International Journal of Multidisciplinary Research and Growth Evaluation.

Image mining using smart multi-agent system

Miliind Deshkar [1], Dr. Manoj Kumar Choubey [2], Kirti Agrawal [3]
[1-3] Department of CEA, GLA University, Mathura, Uttar Pradesh, India

Corresponding Author: **Miliind Deshkar**

Article Info

ISSN (online): 2582-7138
Volume: 03
Issue: 01
January-February 2022
Received: 15-12-2021
Accepted: 01-01-2022
Page No: 26-31

Abstract
In the current scenario, the fast growth and demands for remote sensing databases combined with human limits to analyze and extract knowledge from huge datasets lead to a need to investigate tools, techniques, methodologies, and theories capable of assisting humans. Image mining arises as a solution to extract implicit knowledge intelligently in huge image databases. Due to the growth in the volume of spatial information which is produced many times a day, demands other means for knowledge extraction. Spatial databases are among the ones with the fastest growth for investigation. Multiagent systems are composed of multiple computing elements known as agents that interact to pursue their goals. Agents have been used to explore information in distributed, open, large, and heterogeneous platforms. To solve relevant issues more precisely, accurately, and fastly, Agent mining studies ways of interaction and integration between data mining and agents. This area brought advances to the technologies involved such as theories, methodologies, and solutions. AgentGeo is one such technology enhancement consisting of relevant functions to extract knowledge from spatial databases. Satellite image mining promotes advances in the state of art of agent mining.

Keywords: remote sensing, database, image mining, multiagent system, agent, agent mining, data mining, AgentGeo.

1. Introduction

Now-a-days, with the help of advanced technological methods, there are numerous ways to collect spatial data: satellites, radars, unmanned air vehicles, balloons, and many others. The volume of spatial information produced all day long accumulates a large amount of images data on remote sensing databases. These remote sensing images when analyzed systematically, provide a unique opportunity for understanding how, when, and where changes take place in our world. Precious information exploited from spatial repositories has been promoting benefits in many areas, such as agricultural [1, 2] (forecast of harvests and soil erosion), hydric [3] (use of water resources and verification of the water quality), urban [4] (urban planning and demographic inferences), forest [5-7] (monitoring deforestation and biomass control), limnology [8] (characterization of aquatic vegetation and identification of water types), meteorology [9] (weather and climate studies), air traffic [10] (information for safety in the air), and national security.

However, for human experts, the manual analysis of huge databases is an extremely inconvenient task. The semi-automatic and intelligent interpretation of these data can become a useful and handy tool to control the earth's surface monitoring process. To detect these patterns semi-automatically and intelligently in huge databases, Data mining (DM) provides a solution. DM is used to find potentially useful, novel, and valid patterns in data [12, 13]. Traditional data mining can be extended for mining from structured data to unstructured data with the help of Image mining (IM) [14].

IM deals with the extraction of implicit knowledge, image data relationships, or other patterns not explicitly stored in huge image databases such as remote sensing and medical database. The research community of DM has dealt with some issues like the diversity of database models, efficiency and scalability, mining methodology, and user interaction [15]. Since the amount of currently available data is increasing day by day, the efficiency and scalability issue is very significant. Hence to improve some of these issues, it is necessary to make use of these new technological resources.

A collection of multiple computing elements is known as a multiagent system (MAS). These computing elements are known as agents that interact with each other to pursue their goals. They have decision-making software architecture systems that are embedded in an environment. In various significant issues of applications, this agent-based technology has been widely adopted.

The embedded technology of intelligent agents mining and Data mining aims to solve relevant problems that cannot be tackled by a single technology with the same quality and performance. Due to the involvement of both technologies, it promises to solve specific issues. In this chapter, we present an introduction to image mining, multiagent systems, and agent mining, as well as an overview of these areas.

In this paper, we present an introduction to image mining, multiagent systems, and agent mining, as well as an overview of these areas. AgentGeo, a multiagent system for satellite image mining uses the agent resources to mine image data in remote sensing databases and improves the analysis and application of satellite image mining as compared to other systems. Due to properties such as autonomy, interaction, reaction, and initiative of the agents controls the process of image mining. The agents are capable of handling the tasks of the image mining process, as well as improving the performance of the steps of preprocessing, transformation, feature extraction, classification, and evaluation.

2. Image Mining

Since the data from computing systems are produced constantly, it tremendously increases the uncontrollable growth in the institutions, industries, and corporations databases. This enormous development is caused by several factors, including internet versatility, reduction in the price of data storage devices, improvement of data collection tools, the popularity of embedded systems, increasing of online work. Also, data are being made available in various file formats like spreadsheets, texts, videos, and images.

The data volume is so complex that pertinent information becomes hidden within databases and hence becomes difficult or even impossible for a human being to detect patterns in expanded databases. Communication engineers, economists, statisticians, and forecasters worked with the indication that patterns from data can be identified, validated, and reached automatically for various applications [19]. Hence, to extract the knowledge from huge databases the knowledge discovery databases (KDD) originated.

Data mining (DM) is a particular step in the non-trivial KDD process of identifying unique and understandable patterns in data, where specific algorithms for extracting patterns are applied [12, 13]. However, the term data mining has become popular in the database field, used by statisticians and data analyst.

Data mining technology uses various algorithms, techniques, and tools to extract relevant information from a huge collection of databases. Data mining (DM) uses Image Mining (IM) technology that extends traditional data mining from structured data to unstructured data such as satellite images, medical images, and digital pictures. Extracted patterns from image databases are not easily interpreted and understood. Consequently, IM becomes an interdisciplinary endeavor that incorporates knowledge of important areas such as machine learning, image processing, computer vision, data mining, database, and artificial intelligence [20, 21]. The focus of image processing and computer vision is extracting specific features from a single image, whereas the IM makes efforts for extraction patterns stored in the huge image databases.

Figure 1 shows the image mining process. Everything starts from an image database in which is often inconsistent data that need to pass through a preprocessing step in order to improve the level of database quality. Image processing techniques are applied in this step, which are mathematical operations to change the pixel values of images, such as filtering, histogram equalization, image subtraction, image restoration, and others [22].

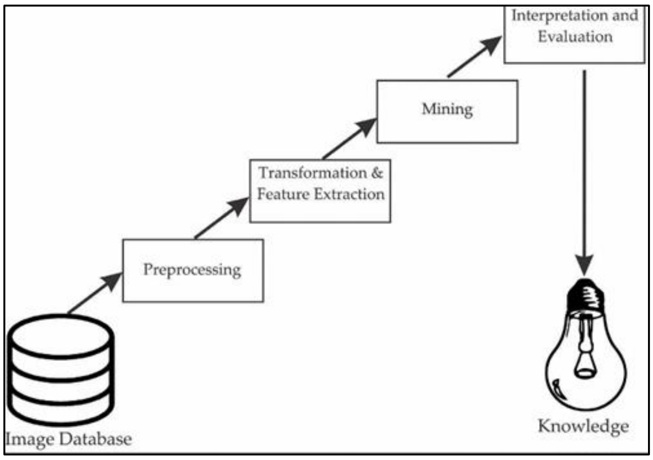

Fig 1: Image mining process

The image mining process works as follows

From the huge database available, some preprocessing is done to identify the data objects. The preprocessed data objects are then transformed by using relevant functions and methods till required objects are identified. Then the process of feature extraction like edge shape, texture, and length are used to extract the objects. Then using data mining techniques, the significant patterns are identified automatically or semi-automatically and evaluated by a specialist to obtain the knowledge, which can be applied to applications and can be useful in decision-making processes or on problem understanding.

Though the image mining process is analogous to the data mining process, but there are important differences between relational databases and image databases [20]. Some differences are as follows.

2.1 Spatial information: In image databases, several pixels constitute a simple image. As each pixel is related to its neighbors, the image miners have to extract position-independent features from images before mining useful patterns.

2.2 Domain dependency: Real-world objects are generalized and collected in terms of the specific features and elements. In relational databases, the data values are semantically arranged, whereas in image databases, the identification of elements, their classes, and relationships are linked to the context itself, and an image can have different information related to different domains.

2.3 Unique versus multiple interpretations: In relational databases, the data is identified and differentiated by the values. For example, the field person is Miliind; clearly states that the field stores a person's name. But in image databases, image data may reproduce ambiguous interpretations for the same visual pattern. For example, simple intensity data can be seen with different colors patterns.

Despite the fact that Image mining is a vast field and under development, there are many techniques that are frequently used for object recognition, like, image indexing and retrieval, image classification, and clustering [23-26]. Image mining has applications in many areas such as spatial research, aerospace, meteorology, health, agriculture, air traffic, among others.

3. MAS-Multiagent Systems

An agent is a computer program that acts for a user or other program in a relationship of agency, and it can be viewed as perceiving its environment through sensors and acting upon that environment through actuators [36]. The information received by agents at a certain moment is referred to as Perception. The perceptions can occur through the physical world, via a graphical interface, a collection of agents, the internet, or perhaps all combined. Agents have properties such as autonomy, social ability, reactivity, and proactiveness [37].

3.1.1 Autonomy is the ability to analyze the environment and make their own decisions without the intervention of humans or other agents, controlling their acts and internal state.

3.1.2 Social ability is the agent's ability to communicate with other agents using communication language.

3.1.3 Reactivity is the agent's ability to respond in a timely manner.

3.1.4 Proactiveness (or initiative) refers to the agent's capacity of taking initiative in order to achieve the needed goals.

These properties make agents of a technology capable of the following: cooperate in solving problems; share expertise; work in parallel on common problems; develop and implement modularly; be fault-tolerant through redundancy; represent multiple viewpoints and the knowledge of multiple experts, and mainly be reusable.

There are four kinds of agents: goal-based agents, utility-based agents, simple reflex agents and model-based reflex agents, and [36].

3.2.1 Goal-based agents: These agents know some sort of information about their goals. Based on this information and with the internal state, these agents analyze and take their actions.

3.2.2 Utility-based agents: These agents know some sort of information about their goals and have an internal state.

3.2.3 Simple reflex agents: Based on current perception, these agents select their actions, ignoring the perceptual history.

3.2.4 Model-based reflex agents: Maintains an internal state based on the perception history and reflects the unobserved aspects of the current state.

Agents can adapt and anticipate the needs of different users by learning new concepts and techniques. During the last years, agents have become a powerful technology, which has been adopted in several applications as a solution to solve complex issues that cannot be solved by humans. In the area of remote sensing, for example, a complex task is to analyze remote sensing images, a human being is able to analyze a single remote sensing image, but analyzing a significant amount of this data is unlikely because of limited human ability to reason and interpret huge information volumes.

However, when a simple agent cannot solve the problem, a MAS can be implemented. These systems have been studied since the 1980s but were only recognized in the mid of the 1990s. At that time, scientific and industrial interests were raised due to the need of exploiting information and modern computing platforms, as well as distributed, open, large, and heterogeneous ones. If two or more agents interact with each other to solve a problem, a MAS is formed. These agents act on behalf of users with different goals. However, the purpose of the system determines whether agents having the same goals can interact or not. The interaction is based on the messages passed and determined on the basis of the ability to coordinate, cooperate, and negotiate between agents [38].

Coordination is a property that aims to perform some activity in a shared environment. Avoiding livelock and deadlock, maintaining applicable safety conditions, the extent to which they avoid extraneous activity by reducing resource contention determines the degree of coordination [39]. In general, we consider a relevant degree of coordination, when agent activities agents are well-balanced inside of the environment as well as the operations are being distributed and involved among agents without any failure. There are some reasons why multiple agents need to be coordinated [40], which are as follows:

- Their goals may be conflicting.
- Their goals may be interdependent.
- Agents may have different capabilities and different knowledge.
- Their goals can quickly achieve if different agents work together in a coordinated way.

Cooperation is coordination among non-antagonistic agents, while the negotiation is coordination among competitive or simply self-interested agents [39]. Agents can cooperate with each other to the general goals of the system, or they can compete for their individual goals, at MAS. Both features must be determined according to the general purpose of the agents in the application.

In the MAS, a key issue is how the agents will communicate. There are two types of communication in multiagent systems [42]. Direct communication: agents are able to communicate with each other using direct message exchange mechanisms between them.

Indirect communication: consists of indirect transfer of information between agents. For example, when an agent wants to send a message to another agent, it relies on the mediating agent who is responsible for the exchange of information within the environment.

MAS have become more and more important in many aspects of computer science such as distributed artificial intelligence, distributed computing systems, robotics, and artificial life. MAS can be implemented because of following reasons:

- When the system is complex and the human being is unable to predict the behavior of that system.
- When it is expensive to keep a team of specialists working.
- When the activity involved put humans at risk.
- When the decision-making process requires performance, agents can solve the problem quickly using parallel

processing.
- When it is necessary to ensure information privacy.

4. Multi-agent system for image mining

To solve a very complex problem in distributed environments, Agents technology is powerful amongst all. Here agents can cooperate, coordinate, and communicate their activities in order to reduce the complexity of the problem. Agent research focuses on theoretical, methodological, technical, experimental, and practical issues and the means to handle system complexities [44]. Data mining is an application-oriented technology that uses standard algorithms and techniques, for extracting relevant information (or patterns) intelligently from huge and varied datasets collection.

The application areas of Agent technologies are user interfaces, e-commerce, information retrieval, robotics, computer games, education, and training, whereas the application areas of Data mining are in government services, fraud detections, securities, bioinformatics, web, text, medical data and meteorological data mining.

Agents and data mining have their specific problems and limitations. Both areas can contribute to tackling the critical challenges of other technology. Agents are an excellent option for distributed computing and peer-to-peer, parallel computing and hence can force the KDD process on data selection, extraction, preprocessing, and integration. They can be an interface between software systems and humans and hence can affect multisource mining or human-mining needs [44]. In the same way, the knowledge acquired through data mining processes can be used for learning on multiagent systems that provide more stable, predictable, and controllable models for dispatching and planning.

Therefore, agents are elements that can influence the data mining process, and data mining can contribute significantly to the agent's area. This interaction and integration of Agent and Data mining technology gave a new research field known as Agent Mining [44-49]. Agent mining refers to the methodologies, principles, techniques, and applications for the integration and interaction of agents and data mining, as well as the community that focuses on the study of the complementarity between these two technologies, for better addressing issues that cannot be tackled by a single technology with the same quality and performance [44, 49]. The Agent-mining area is under development; therefore, some issues demand research on theoretical, technological, and methodological aspects. This gave rise to two fields of research.

Agent-driven distributed data mining or multiagent data mining: It involves the ways to use agents to enhance data mining processes and systems. Agents can be used in data mining for different purposes such as agent-based data mining systems, agent-based data warehouses, agents for information retrieval, mobile agents for distributed data mining, among others [49].

Data mining driven agents: It deals with the issues related to collaborative learning, reasoning, communication between two agents, adaptation and evolution, planning and dispatching, data mining agent intelligence enhancement [52].

Agents can enhance the image mining process in some steps:

Preprocessing: If Agents perceive anomalies on the stored images, they can automatically use digital image processing techniques to deal with them. This can reduce human efforts.

Transformation and feature extraction: Multiple agents can be assigned the tasks of transformation and feature extraction, which will increase the speed of data transformation, detecting objects, and the segmentation process.

Mining: Agents can discover significant patterns automatically or semi-automatically in image databases. This will increase the performance and accuracy of the classification process.

Interpretation and evaluation: Agents can learn with specialists about a particular domain, that is, they can perform this task to support or even substitute the specialists.

5. Geo-agent

GeoAgent emerged based on the ideas proposed by [16, 17], which uses agents to mine remote sensing images. This tool is being developed in Java language, and its initial version works along with GeoDMA and TerraView. This system brings advances in the state-of-the-art of multiagent systems and image mining, presenting relevant resources and precious functionalities. For example, this system implements functions such as creation, edition, and selection of agents, selection and creation of the environment, and use of agents for image mining. Moreover, it is capable of performing the classification process with multiple images at a time, differently from GeoDMA that is limited to only one image during the entire image mining process.

5.1 System Architecture

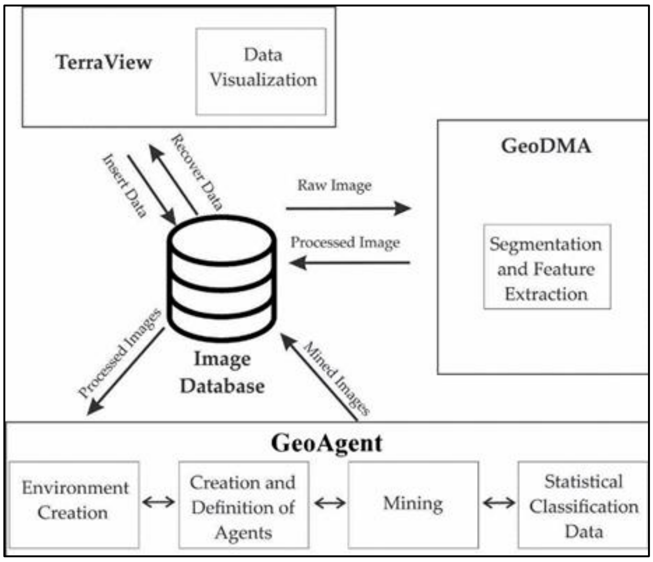

Fig 2: Image mining system architecture.

TerraView is the first step from where the image mining process starts. It is a Data Visualization step that provides the structure to insert and view several images in the database. Each individual image is then processed and segmented by GeoDMA, which extracts the characteristics of each one. The image database holds these processed images. GeoAgent is connected to the same image database and receives the processed data as input which creates the environment for the agents through GeoDMA. After that, the specific mining agents that will perform automatic mining on these images are defined; which gives the mined images as output. These images are in turn stored in the image database and can be visualized through TerraView. GeoAgent provides statistical data about the mining performed by agents, such as the number of segments classified by agents.

5.2 Agent's Structure

There are two agent types implemented at GeoAgent based on the features, goal and functions:

1. The simple agent: These agents select their actions based on current perception, ignoring the rest of the perceptual history. As soon as a user selects an agent on the GeoAgent, defining the environment, the agent's perception starts. When the agents are pursuing their particular goals in the mining process, they enjoy a level of autonomy. The agents decide the goals and hence the user is not able to forecast the agent actions. Besides that, every simple agent competes for its goals in the environment, and when their goals conflict with other agents, a monitor agent is responsible for coordinating environmental disputes.
2. A monitor agent: On conflicts between two agents, the monitor agent takes the initiative to resolve it. The monitor agent has the ability to communicate with other agents.

The communication in the environment occurs indirectly as shown in Figure 3. Simple agents can only store information about their goals, whereas the monitor agent has the summary of the environment and hence knows the other agents. This communication architecture solves the communication, coordination, and negotiation problems partially and reduces the complexity of the MAS considerably.

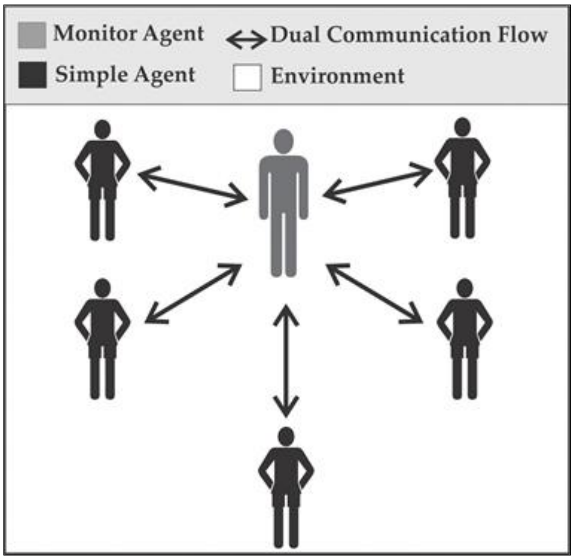

Fig 3: Communication architecture at GeoAgent.

The thread structure is used to implement agents and they can perceive and act simultaneously in a certain environment. The simple agents are composed of internal structure, with its goal information and metadata, and are built by the user. This development has four steps as illustrated in Figure 4.

Any agent needs basic information in the form of a knowledge base to reason about their acts. The training phase of GeoDMA u responsible to build this knowledge base, after which the segmentation and extraction of features are carried out. The decision tree structure is used to create an agent's knowledge base.

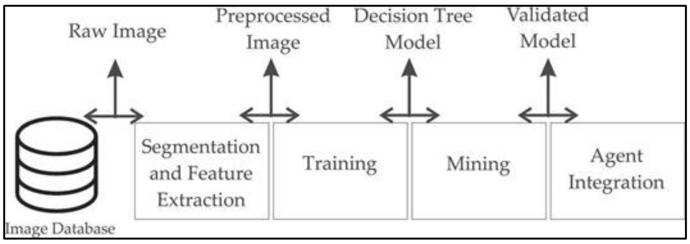

Fig 4: Steps for agent construction at GeoAgent.

Firstly, it is defined the image database, and the images are preprocessed on the segmentation and feature extraction step. At the training step, segment samples that are known to users referring to a specific class are selected by them, according to the agent goals referring to a specific class according to the agent goals. After the creation of specific classes, the users can generate a decision tree using GeoDMA, which uses the supervised algorithm C4.5, for spatial data mining, and hence a resource is built.

The mining step happens automatically, with the decision tree model generated by GeoDMA. By referring to the spatial and spectral attributes of the segments present in the image, the process is performed by means of thresholds. To get a consistent model, the steps are performed several times. If the user identifies that the results of the model are not satisfactory, he can backtrack to previous steps and re-perform it. If re-performing the steps doesn't solve the issue then the model is used as a knowledge base for the agent. The user can inform the agent about the name, metadata, and knowledge base.

6. Conclusion

In this work, we proposed an approach to image mining, multiagent system, agent mining, and presented GeoAgent. We have suggested improvements in the process of integration and interaction of agents, data and image mining. Also, GeoAgent was introduced, which exploit remote sensing databases using the image mining process. GeoAgent brings advances in state-of-the-art multiagent systems and image mining, by providing relevant resources that leverage the image mining process.

As future work, we hope to integrate more resources into GeoAgent to provide operational advantages, optimization, and innovation. We conclude that the agents can be used at various image processing domains for performing various steps in the image mining process. These agents can be used to build modules at GeoAgent for the steps of transformation, feature extraction, and evaluation. We can also combine various agents with different goals to work in a synchronous way for applications of agriculture mining, deforestation mining, water mining, oil extraction mining, cloud mining, or to mine any spatial object.

7. References

1. Aguiar DA, Rudorff BFT, Silva WF, Adami M, Mello MP. Remote sensing images in sup- port of environmental protocol: Monitoring the sugarcane harvest in Sao Paulo State, Brazil. Remote Sensing. 2011;3(12):2682-2703. DOI: 10.3390/rs3122682
2. Alba JMF, Schroder VF, Nóbrega MRR. Land cover change detection in Southern Brazil through orbital imagery classification methods. In: Escalante B, editor. Remote Sensing-Applications. 1st ed. In Tech; c2012. p. 99-116. DOI: 10.5772/36940. Available from: https://

www.intechopen.com/books/remote-sensing-applications/strategies-of-change-detection- in-southern-brazil-by-orbital-imagery-classification-methods
3. Ding Z, Qi N, Dong F, Jinhui L, Wei Y, Shenggui Y. Application of multispectral remote sensing technology in surface water body extraction. In: International Conference Audio, Language and Image Processing (ICALIP); c2016. p. 141-144. DOI: 10.1109/ICALIP.2016.7846565
4. Matthieu KWK, Maeyer M, Wolff E. The mapping of the urban growth of Kinshasa (DRC) through high resolution remote sensing between 1995 and 2005. In: Escalante B, editor. Remote Sensing-Applications. 1st ed. In Tech; c2012. p. 463-478. DOI: 10.5772/38435. Available from: https://www.intechopen.com/books/remote-sensing-applications/the-mapping-of-the-urban-growth-of-kinshasa
5. Sousa WRN, Couto MS, Castro AF, Silva MPS. Evaluation of desertification processes in Ouricuri-PE through trend estimates of times series. IEEE Latin America Transactions. 2013;11(1):602-606. DOI: 10.1109/TLA.2013.6502869
6. Pizaña JMG, Hernández JMN, Romero NC. Remote sensing-based biomass estimation. Marghany M, editor. Environmental Applications of Remote Sensing. In Technology; c2016. p. 1-40. DOI: 10.5772/61813. Available from: https://www.intechopen.com/books/environmental-applications-of-remote-sensing/remote-sensing-based-biomass-estimation.
7. Maciel AM, Silva MPS, Escada MIS. Mining frequent substructures from deforestation objects. In: IEEE International. Geoscience and Remote Sensing Symposium (IGARSS); c2012. p. 6745-6748. DOI: 10.1109/IGARSS.2012.6352557
8. Cho HJ, Mishra D, Wood J. Remote sensing of submerged aquatic vegetation. In: Escalante B, editor. Remote Sensing-Applications. 1st ed. In Tech; c2012. p. 297-308. DOI:10.5772/35156. Available from: https://www.intechopen.com/books/remote-sensing-applications/remote-sensing-of-submerged-aquatic-vegetation
9. Ashraf A, Rustam M, Khan SI, Adnan M, Naz R. Remote sensing of the glacial environ- ment influenced by climate change. In: Marghany M, editors. Environmental Applications of Remote Sensing. 1st ed. In Tech; c2016. p. 99-129. DOI:10.5772/62134. Available from: https://www.intechopen.com/books/environmental-applications-of-remote-sensing/ remote-sensing-of-the-glacial-environment-influenced-by-climate-change
10. Zhang L, Zhang Y. Airport detection and aircraft recognition based on two-layer saliency model in high spatial resolution remote-sensing images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2016;10(4):1511-1524. DOI: 10.1109/ JSTARS.2016.2620900
11. Ardouin JP, Lévesque J, Roy V, Chestein YV, Faust A. Demonstration of hyperspec- tral image exploitation for military applications. In: Escalante B., editor. Remote Sensing-Applications. 1st ed. In Tech; c2012. p. 493-517. DOI:10.5772/37681. Available from: https://www.intechopen.com/books/remote-sensing-applications/ demonstration-of-hyperspectral-image-exploitation-for-military-applications
12. Han J, Pei J, Kamber M. Data Mining: Concepts and Techniques. 3rd ed. Waltham, MA, USA: Elsevier; c2011. p. 703.
13. Moura CDG, De Freitas NCA, Silva MPS. Agentes de Mineração de Imagens. In: Proceeding of the XV Brazilian Symposium on GeoInformatics (GEOINFO), SP-Brazil; c2014. p. 1-6.
14. De Freitas NCA, Filho PPR, Moura CDG, Silva MPS. GeoAgent: Multi-agent system of satellite images mining. IEEE Latin America Transactions. 2016;14(3):1343-1351. DOI: 10.1109/TLA.2016.7459619.
15. Zhang J, Wynne H, Lee ML. Image mining: Issues, frameworks and techniques. In: Proceedings of the Second International Conference on Multimedia Data Mining. Springer- Verlag; c2001. p. 13-20.
16. Gonzales RC, Woods RE. Digital Image Processing. 3rd ed. Upper Saddle River, NJ, USA Pearson Education, Prentice Hall; c2008. p. 976.
17. Dey N, Karâa WBA, Chakraborty S, Banerjee S, Salem MMA, Azar AT. Image mining framework and techniques: A review. International Journal of Image Mining. 2015;1(1):45-64. DOI: http://dx.doi.org/10.1504/IJIM.2015.070028
18. Kaur P, Kaur K. Review of different existing image mining techniques. International Journal of Advanced Research in Computer Science and Software Engineering. 2014;4(6):518-524.
19. Shukla VS, Vala J. A Survey on image mining, its techniques and application. International Journal of Computer Applications. 2016;133(9):12-15. DOI: 10.5120/ijca2016907978
20. Silva MPS, Câmara G, Souza RCM, Valeriano DM, Escada MIS. Mining patterns of change in remote sensing image databases. In: Fifth IEEE International Conference on Data Mining; November 2005; Houston, Texas, USA. Houston, Texas, USA: IEEE; c2005. p. 8. DOI: 10.1109/ICDM.2005.98
21. Korting TS, Fonseca LMG, Câmara G. GeoDMA-Geographic data mining analyst. Computers & Geosciences. 2013;57:133-145. DOI: 10.1016/j.cageo.2013.02.007