The Seventeenth International Symposium on Artificial Life and Robotics 2012 (AROB 17th '12),
B-Con Plaza, Beppu, Oita, Japan, January 19-21, 2012

Development of Stream Data Platform in Satellite Image Data Analysis System

Eiji Nunohiro[1], Hayao Mori[2], Masaki Hanada[1], Kenneth J. Mackin[1] ,and Jong Geol Park[1]

[1] Tokyo University of Information Sciences, 4-1 Onaridai, Wakaba-ku, Chiba 265-8501, Japan
[2] Graduate School of Tokyo University of Information Sciences, 4-1 Onaridai, Wakaba-ku, Chiba 265-8501, Japan
(Tel: 81-43-236-1292; Fax : 81-43-236-1292)

[1] nunohiro@rsch.tuis.ac.jp

Abstract: Tokyo University of Information Sciences receives MODIS (Moderate Resolution Imaging Spectroradiometer) data, one of the sensors equipped by NASA's Terra and Aqua satellites, and researches of the analysis on change of environment as part of the academic frontier project. For the information infrastructure of this frontier research, we are currently developing a satellite image data analysis system (SIDAS) implemented as a web system, with parallel distributed system configuration using multiple PC clusters, database for MODIS data, and applications to analyze the environmental changes. SIDAS is used to open the research results and MODIS data for public use. In this research, we are developing SIDAS 1.2 using the previously developed SIDAS 1.1 as the basic system. The main feature of SIDAS1.2 is the function of stream data processing, used to analyze the influence of the East Japan Earthquake and the monitor the regrowth of vegetation for the tsunami-stricken area in Tohoku region.This paper presents the overview of system configuration, the stream data processing and analysis functions of SIDAS 1.2.

Keywords: Satellite image data, MODIS, Stream Data Platform.

1 Introduction

Tokyo University of Information Sciences (TUIS) receives MODIS (Moderate Resolution Imaging Spectroradiometer) data, one of the sensors on NASA's Terra and Aqua satellites, and provides the processed data to universities and research institutes as part of the academic frontier project. The main research theme of this project is on "the sustainable development of economic and social structure dependent on the environment in eastern Asia". One of the major fields of research using MODIS data is the analysis of change in the environment. In this research, we are developing satellite image data analysis system (SIDAS) to support the frontier research such as weather change and environmental changes. In order to improve the accuracy of the analysis, we have developed SIDAS 1.1 which is implemented as a web system, with a parallel distributed system configuration using multiple PC clusters and a database system. SIDAS 1.1 has already supported several applications to analyze the satellite data such as "search system for fire regions in forests and fields" and "search system for similar satellite image data".

The purpose of our present study is to support the recovery from the East Japan Earthquake. In the research, we examined methods to observe the reconstruction of northeastern Japan after the East Japan Earthquake of March 11, 2011 by using the MODIS data. From these results, we are developing SIDAS 1.2 in order to provide the change of vegetation in time series of the north eastern regions, using the previously developed SIDAS 1.1 as the basic system. The main feature of the updated system is the addition of time series data processing to allow the analysis of the influence of the East Japan Earthquake and the revival of vegetation in the tsunami stricken area in north eastern regions. This paper presents the overview of system configuration, the stream data processing and analysis features of SIDAS V1.2 enhancements for the observation of the reconstruction from the East Japan Earthquake in the north eastern regions.

2 Overview of satellite data analysis system (SIDAS)

In this section, we describe the satellite data analysis system (SIDAS) to support the research environment and open the research results for public use. Figure 1 shows the overview of SIDAS 1.1. SIDAS 1.1 consists of a web server, an application server consisting of multiple PC clusters, and a database server. The application server schedules the execution of applications, manages the satellite data, and controls PC clusters which execute applications. The user can access the system, request data analysis and check the results over the network via a Web browser or other web applications, and the user can also download satellite image data through the system.

The processing flow of SIDAS1.1 is shown as follows.

(1) The Web server accepts request of the user and submits processing of the request to the application server.
(2) The application server transmits the MODIS data used by the processing from the MODIS database to the local disk in the application server.
(3) The application server executes application for the request and returns the results to Web server.

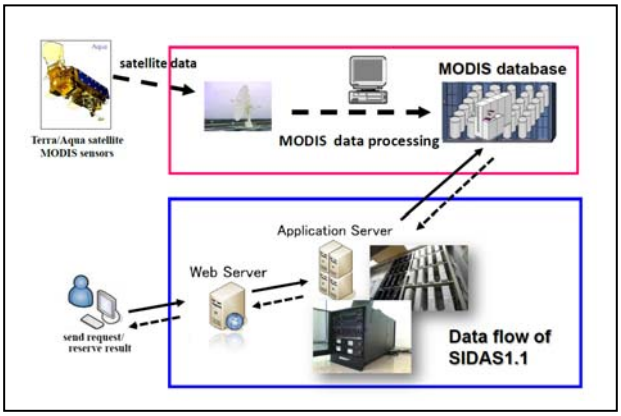

Fig. 1 Overview of SIDAS1.1

3 Overview of Satellite data analysis system 1.2

The purpose of our present study is to support the recovery from East Japan Earthquake. SIDAS 1.2 is being developed for this purpose, and the main feature of this updated system is the stream data processing, to enable the analysis of the influence of East Japan Earthquake and to monitor the regrowth of vegetation for the tsunami stricken area in Tohoku region. The analyzed results are automatically produced for open public access.

The information that we have planned to provide are the next 4 items to identify the reconstruction of land in the tsunami stricken area in the north eastern region from March 11, 2011.

(a) Image information from March, 2010 to the end of February, 2011 (Twice a month: In the first half and the latter half of each month).
(b) Image information after March 1, 2011 (Every day).
(c) Visible information of Normalized Difference Vegetation Index (NDVI), Sea Surface Temperature (SST) and Land Surface Temperature (LST) after March 1, 2011 (Every day).
(d) Visible information of difference between NDVI after March 1, 2011 and NDVI on the same day for one year ago (Every day).

Figure 2 shows the overview of SIDAS 1.2. SIDAS 1.2 executes the following processes every day.

(1) The application server receives the MODIS data stored in the MODIS database, and analyzes the information described in the above (a) ~ (d).
(2) The application server stores the analytical result in the database server.
(3) The Web server retrieves the analysis results from the database server and provides the data for open public access.

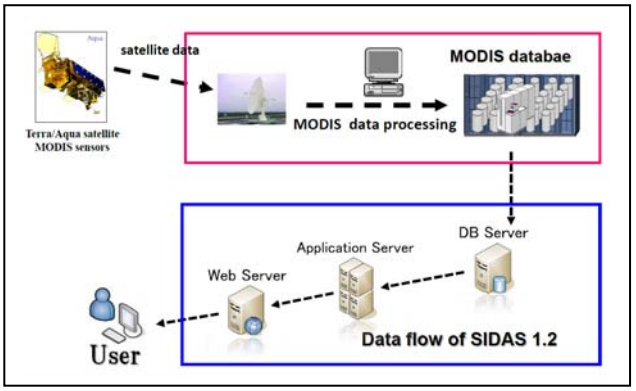

Fig. 2 Overview of SIDAS1.2

4 Summary of MODIS data in SIDAS

MODIS is a multiple band radiometric sensor instrument aboard Terra (NASA launched on December 18, 1999, at local time 10:30AM) and Aqua (NASA launched on May 4, 2002, at local time 1:30PM) satellites to monitor the environment of a large region of the earth covering a width of 2,330 km. It measures 36 spectral bands between 0.405 and 14.385 μm with three spatial resolutions (250m, 500m, 1,000m).

Tokyo University of Information Sciences (TUIS) is receiving MODIS data from three places in Hokkaido, Tokyo and Okinawa, and received the data 3 or 4 times in the daytime and 3 or 4 times in the nighttime from each place. The observing areas from the three places are from 100 degrees to 180 degrees of east longitude and from 10 degrees to 70 degrees of north latitude.

In this research, we use the following standard MODIS data products, MOD02 (MODIS calibrated radiances product: spatial resolution), MOD03 (geological location product), MOD11 (land surface temperature), MOD13 (vegetation index), MOD28 (Sea Surface Temperature), MOD35 (cloud mask) from MODIS products. Table 1 shows the band information of MODIS data.

Table 1 Relationship between the feature of the wavelength and the resolution of the bands

BAND	WAVELENGTH(nm)	SPATIAL RESOLUTIONS	KEY USE
1	620 ~ 670	250m	Absolute Land Cover change, Vegetation Chlorophyll
2	841 ~ 876	250m	Cloud Amount, Vegetation Land Cover change
3	459 ~ 479	500m	Soil/ Vegetation Differences
4	545 ~ 565	500m	Green Vegetation
5	1230 ~ 1250	500m	Leaf/Canopy Differences
6	1628 ~ 1652	500m	Snow/Cloud Differences
7	2105 ~ 2155	500m	Cloud Properties, Land Properties
21	3929 ~ 3989	1000m	Forest Fires & Volcanoes
22	3929 ~ 3989	1000m	Cloud Temperature, Surface Temperature
31	10780 ~ 11280	1000m	Cloud Temperature, Forest Fires & Volcanoes, Surface Temperature
32	11770 ~ 12270	1000m	Cloud High, Forest Fires & Volcanoes, Surface Temperature

5 Process of function in Stream Data Platform

In this research, we examined the observation method for the reconstruction of land in the Tohoku region from the East Japan Earthquake that had occurred on March 11, 2011 by using the MODIS data. The developed analysis functions of SIDAS1.1 is used to open the satellite data described in Chapter 3 for public retrieval. SIDAS1.1 that is base system of this research receives MODIS data, and this system executes application according the request of user such as "Search for fire regions in forests and fields" or "Search for similar image data" by using the MODIS data. Fig. 3 shows the configuration of SIDAS1.1.

In this research, we are examining stream data processing function that analyze reconstruction of the tsunami stricken areas in the Tohoku region, and are developing SIDAS1.2 to open the analytical result in real time by using the satellite data received every day.

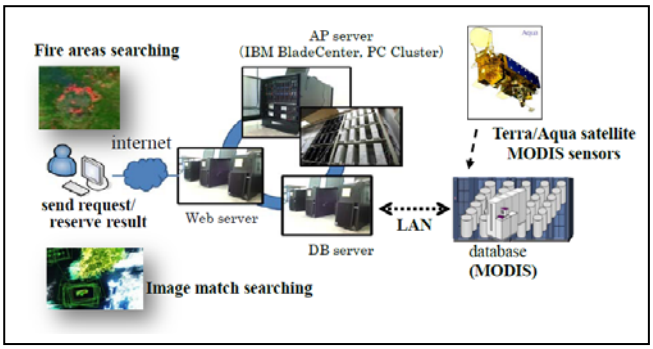

Fig. 3 Configuration of SIDAS1.1

Fig. 4 shows the configuration of SIDAS1.2. We used the stream data processing model for the data processing in SIDAS1.2. In this model, data processing scenarios are registered in the system beforehand, and the system receives and processes huge satellite data in real-time according to the predefined process scenario.

In the stream data processing function of SIDAS1.2, the analysis and web publishing function for the following item were implemented.

(a) Analysis from March in 2010 to the end of February in 2011

(b) Analysis after March 1, 2011

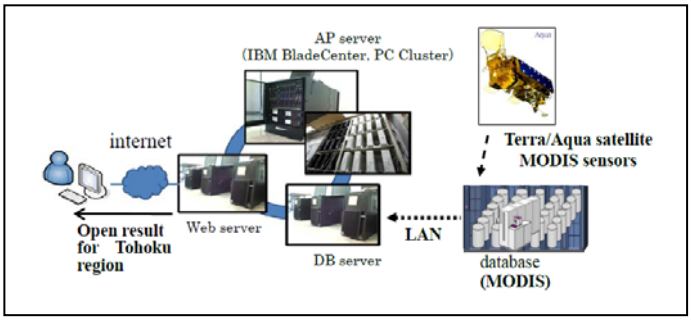

Fig. 4 Configuration of SIDAS1.2

MODIS data is analyzed beforehand and stored into the database for (a). We developed the data stream function to execute the reception, analysis and web data release of the satellite data in real time for (b). The process in the application server and Web server for (a) and (b) is shown as follows.

(a) Create analytical information from March, 2010 to the end of February, 2011

(a1) Application server

The following image data are created by using the MODIS data from March in 2010 to the end of February in 2011, and the image data are stored into the database.

(a1.1) RGB image data (GIF image : true color image)

The color image data in the Tohoku region is created by using radiance data (Band 1: Red and band 4: Green and band 3: Blue) of MOD02 (spatial resolution: 500m). The created data is GIF image data for 1- day composite data (received every day), 15- day composite data (removed cloud effects of 15-days) and Monthly composite data (removed cloud effects of 30-days). Fig. 5 shows example of display for RGB image data.

(a1.2) Visualized data of NDVI,SST and LST (GIF image : pseudo color image)

NDVI, SST and LST are single band data and thus the image data is normally shown as a grayscale image. In order allow make the data more intuitive to non-experts, the

vegetation indexes and temperature values were mapped to color values to create pseudo color images.

(a2) Web server

The Web server publishes data (a), (b) from March, 2010 to February, 2011.

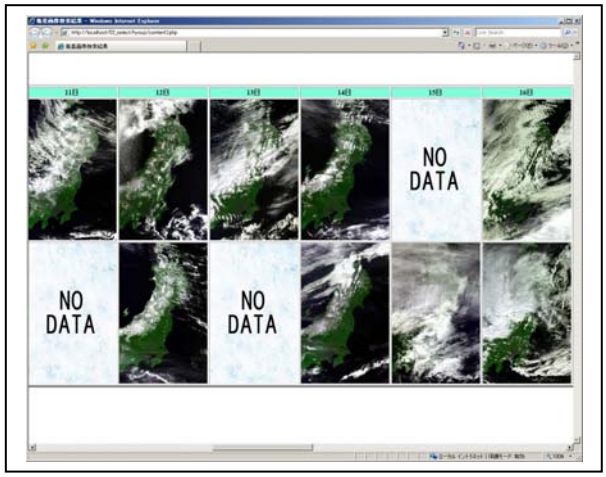

Fig. 5 Example of display for RGB image data

(b) Publish of analytical information after March 1, 2011 (The following process is executed once a day).

(b1) Application server

(b1.1) The application server retrieves the satellite data containing the Tohoku region for a single day. The number of reception from MODIS satellites is three or four times a day.

(b1.2) From the data received in (b1.1), the Tohoku region (north latitude 34°00' - 42°00' and longitude 138°00' - 43°00') is selected, and converts the HDF (Hierarchical Data Format) data to BSQ(binary Band SeQuential) data.

(b1.3) 1-day composite process is executed for the BSQ data created in (b1.2). 1-day composite processing combines the BSQ data with the best reception status (low cloud interface) from among the several MODIS data received that day and synthesizes them to one BSQ data for the selected day.

(b1.4) Create RGB image data from BSQ data (MOD02) and visualized data of NDVI, SST and LST from BSQ data (MOD11, MOS13, MOD28) created in (b1.3), and stores them into the database.

(b1.5) Execute the 15-days composite process once in the first half and once in the latter half of month, and the execute 30-days composite process once a month for BSQ data from (b1.3).

(b1.6) Create RGB image data from BSQ data (MOD02) from (b1.5).

(b1.7) Calculate the difference between the BSQ data (for LST, NDVI, and SST) of the selected observation date and the BSQ data (for LST, NDVI, and SST) created in (a) for exactly one year previous to the selected observation date, and create visualized data from the calculation results and store them into database.

(b2) Web server

Take out image data and visualization data that created in (1.4), (1.6), (1.7) from database and publish these data on the web server.

6 Conclusion and future works

In this research, we examined what information could be used to support the recovery from the East Japan Earthquake, and developed satellite data analysis system (SIDAS1.2) for this purpose. We developed SIDAS 1.2 to implement stream data processing in order to automatically publish up-to-date information on the influence of the East Japan Earthquake and to monitor the regrowth of vegetation in the tsunami stricken areas.

In future works, we intend to support the functions so that the user can acquire necessary information on demand, such as improved features to select analytical region and analytical information arbitrarily. Moreover, we plan to publish visual information incorporating other information sources such as aerial photographs, to examine the situation of the disaster and the reconstruction situation.

Acknowledgments

This research was supported by the Research project of Tokyo University of Information Sciences for the sustainable development of economic and social structure dependent on the environment in eastern Asia.

References

[1] Akihiro Nakamura, Jong Geol Park, Kenneth J. Mackin, Eiji Nunohiro, et al, "Development and Evaluation of Satellite Image Data Analysis Infrastructure", The Sixteenth International Symposium on Artificial Life and Robotics, Proceeding Index OS18-3, (AROB 16th' 2011)

[2] Eiji Nunohiro, Jong Geol Park, Kenneth J. Mackin, et al, "Development of satellite data analysis system", Proceedings of the 5th International Conference on Information (2009), pp.143-146

[3] Eiji Nunohiro, Kei Katayama, Kenneth J. Mackin, Jong Geol Park, "Forest and Field Fire Search System using MODIS Data", Journal of Advanced Computational Intelligence and Intelligent Informatics (JAC 3, 2007), vol.11 No.8, pp.1043 ~ 1048

[4] National Aeronautics and Space Administration (NASA), MODIS Web, http://modis.gsfc.nasa.gov/