 2, 5, 32038, 2013.
- [9] Kohei Arai, Data collection and active database for tsunami warning system, Proceedings of the 1st International Workshop on Knowledge Cluster Systems, 2007.
- [10] Kohei Arai, Remote sensing satellite image database system allowing image portion retrievals utilizing principal component which consists

spectral and spatial features extracted from imagery data, International Journal of Advanced Research in Artificial Intelligence, 2, 5, 32038, 2013.

- [11] Kohei Arai, Numerical representation of web sites of remote sensing satellite data providers and its application to knowledge based information retrievals with natural language, International Journal of Advanced Research in Artificial Intelligence, 2, 10, 26-31, 2013.
- [12] Kohei Arai, Indra Nugraha Abudullar, Hiroschi Okumura, Image retrieval based on color, shape and texture for ornamental leaf with medicinal functionality, International journal of Image, Graphics and Signal Processing, Vol.6, No.7, June 2014.
- [13] Cahya Rahmad, Kohei Arai, Comparison contour extraction based on layered structure and Fourier descriptor on image retrieval, International Journal of Advanced Computer Science and Applications, 6, 12, 71-74, 2015.
- [14] Kohei Arai, Cahya Rahmad, Image Retrieval Method Utilizing Texture Information Derived from Discrete Wavelet Transformation Together with Color Information, International Journal of Advanced Research on Artificial Intelligence, 5, 10, 1-6, 2016.
- [15] S.Sobue and Kohei Arai, Metadata Definition and Retrieval of Earth Observation Satellite Data, Proceedings of the IEEE Metadata Conference, (1997).
- [16] Kohei Arai, Geographic information system: GIS based on neural network for appropriate parameter estimation of geophysical retrieval equations with satellite remote sensing data, Proceedings of the IEEE Geoscience and Remote Sensing, PID 220128, 2006.
- [17] Kohei Arai, Yuji Yamada, Image retrieval method based on hue information and wavelet description based shape information as well as texture information of the objects extracted with dyadic wavelet transformation, Proceedings of the 11th Asian Symposium on Visualization, ASV-11-08-10, 1-8, 2011.
- [18] Kohei Arai, C.Rahmad, Wavelet based image retrievals, Proceedings of the 260th conference in Saga of Image and Electronics Engineering Society of Japan, 243-247, 2012.
- [19] Kohei Arai, Image Retrieval Method Based on Back-Projection, Proceedings of the Computer Vision Conference 2019.
- [20] ANCI / NISO, Z39.50, Information Retrieval (Z39.50): Application Searvice Difinition and Protocol Specification, URL: <http://lcweb.loc.gov/z39.50/agency/>, 1995.
- [21] Terano, J., Asai, K., Sugano, M., Introduction to applied fuzzy systems, Ohmsha, 1989.
- [22] Naoya Miya, Kazuo Tanaka, Practical Fuzzy Book in C Language, Russell, 1989.
- [23] Kaoru Hirota, Introduction to Fuzzy Expert System, Ohmsha, 1993.
- [24] Masakazu Saka, Fundamentals and Applications of Fuzzy Theory, Morikita Publishing, 1989.

AUTHOR'S PROFILE

Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Science Commission "A" of ICSU/COSPAR since 2008 then he is now award committee member of ICSU/COSPAR. He wrote 37 books and published 570 journal papers. He received 30 of awards including ICSU/COSPAR Vikram Sarabhai Medal in 2016, and Science award of Ministry of Mister of Education of Japan in 2015. He is now Editor-in-Chief of IJACSA and IJISA. <http://teagis.ip.is.saga-u.ac.jp/index.html>.