UPDATING OF A GEOGRAPHIC DATABASE: AN APPLICATION AND DESIGN OF A GEOGRAPHIC INFORMATION SYSTEM

Hakan MARAS*, M.Orhan ALTAN**

*Head of Digital Geographic Data Production Section, General Command of Mapping, Turkey
hmaras@hgk.mil.tr
**Head of Photogrammetry Division of Geodesy and Photogrammetry Department, Civil Engineering Faculty, Istanbul Technical University, Turkey
oaltan@srv.ins.itu.edu.tr

KEY WORDS: Geographic Information Systems, Database, Design, Updating, Monitoring.

ABSTRACT

The creation of a GIS database can cost more than the GIS hardware and software itself. Although it takes several years to create, the information used in the databases of GIS represents the status of the world instantly. GIS databases must appropriately be organized and constantly be updated to make better use. In this study, methods were investigated with regard to updating of GIS databases to be utilized currently at various applications for a great number of disciplines. Different methods were determined for updating either spatial or non-spatial component of data in geographic databases. As a result, a user-friendly software environment was developed which is also able to use raster data as source, in order to update simultaneously spatial and non-spatial components of data in databases, to store time information with regard to updating, to provide the update of data model and to change its environment automatically according to data model.

1 INTRODUCTION

The effects caused by natural and artificial events result in changing of the earth continuously. Although changes resulting from natural events occur very slowly and they are almost rarely detectable, artificial changes brought by the increase in the world population and growth in the rural areas occur much rapidly. The Earth must be represented in a certain medium to keep track of the occurring changes aimed to make plans which help to take necessary precautions via monitoring them and making predictions for the future. Representation of spatial features classically on map sheets, which are used in the past, is not suitable for this purpose. The concept of Geographic Information System (GIS) is developed as a result of the studies initiated to solve this problem and rapid development in computer technology. Today, a great deal of disciplines and technology ie., computer science, information management, geodesy, photogrammetry, remote sensing, cartography, data communication etc. affects and gives shape to GIS technology (Altan et al., 1996).

On the other hand, GIS is superior to classical methods such that it has a larger capacity for storage and manipulation, and derives effective results on making decisions by analysis. In order to make the most use of these advantages, data included in GIS database must constantly be updated. Besides, the time information of update processes should also be recorded in the database. Because, information relevant to date of data capture and modification have a great impact on the reliability and usage of itself. Ignoring the updating process will obviously decrease the usefulness of GIS.

2 DESIGN AND UPDATING OF GEOGRAPHIC DATABASES

2.1 Geographic Database Design

GIS cannot generally be used without the prior design of a database and its creation. In the design of geographic databases, after user requirement determination, database structure is created conceptually, then to meet these requirements physical design is documented, and finally the database is implemented in the computer medium. To come up to future expectations expansion capabilities must be taken into account during the conceptual and physical database design stages. The creation phase almost always represents the most significant investment in the implementation of a GIS.

2.2 Updating Geographic Database

Although change on the Earth's surface may be slow and difficult to detect, changes will and do occur. This change may be global (climate warming) or local (salinization, erosion), it may be gradual and difficult to observe (loss of soil by wind or water) or dramatic and catastrophic (earthquakes, volcanism, floods or landslides) (Burrough, 1997). While other databases may seldom need changing, geographic databases must be changed regularly, perhaps as often as several times a day. The need for updating results from the changes of feature locations and/or of feature attributes. In addition, data dictionary, which consists of meta-data and describes the database, might also be updated when the expectations and circumstances are changed.

Most database management systems have functions allowing the user to keep the database up-to-date. Methodologies aiming to update data should not only update the spatial information according to the main principles, but also enable the GIS user to store the updating status of the database. Davis and Zuppo (1989) stated that the development of a system that is able to store historical records of updating processes is mandatory in order to develop any updating methodology. When designing a GIS, care should be given to using efficient and useful methodologies enabling to input new information to the system and to store the time information of changes.

Members of European Organization for Experimental Photogrammetric Research (OEEPE: *Organisation Européenne d'Etudes Photogrammétriques Expérimentales*) detected the following main problem areas on updating digital complex topographic databases (Gray, 1995):

- handling the time/historical dimension
- development of user friendly, cost effective procedures
- production of software to produce and accept change-only information
- data compatibility

2.2.1 Updating Spatial Data. Spatial data related with the spatial objects are stored and displayed in geographic databases with graphical features. Therefore, updating geographic database consists of adding new graphic feature, modifying, deleting or suppressing the old ones. Adding a new feature can be done by digitizing new points, lines and polygons on the source material, or by data conversion. Suppression or deletion of a feature is the process of removing a graphic feature from the current database with or without having a backup copy. All features have an independent record in the GIS software's attribute table, so deleting the feature results in removing the record from the table. Modification of a feature can be done by either changing its location or dimension. If the existing features in the database are not affected by the changes in feature location, only the coordinate data of relevant feature is changed. Changing the dimension of a feature is followed by the change in coordinate data as well as its other graphical attributes (i.e. perimeter, length, area, etc.). If the modification causes an effect on existing features, all topological attributes (i.e. starting-nodes, ending-nodes, polygons on the left and right, etc.) and graphical attributes must be corrected according to the new situation. All graphical data manipulation processes (i.e. addition, suppression and modification) can change the number of graphical elements because of splitting and joining. In this case, number of records in attribute tables must be altered to ensure the compatibility between spatial and non-spatial data components.

2.2.2 Updating Attribute Data. Attributes are alphanumerical data that are associated with graphical features and stored in separate database files called attribute tables. The updates required from a change in the value of an attribute can be done by modifying the graphical and topological attribute values. To make new decisions by means of new analysis, user might need to add new attributes or remove some of the existing ones. Since the attribute values can be initialized or changed, attributes, which are defined by the fields of database tables, can be created, deleted or their characteristics can be altered. Deleting an attribute value causes it to be changed with zero or nothing can be considered as modification. On the contrary, deleting an attribute means removing the column it belongs to. Therefore, all the graphic elements using the same attribute table will be subjected to this change.

Some meta-data that might be defined as attribute such as recording and updating date can be updated automatically by the use of GIS software facilities. Other meta-data like the source or accuracy of the data can be updated by the user during the data capture process interactively.

2.2.3 Updating the Data Dictionary. Adding, deleting and modifying records and fields are all common database operations. They are mostly easy to implement. But the data in the data dictionary are dependent not only on each other but also on the data in the geographic database. Hence updating of the data dictionary is not as easy as updating spatial data or attribute values. For instance, if someone wants to remove a layer of geographic data model, he has to delete all features within this layer, or move them to other layers. Therefore, the meta-data forming the data dictionary should be updated meticulously. Updating processes of a data dictionary are as follows:

- adding, deleting, merging and decomposing layers,
- adding and deleting features, modifying feature definitions,

- changing the feature contents of layers,
- adding and deleting attributes, modifying attribute definitions,
- determination or modification of attribute value domain,
- assigning attributes to features or modification of feature attributes,
- determination or modification of feature attribute value domain,

2.2.4 Monitoring the Database-Updating Process. The course of updating processes should be stored with their dates so that they can be more significant and queried. In monitoring the history of updates, the worst handicap takes place when a deletion is made. The deletion causes the data to be inaccessible at computer medium. Access to the deleted data or undoing of deletions can be done by recording backup files. The backup file that is created before deletion should include the information, i.e. date and reason for backup, the user ID recording backup, etc. Backup files enable the users to monitor additions, deletions and modifications in chronological order.

There are two major approaches to monitor database-updating processes (Maras, 1999):

- Creation of a class of geographic area objects which will receive all information regarding a specific updating task.
- Recording of historical data of an updating task of each feature individually in its attributes.

The latter can be called as feature-oriented method and handled by adding user attributes carrying time stamps. In this approach, to monitor updating processes dates related with addition or modification of the feature are stored in its attributes. The modification date of attribute values should also be assigned to the attributes of the feature associated with the attribute (Figure 1). Date tagging can be done by the user interactively, which, of course, is not ideal. To have better results, routines capable of automated date tagging should be developed. To automatically handle historic information of updating processes during data manipulation, user interfaces enabling both manipulation and recording should be developed by using GIS macro language.

		FOR GRAPHICAL UPDATING PROCESSES			FOR ATTRIBUTE VALUE UPDATING PROCESSES		
F_CODE	NAME	CREAT_D	MODIF_D	POPULATN	POP#ENTR_D	POP#MODF_D	POP#DEL_D
PP0201	SAKARYA	12.01.1995		732,414	02.10.1995		
PP0201	SAMSUN	30.04.1995	27.05.1998	1,168,574	02.10.1995	12.11.1998	
PP0201	ANKARA	05.08.1995		3,631,612	02.10.1995	12.11.1998	
PP0201	ISTANBUL	30.04.1995	27.05.1998	9,057,747	02.10.1995		
PP0201	IZMIR	24.07.1995			02.10.1995		24.09.1998

Figure 1. A sample attribute table for monitoring both the feature and the attribute updating

Deleting or changing of location or shape of a graphical feature results in loosing its previous information. Therefore, it is not possible to store information prior to updating in the current database files. Before significant deletion processes, data file should be saved as another data file which also includes date information. On the other hand, recording backups much frequently results in waste of disk space. However, recording backups seldom can cause loss of information or access to misleading information. So, the answer to the question of “Under which condition and to what extent a backup should be taken?” can be given according to the expectations and the usage of GIS. Since the backup time is not specific, the user should decide it meticulously. Another way to access deleted data is to create a file for storing the features to be deleted before deletion. By using this file, the user can query and access deleted features. A sample feature-oriented updating process is exposed by the flow chart on Figure 2.

In this approach, after modification or deletion user can not instantly access the information prior to these operations. Going through the backup files is necessary to access this information. Creating two additional attributes for each table and three additional attribute fields for each attribute which is being monitored, causes extra use of disk space. Although this seems to be a waste of disk space, they enable the user to keep the time dimension in the database and to query regarding time.

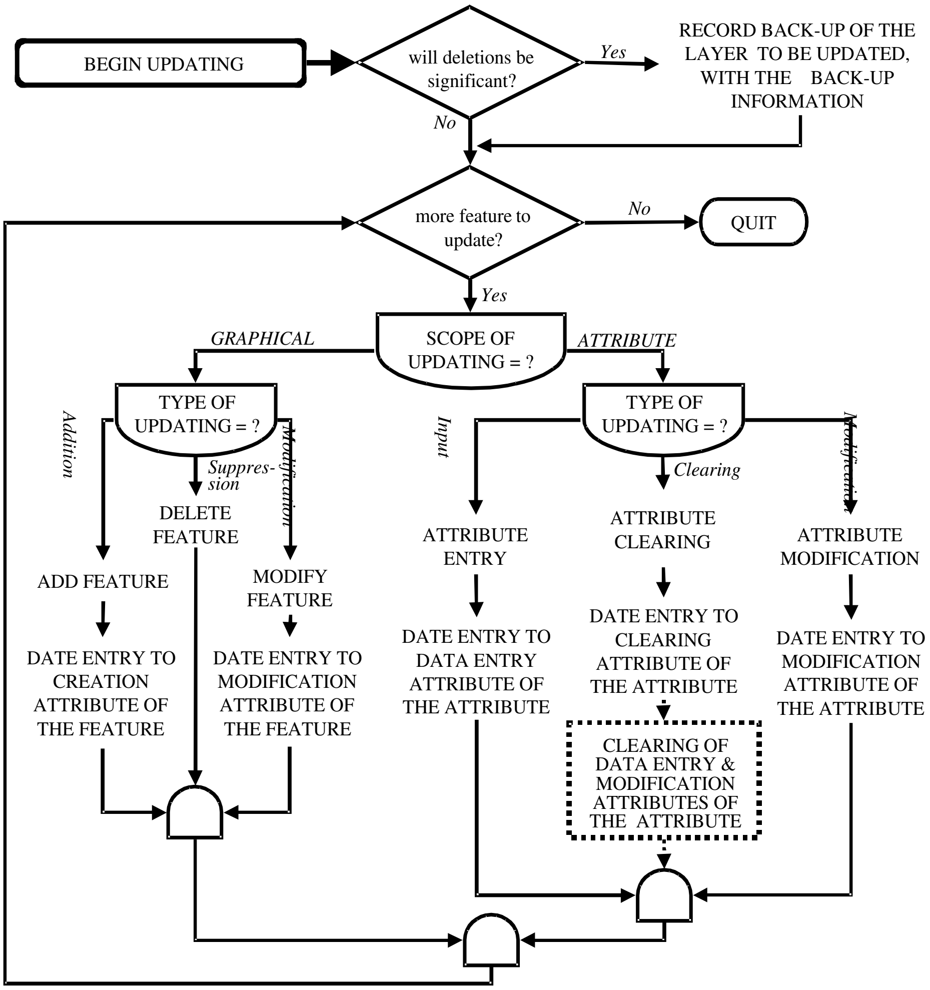

Figure 2. Feature-oriented updating process enabling to monitor updates

As a result; to monitor database-updating process, to undo modifications and deletions, and to keep time information of updates, the second approach, which also includes first one partially, is preferably to be used. If space efficiency is preferred to instant monitoring of database updates, the first approach is to be preferred.

3 DESIGN AND APPLICATION STUDY FOR DATABASE UPDATING

Studies to update databases which utilize various applications of a great number of disciplines has been carried out. Different methods were determined for updating either spatial or non-spatial component of data in geographic databases. Methods were examined for updating spatial data component via graphical data manipulation functions included in GIS software, and of non-spatial data component via database management system. It was observed that the built-in functions of the GIS softwares were away from updating and recording date information together automatically, and susceptible to user errors. Furthermore, in the applications held so far, it is observed that the data model (from which system design and development is derived) is fixed causing to prevent flexibility in modifying the data model which comes from the changing requirements of the users. Hence new systems that give this ability should be developed. The available methods and structures utilized for the representation of geographical features of the earth in computer medium were examined to develop more useful data storage methods and data structures when updating is carried out. Since the changes occur in units of geographical features world-wide, in order to store and update data in databases, the use of vector data structures representing the geographical features in detail was considered to be more useful. Satellite images, and raster data obtained from scanning of the classical paper maps or aerial photographs were considered to be the ideal data sources for update processes.

An appropriate database and application software programs were designed utilizing system development techniques to;

make simultaneous updates in spatial and non-spatial components of data in geographical databases
store time information relevant to updating
provide the update of data model
change its environment automatically according to data model.

Figure 3 shows the functional architecture of the system. The software required by the system design was developed via Arc Macro Programming Language of Arc/Info Geographic Information System software and Access Basic Programming Language of Access database management system software. Because data collection needs more time than other phases of system design, a new data set was obtained appropriate to use in the application by transforming another data set in a different model. In the application constructed in accordance with the system design, a user-friendly software environment was prepared, which enables to update database through addition, suppression and modification of spatial and non-spatial data. Recording of time information of updating processes automatically into database during data manipulation is optional in the system. In the application, spatial queries are also performed regarding the dates of updating process. Besides, data model update resulting from the user requirements can be carried out in an easier-to-use environment, leading to automatic update of the interface based on update of data model.

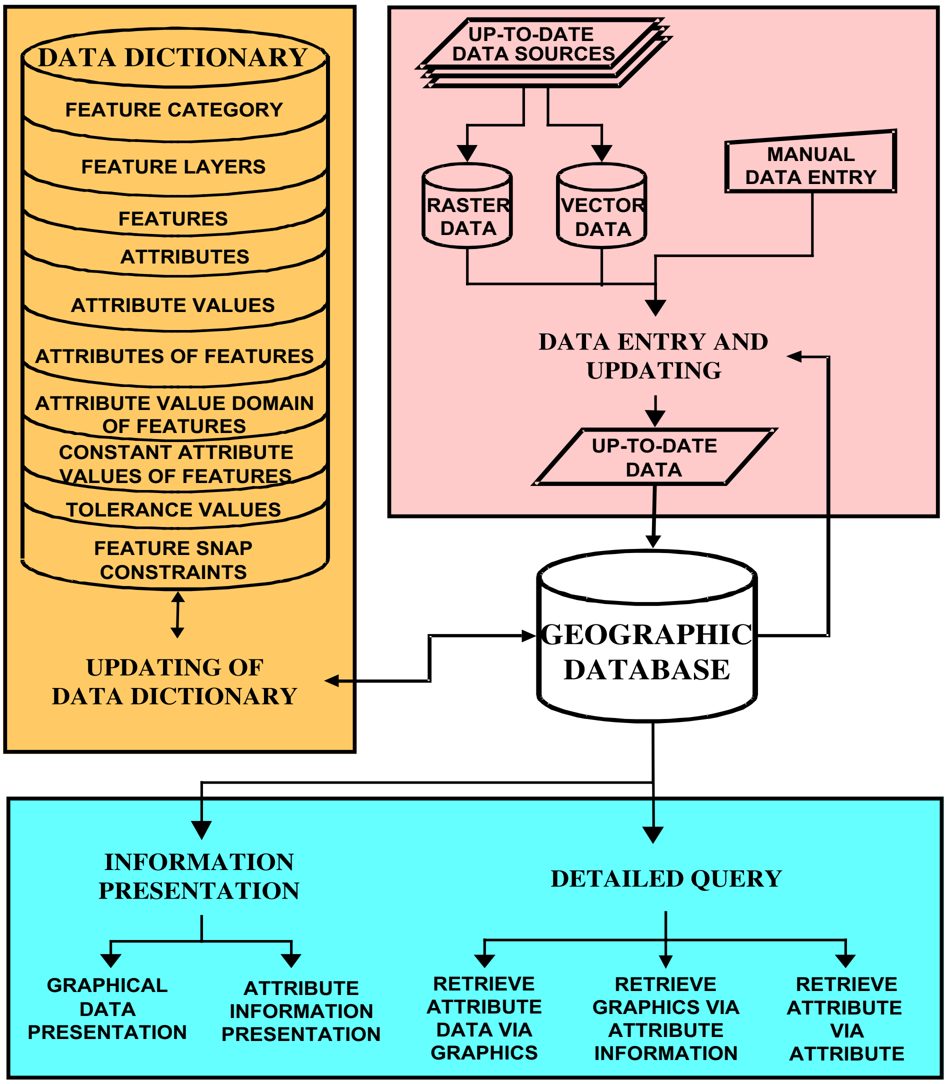

Figure 3. Functional architecture of the system

Two snapshots of the system during data capture/updating and data query are presented in Figure 4 and 5 respectively. Figure 6 shows data entry and updating window of the data dictionary of geographic database. Data dictionary query window is given in Figure 7.

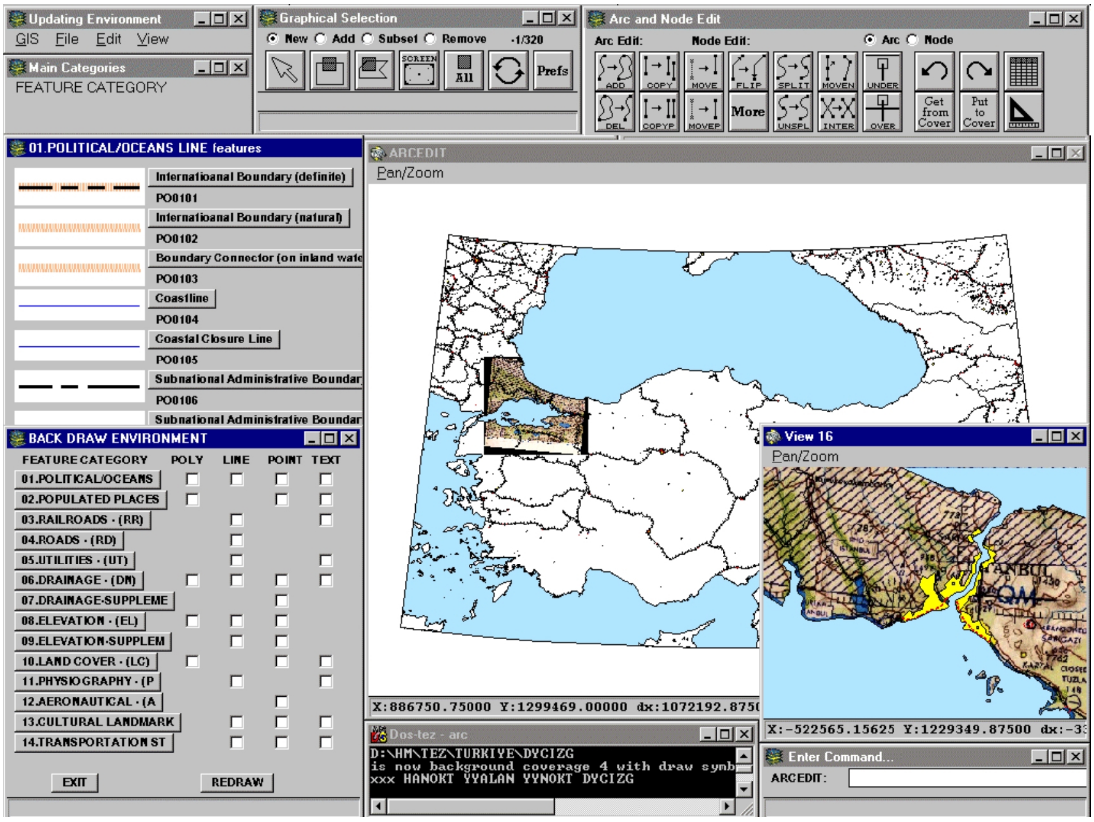

Figure 4. Screendump from the system during data capturing and updating

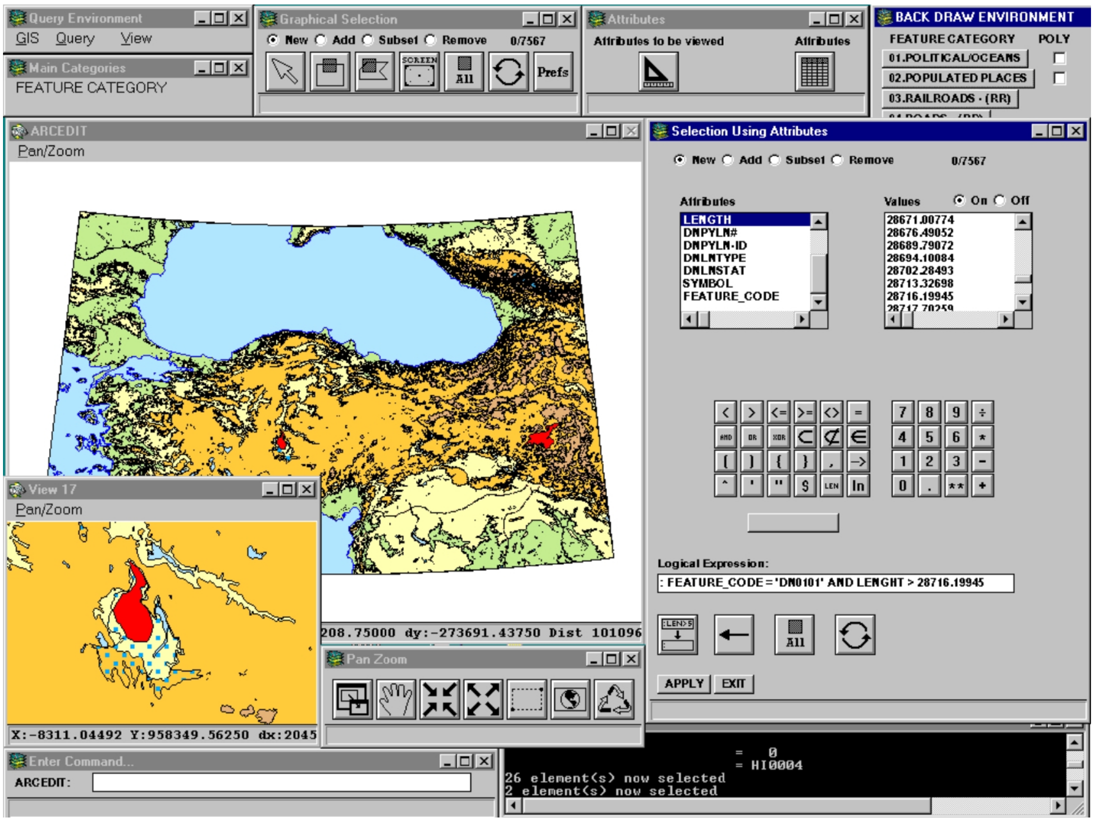

Figure 5. Screendump from the system during data query.

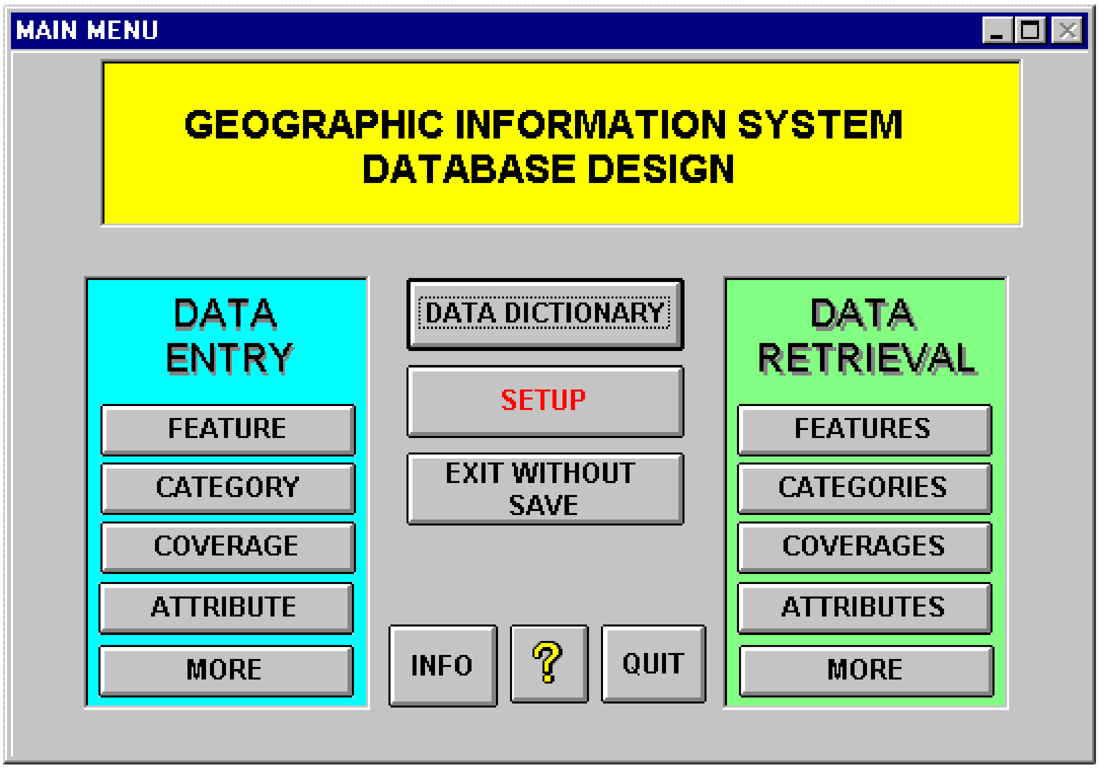

Figure 6. Data dictionary updating window

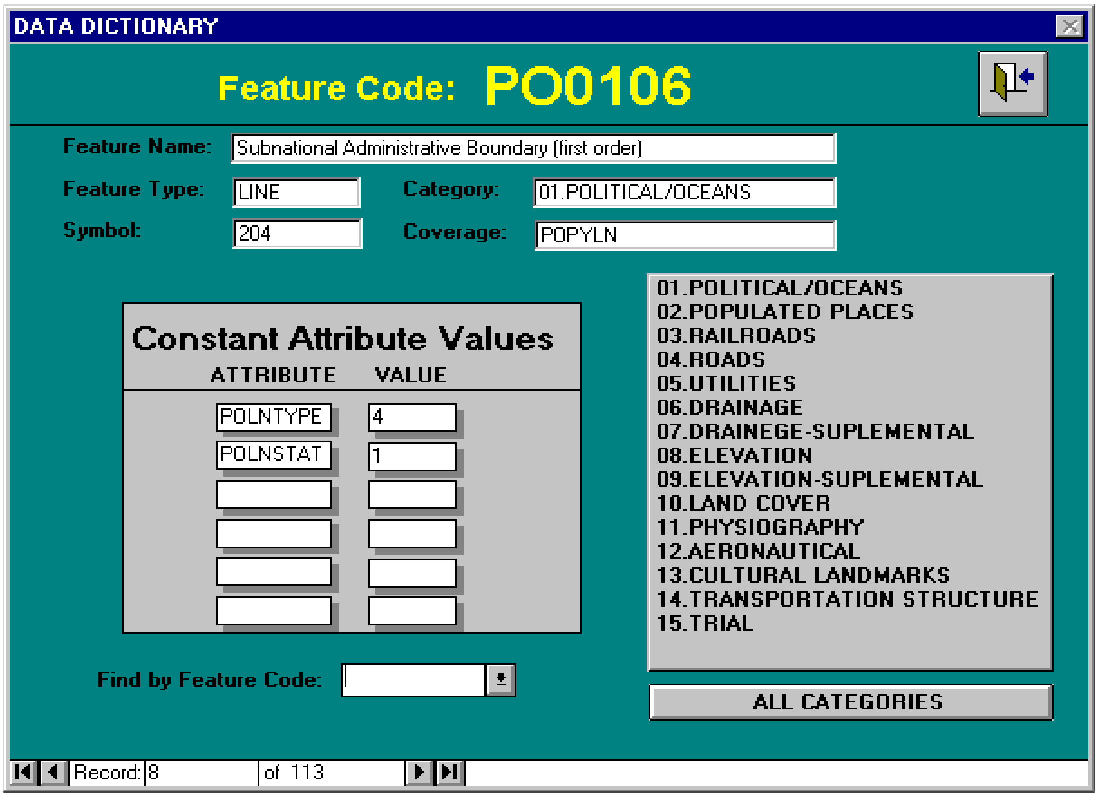

Figure 7. Data dictionary query window

4 CONCLUSIONS

The need of up-to-date data should be considered to provide accurate and reliable information. To explain our observations from the study and from other references, the main conclusions can be summarized as follows:

- Today, to develop easier-to-use GIS applications is possible even in PC environment,
- It is possible to develop application software enabling its interface automatically based on update of data model,
- The more complex the database is, the harder its updating,
- After developing or purchasing of application softwares, the updating and its monitoring will be quite easy, economical and fast.
- While updating the database the user should be very meticulous to prevent unrecoverable mistakes,
- To be reliable, the GIS database used in query and analysis which will be used in decision support, should be updated.
- Because there is a few methodology for GIS database updating in literature, knowledge and experience on use of GIS software and Database Management Systems is needed in developing applications,
- To make deleted data accessible, backup should be made,
- User friendly interfaces need much programming and more program codes.

REFERENCES

Altan, M.O., Toz, F.G., ve Külür, S., 1996. Developments in Information Systems and Photogrammetry, CBS 96 - Geographic Information Systems Symposium, Istanbul, pp.63-69,

Burrough, P. A., 1997. Geo-information needs: effect of scale, ITC Journal, 3/4, pp.235-236.

Davis Jr., C.A., and Zuppo, C.A., 1995. Updating Urban Geographic Databases: Methodology and Challenges, Proceedings of GIS/LIS '95 Annual Conference and Exposition, Volume I, Nashville Convention Centre, Tennessee, November 14-16, pp. 269-276.

Gray, S., 1995. Updating of Complex Digital Topographic Databases, European Organisation for Experimental Photogrammetric Research (OEEPE) Report, Institut für Angewandte Geodäsie, Frankfurt am Main, March, pp.13.

Maras, H., 1999. An Application and Design of a Geographic Information System for Updating of a Geographic Database, Ph.D. Thesis, (in Turkish).