國立臺灣大學地理學系地理學報 第十六期：87-101 (1993)
DEPARTMENT OF GEOGRAPHY, NATIONAL TAIWAN UNIVERSITY,
JOURNAL OF GEOGRAPHICAL SCIENCE (16): 87-101 (1993)

應用系統化空間查詢觀念於地理資料供應及管理系統之新設計

A NEW DESIGN OF A GEOGRAPHIC DATA SUPPLY MANAGEMENT SYSTEM APPLYING SYSTEMATIC GEOGRAPHIC QUERY CONCEPTS

朱 子 豪

Tzu-How Chu

ABSTRACT

In Taiwan, aerophotos and 1/5000 topological maps are very important data sources for resource management and research. In practice, however, it is very hard to find exact sheets for a certain task. In order to make these important types of data supply more flexible, GIS techniques were applied to establish a data supply management system.

In this system, a new query approach was designed according to users natural query behaviors. Users can define their needs by various ways to make the search more natural and efficient. There are three basic types of query methods: 1) point-type, 2) line-type and 3) polygon-type. Also the query modes can be classified into three categories: 1) batch, 2) interactive and 3) semi-batch. Moreover, for interactive and semi-batch modes, the helps for locating the target areas can be in the forms of vector or image data of different degrees of detail on the background. Finally, an uncertain query concept and spatial-spectral interleaved data filtery were introduced the query process. Besides, for the data supply clerks, they can easily bookkeep data supplied and finish all the offical processes for data purchase automatically.

The structure of the system is flexible enough that users can adapt and modify it to manage different types of data sources and promote the use of those data. Therefore, this system concept can be developed to support the national geographic data dictionary in Taiwan and make most of data visible and readily available to the public.

Keywords: spatial query, data supply, Geographic Information Systems, query certainty.

摘 要

在國內，航空照片及不同比例尺（如1/5,000，1/25,000）之基本圖是資源管理及研究不可或缺的重要資料；但在實務上，取得這些資料的程序卻不夠方便與簡易。故爲了使此重要之資料供應工作能更有彈性，地理資訊系統（GIS）技術被引入，而應用在此一地理資料供應管理系統之建構中。

在此系統中，依據使用者之自然查詢行爲，作者設計了一新的資料查詢方式。使用者可以透過多元之途徑，更自然而有效地滿足他們的需求。在此新設計中，查詢可由不同之角度來分類。本研究依查詢對象之類型（點、線、面），查詢區之界定方式（成批式，半成批式或互動式），查詢之確定性，查詢之步驟多寡，查詢之位置參考資料類型等，來細分查詢之每一查詢皆可由上述之分項加以組合描述。本系統中，管

Associate professor, Department of Geography, National Taiwan University.

理者亦可透過方便之查詢及輔助之管理功能，增加其管理及資料供應上之效率。
而此系統之設計在結構上十分彈性，故可以擴增以接受不同資料爲查詢及管理之對象，進而使該資料更易爲使用。因此，此系統之觀念便可發展，以支援國內國土資訊系統之資料典及其資料供應，使資料流通更加方便。

關鍵詞：空間查詢、資料供應、地理資訊系統、查詢確定性.

INTRODUCTION

Since the government has been promoting the national GIS for more than 5 years, most GIS related departments have increased their amount of geo-data usage and production. Among all, different scales of topo-maps and aerophotos are the most frequently used data. Due to this fact, the Aero-Survey Department for Agriculture and Forest, which produces 1/5000 topo-maps and most of Taiwan's aerophotos, and the Information Center of the Ministry of Interior, which produces 1/25000 and 1/5000 digital topo-map data, both decided to develop a data supply management system to make the data supply task more efficient to handle the increasing demands for data.

In Taiwan, the total number of 1/5000 topo-maps is 3773, and the number is increasing currently. On the other hand, the amount of aerophotos has been accumulated to be more than 400,000, and it is increasing in the rate of roughly 60,000 per year. In spite of the huge amount of data, various users acquire their aerophotos or topo-maps in a very cumbersome way. Users have to register their needs on a set of data index maps and copy the photo or map numbers. Since this kind of data query method is inconvenient for both supplier and user, GIS technology was introduced to help resolve this problem.

This data supply management system was designed to accomplish the following goals;

1. Users can specify their needs easier, faster and with greater flexibility.
2. The purchasing-process can be as automated as possible.
3. The management of aerophotos and topo-maps can be computer-assisted.
4. In the future, the production of the index maps of aerophotos and topo-maps can be automated in order to expedite requests.

The purpose of this paper is to explain role of query techniques of GIS in this system.

THE CONCEPTS OF THE SYSTEM DESIGN

A system analysis (SA) was performed and the system needs were recorded through series of interviews. Owing to the variation of types of users and their methods of query, a flexible system with different query entries and procedures are the key requirement of the system design (Egenhofer, 1988).

As a result of a series of observation and induction, the query process was classified by different aspects, such as the ways of defining the query areas, the query data types, types of referring data, the number of query steps, the degree of query certainty, and completeness of the feature selection. Every query action can be described with all those aspects. In this paper, the features defining the query areas are referred to as query

features, and the data queried as target features. Firstly, according to the ways to define the query features, the query can be divided into three different modes—1) batch, 2) interactive, 3) semi-interactive. In the batch mode, users input their query features from a file, in which the data types (point, line or polygon), coordinates, or feature names are specified (see Figure 1). In the semi-batch mode, the query features are defined by selecting a feature by giving a name or by pointing out its location. In the interactive mode, the query features are defined by pointing or drawing on the screen. In the batch and semi-batch modes, users then can take advantage of pre-existing data to increase the query process.

Secondly, according to the types of the query features, the query process can be categorized into three types: 1) point, 2) line and 3) area. The point type is defined by any point feature or a coordinate pair (x,y); the line type is defined by any linear feature or a series of polyline coordinates; the area type is defined by any area feature or a series of closed polyline coordinates. If the types of query targets are counted, there will be nine combinations of query types (3 query feature types 3 target feature types). Different query types will correspond with different query utilities. Moreover, the point and line types can be buffered as an area type to satisfy users' demand (see

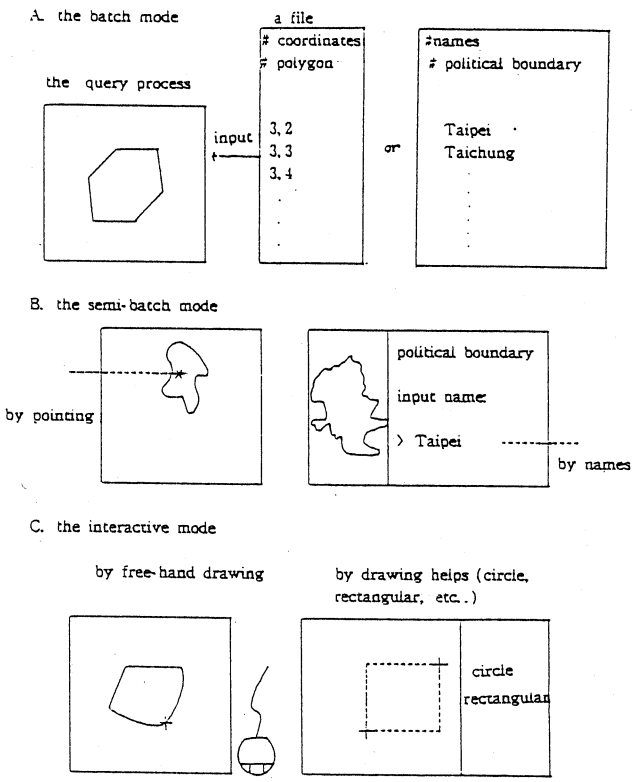


Fig. 1. Three query modes.
(A) The batch mode: inputing the query features by names or coordinates from a file.
(B) The semi-batch mode: defining the qurey features by inputing feature names or by pointing interactively.
(C) The interactive mode: defining the query features by drawing interactively.

Figure 2).

Query can be further classified by the degree of certainty— 1) certain, 2) uncertain. The higher the certainty, the fewer the required search activities. The batch mode and semi-batch mode are usually closely related with high query certainty, while the interactive mode is related with lower query certainty. The query with lower certainty should be assisted with more referencing data, and usually requires more than one step.

The referencing data for query process are very important for the query process with lower certainty. The common types of these data are as follows:

1. the scanned topo-maps (various scales to be selected as needed);
2. the remote sensing images (orthogonal aero-photo or satellite images);
3. the vector data with landmarks (common used vector data and easily memorized landmarks).

Other references such as three-dimensional landscape images are also a good searching help, but such references are more expansive and rarely used. In terms of the number of steps of the query procedures, there are four variations: 1) direct and single step; 2) direct and multi-step (through the logical query operators in Table 1), 3) indirect and move-around (through the moving operators in Table 1), and 4) filtering after spatial query. The first and second ones have higher certainty than the other two. The logical query operators are AND, OR, XOR, NOT etc.., and are used to redefine the query set. The indirect variation always needs reference background and moving assistant tools

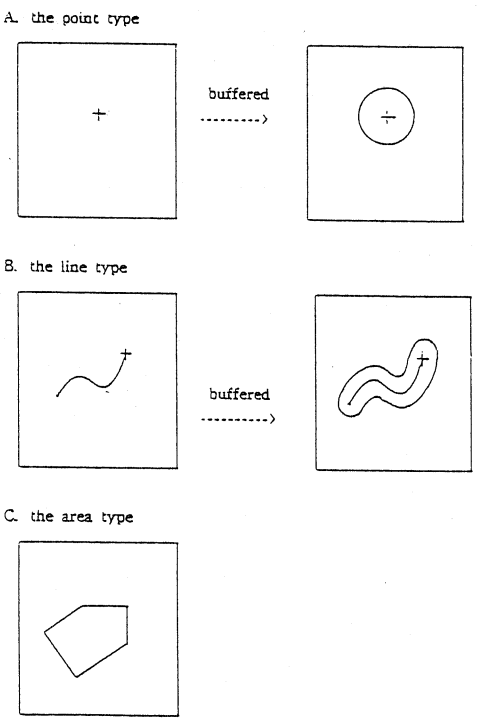


Fig. 2. Three query types.
A. The point type; B. the line type; C. the area type.

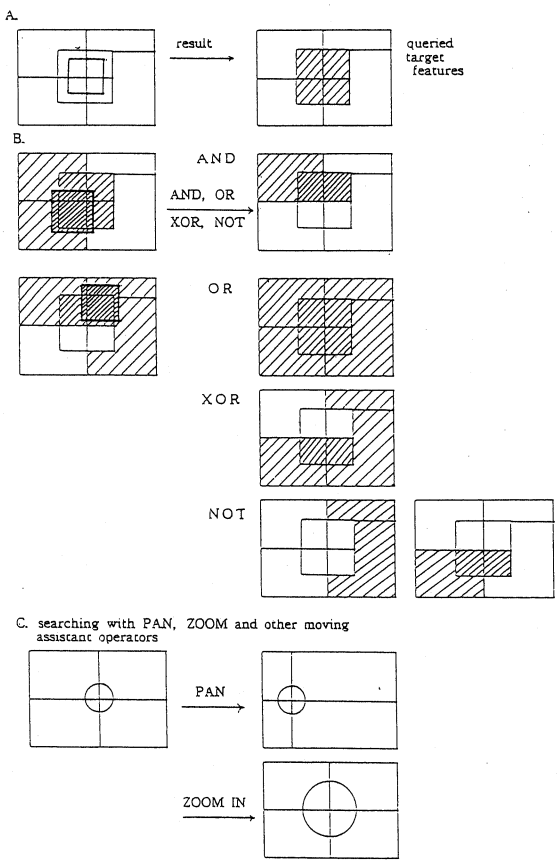
such as ZOOM and PAN. The last variation of query process can be called post-selection, and sub-divided further by the methods of filtering into 1) special formula, 2)optimization (by attributes or spatial attributes), 3) interactive way, and 4) interleaved query, a combination of spatial and attribute query(see Figure 3).

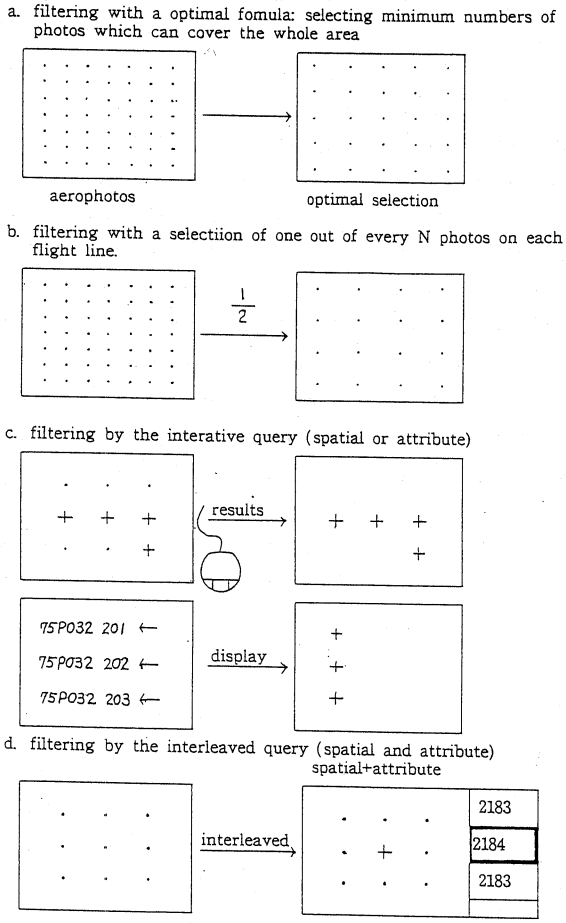
Finally, according to the relationship between the selected objects and specified query areas, the selection methods can be divided into three types: 1) included, 2) excluded, and 3) partial selected (see Figure 4). Usually users would only select one of the first two types because the last type reforms the target features by the query features.

While attribute query is much simpler, and can be classified into 1) direct query and 2) filtering after query (including spatial query afterward). Then the spatial query and attribute query can be combined to make interaction, and are possibly interleaved in a query process to make the query process more natural(Figure 5).

It is necessary to build a data base for frequently used features to make semi-batch query mode easier. The system is flexible to add on any new features to help query activities.

In the system, query features included are political boundaries, river basin boundaries, rivers, transportation lines, and place names. The target features are in point





D. the filtering (post-selection).

Fig. 3. The different steps of query procedures:
A. the direct and single step;
B the direct and multi-step (with logical operators);
C. the indirect query (with moving operators);

data for the aerophotos and in polygon data for 1/5000 topo-maps. The reference data for locating are political boundaries in vector type, 1/2,000,000, 1/250,000 and 1/100,000 and scanned topo-maps in raster type. The target features — aerophotos and topo-maps were only partially scanned for demonstration of the future system development. This data bundled concept should be a new approach for system development in order to make the system available for immediately and convenient use.

SYSTEM IMPLEMENTATION

The system was developed under MS DOS and on a PC with super VGA. The soft-

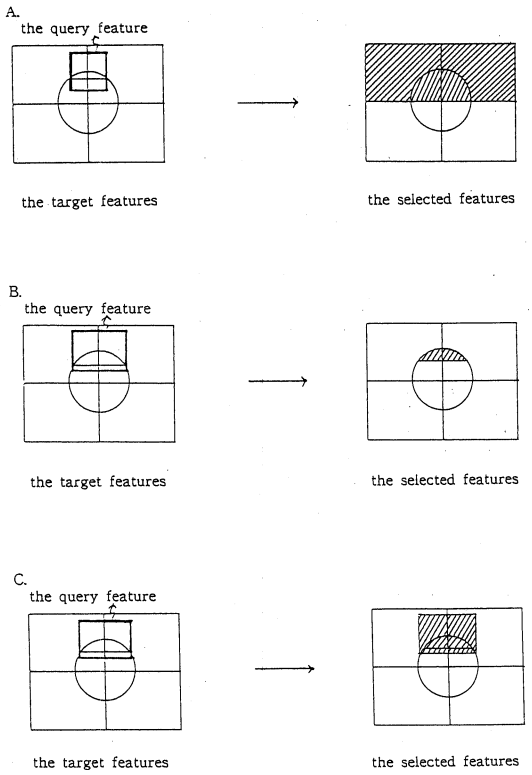


Fig. 4. Three different object-selection types:
A. included (select every target features intersected or included in the query feature);
B. excluded (select only target features fully included in the query feature);
C. partial selected (cut the target features by the query feature)

ware was written in MS C language, and the graphics and image utility library used was built by geography students. All the user interfaces were designed in the pop-up or pull down menu forms. The whole system is presented in Chinese using the language system developed by the research assistants.

The system was divided into three major modules — the sales and management, the query, and I/O modules. In this paper, only the query module is discussed. In the query module, spatial and attribute query sub-modules were included, but only the spatial query sub-module is presented. The system function structure is shown in Figure 6.

Digitizing tablets were used for establishing data bases. The aerophoto data were digitized as point data using ARC/INFO and transferred into DBASE format with x, y coordinates. There were 250 1/500,000 index maps with roughly 180,000 aerophotos digitized. The data were stored separately in different 1/500,000 coverages and different years. The 1/5000 topo-maps were also digitized by ARC/INFO and kept in the same format. There are two different sets of coverage grids and two different versions for 1/5000 topo-maps. For flexibility and compatibility of the system, the topo-map data were kept in the original format. For the same reason, all the query features in vector form were copied from ARC/INFO data and kept in the original format. The scanned

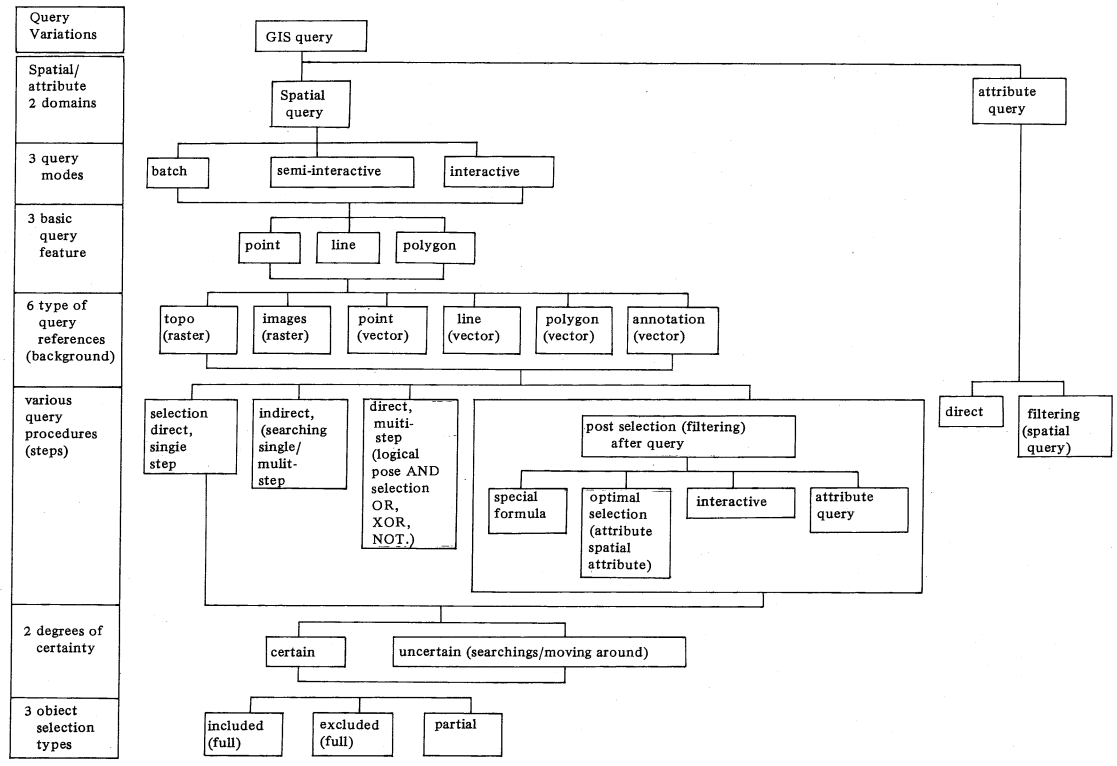


Fig. 5. Different query methods.

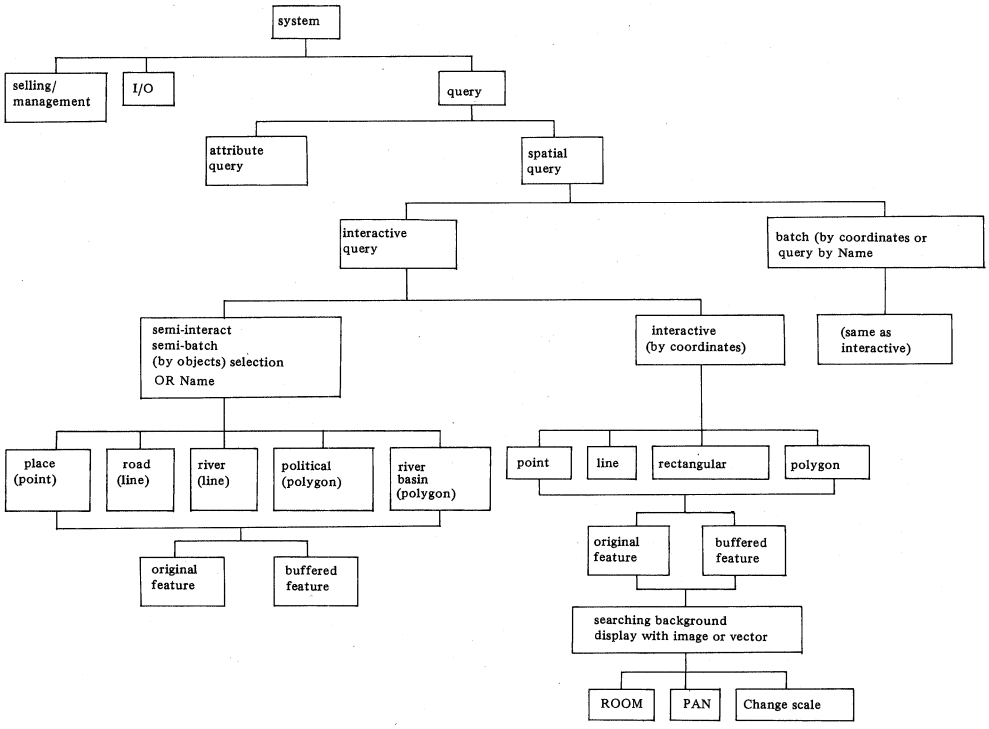


Fig. 6. The system fuction chart.

data however were designed in ERDAS (a remote sensing image processing system) format to be compatible with the image data in the department of geography. The topomaps and aerophotos of various scale were scanned by an A3 size color scanner, and stored in an optical disk for mass storage space. Regarding attribute data, the DBASE format was selected to store various attribute data, but manipulated with our own utilities. The data base structure is presented in Figure 7.

The special features of this system are listed as follows;

1. A user defined menu system allows users to change the system function configuration and expanded functions.
2. It is highly compatible with user written utilities, and there is no limitation of function expansion.
3. It already includes the most possible query utilities for various types, therefore the users or developers can build their own query functions.
4. It separates the query features and target features, and users can define their own query features and target features to fit their own purposes.
5. There are different locating references in order to make the query (especially query with uncertainty) more natural.
6. The spatial query and attribute query can be interleaved to make the post-query data filtering task much easier and more intuitive.
7. The target features can be selected and shown not only by location and attributes,

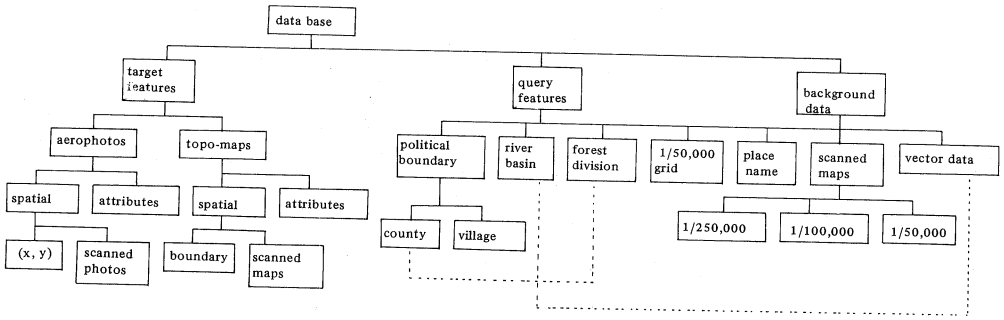


Fig. 7. The data base configuration.

but also by true features by showing their images.

8. It introduces the concept of certainty into query activities. This makes the query process more natural.

Figure 8 shows the menu of different levels, and all the menu can be modified by users. In Figure 9, a query feature— political boundary can be used to directly define the target query area in the batch mode, or help the users to find the rough location in a semi-interactive query mode, and the polygon defined by users was used to show the query area in the interactive mode. In Figure 10, different references — images and vector data were used to help users locate the target query area. In Figure 11, the query set selected after a spatial query can be filtered by their attributes interactively through pointing to the selected element in the attribute table and also through showing its location. This shows that the spatial query and attribute query can be interleaved to make the query more natural. In figure 12, the selected aerophotos can be checked upon request, in order to help users to filter the query sets, or to pre-view before they make the decision to order them.

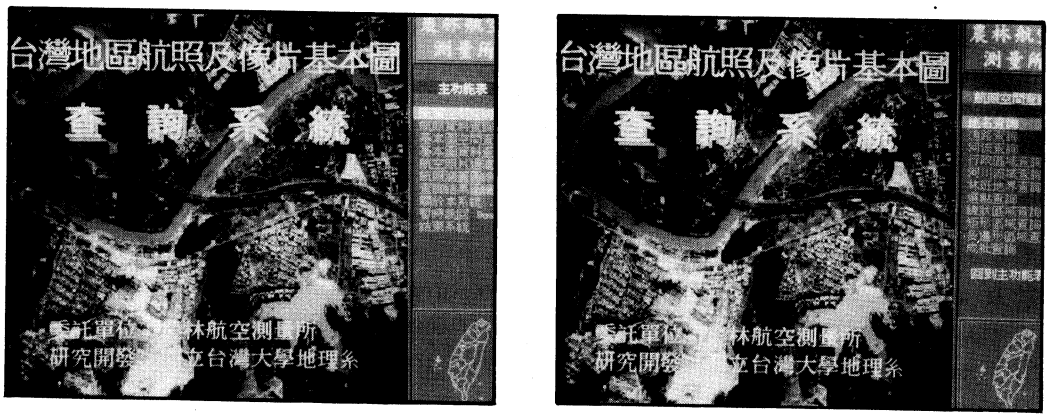


Fig. 8. User definable menu of differ A. major menu B. sub-menu.

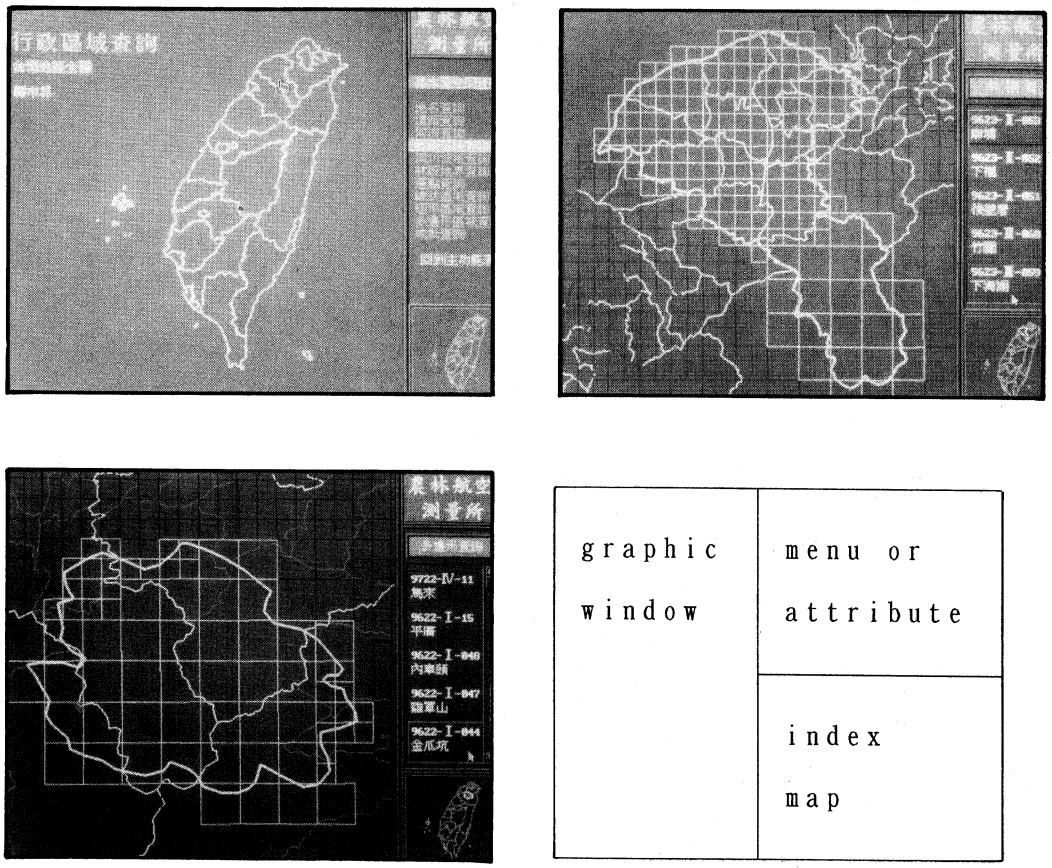


Fig. 9. Three different query modes.
A. the batch mode using political boundary as the query feature;
B. the semi-interactive mode also using political boundary to help select the rough query location;
C. the interactive mode using mouse to define query areas.

FUTURE DEVELOPMENT OF THE SYSTEM

The concepts behind the system can be developed further in many ways, such as reclassifying the query methods by further induction, defining moving or searching tools for uncertain query, and query operators for multi-steps query process. Regardless of approaches taken, a formalization process is necessary. Therefore, the formalization of operator-like form is proposed for future system development.

The query operators can be classified into two major types. One type is the area/feature selection operators (selection operators in short), and the other type is the area/feature searching operators (searching operators in short). A preliminary function set is shown in the Table 1.

On the other hand, the system can be more powerful by enriching its query data feature bases, especially those frequently used by most users, such as the names printed in base maps.

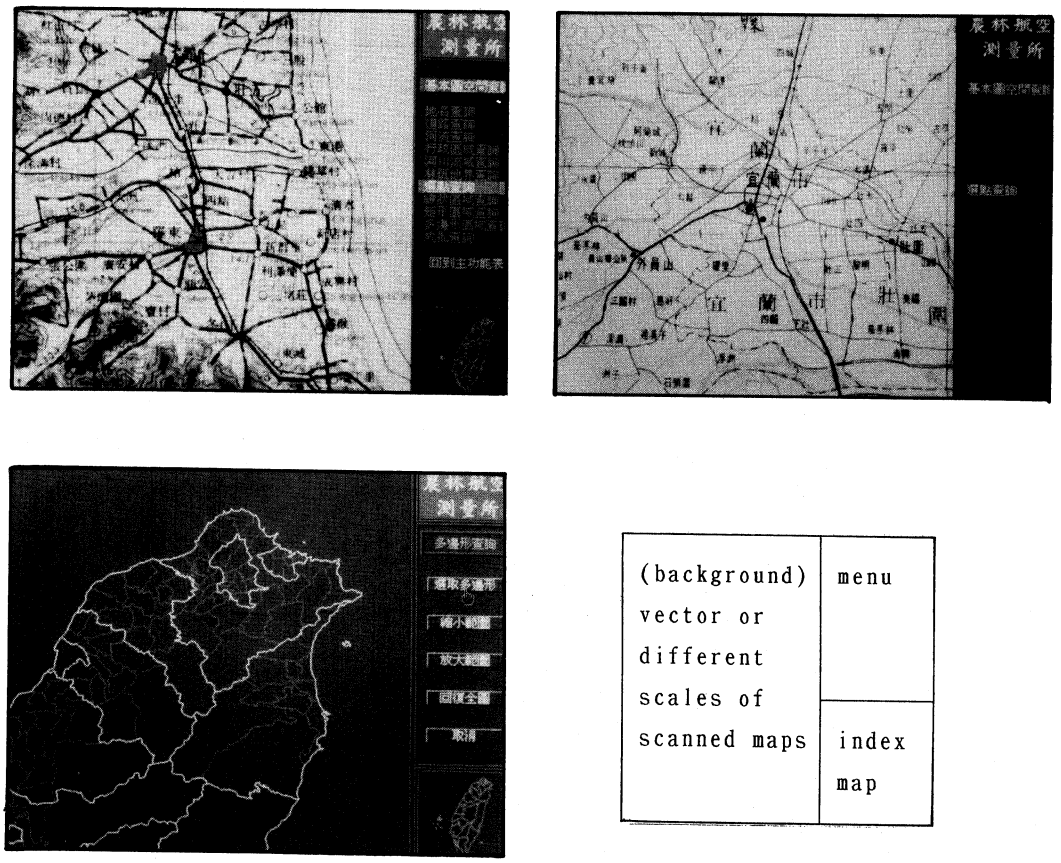


Fig. 10. Different references data for locating query areas.
A. a 1/250,000 topo-map in a raster type.
B. a 1/100,000 topo-map in a raster type.
C. a political map in a vector type.

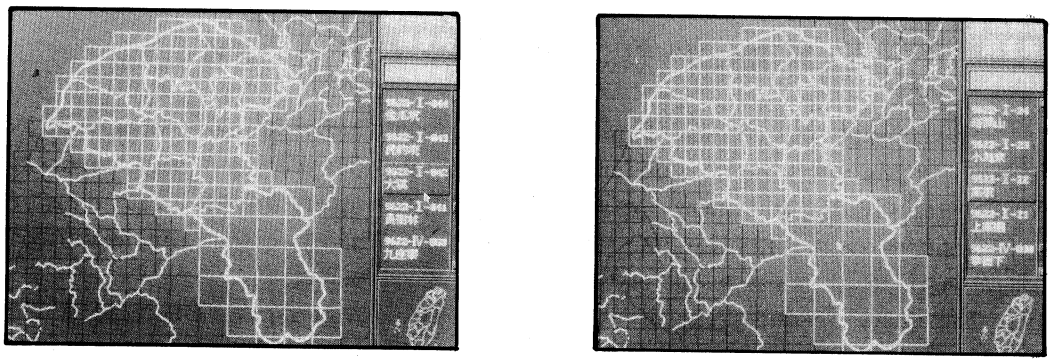


Fig. 11. The spatial query and attribute query can be interleaved through query set filtering. Users can locate one of them by pointing to the map or list of another.
A. using the attribute to check the object location.
B. using the object location to check the attribute.

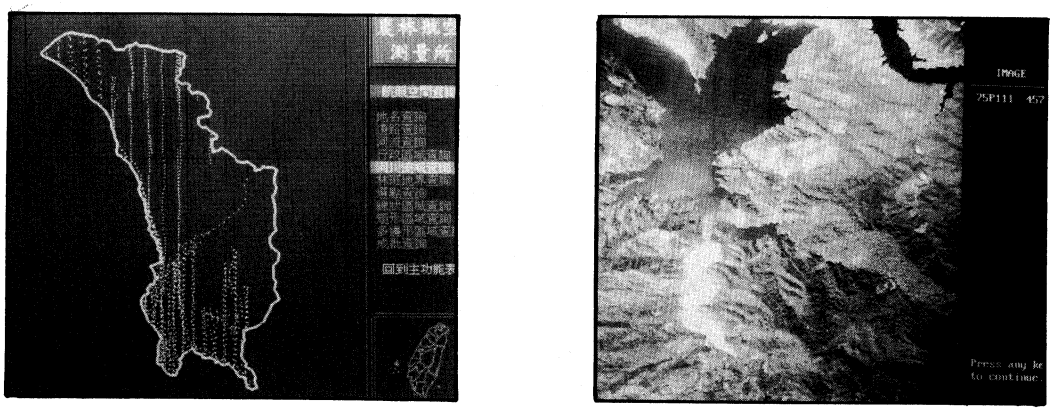


Fig. 12. The aero-photos can be pre-viewed before ordered.
A. selected aerophotos;
B. one of these aerophotos in a scanned form.

The system can be transformed to a common-purpose query system for different data format (with different data reading utilities). The system can be used directly for querying GIS data spatial dictionary with specified data coverages. For example, those data digitized in the topo-map coverage units can be queried as the topo-maps, while other maps digitized in various spatial units can be treated as target features of polygon type. This transformed tools should be able to promote the usage of the data in the data directory of the national GIS.

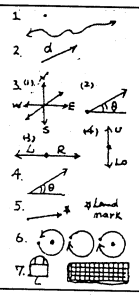
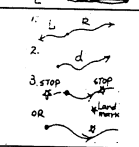
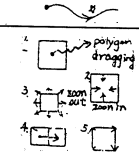
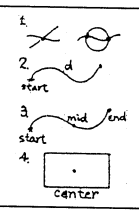
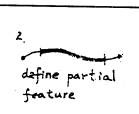
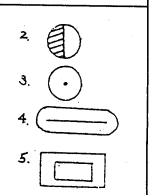
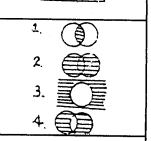
CONCLUSION

In the survey of GIS function needs for the Koushung City GIS integrated plan, most of the GIS related tasks require query functions. Unfortunately, query methods in most GIS packages were insufficiently developed. In this paper, a more flexible and intuitive way was studied and implemented for the supply task of aerophotos and 1/5000 topo-maps. The potential effects lies in saving map purchasing time and raising the service quality, and as well as making the data management more effective.

Because of the importance and high demands for query functions, further serious research should continue. In this paper, only a preliminary research structure was proposed.

This query system with common-used query feature data bundled makes the system effective and immediately available for use. Therefore, managers of GIS of natural resource management and other GIS related tasks and system developers should carefully consider the system to support their existing efforts.

Table 1. A preliminary query operator set.

types	query features	operator name	description	simple graphical description
searching operators (Moving/ pan)	point	MP-FL	1. follow a near-by line	
		MP-DIS	2. move by a given distance (with default direction or instruction from the mouse or keyboard)	
		MP-DIR	3. move by a given direction (with default distance or instruction form the mouse or keyboard, with different option: orientation, direction, left/right, upper/lower.	
		MP-DD	4. move by a given distance & direction	
		MP-GO	5. goto (jump to) a near-by object or landmark	
		MP-CUR	6. move alone a special curve (such as ellipse, circle, etc..)	
		MP-FRE	7. move with the mouse or keyboard (free hand)	
	line	ML-DIR	1. follow the line and move in a given direction (with default distance or instruction from the the mouse or keyboard.)	
		ML-DIS	2. follow the line and move by distance (with default direction or instruction from the the mouse or keyboard.)	
		ML-GO	3. follow the line and go to next/previous stop or a special feature or landmark nearby the line	
	polygon	MPL-POL	1. move as in the point mode	
		MPL-IN	2. move inside the polygon as in the point mode	
		MPL-OUT	3. move outside the polygon as in the point mode	
		MPL-AJ	4. move to a adjacent polygon.	
		MPL-AL	5. move along the boundary of the polygon as in the line mode.	
selection operators (single step)	point	SO-NP	1. select by name (semi-batch mode)	
		SO-ILL	2. select by intersection of lines or lines with polygon	
		SO-ILP		
		SO-CL	3. select by the offset of a line from a starting point	
		SO-OL	4. select the center of a certain line feature or polygon feature.	
		SO-CPL		
		SO-PDP	5. select by known coordinates	
		SO-PP	6. select by pointing a feature with mouse or keyboard.	
		SO-DL	7. select by drawing a point. (interactive mode)	
	line	SO-NL	1. select by name. (batch mode)	
		SO-PL	2. select by specifying partial of a feature (variation of semi-batch mode)	
		SO-PPL	3. select by a polying specified by known.	
		SO-PL	4. select by pointing with mouse or keyboard (semi-batch)	
		SO-DL	5. select by drawing (interactive mode)	
	polygon	SO-NPL	1. select by name. (semi-batch. mode)	
		SO-PPL	2. select by specifying partial of a feature (variation of semi-batch mode)	
		SO-BP	3. select by a buffer of a point.	
		SO-BL	4. select by a buffer of a line.	
		SO-BPL	5. select by a buffer of a polygon.	
		SO-PDPL	6. select by a pre-drawn polygon (batch mode)	
		SO-PPL	7. select by pointing a polygon interactive (semi-batch mode)	
		SO-DPL	8. select by drawing a polygon (interactive mode)	
post-selection operators (logical operation)	all	SL-AND	1. and (two query sets)	
		SL-OR	2. or (two query sets)	
		SL-NOT	3. not (single query set)	
		SL-XOR	4. xor (two query sets)	
post-selection operators (filtering)	all	SF-PT	1. select by pointing (spatial or in attribute table (attribute interleave))	
		SF-SS	2. select by specifying spatial interim	
		SF-RS	3. select by regular skipping (skip every n objects)	
		SF-SA	4. select by specifying attribute characteristics. (attribute interleave)	

REFERENCES

1. Brett J. Larkin, 1991. *An ANSI C Program to Determine in Expected Linear Time the Vertices of the Convex Hull of a Set of Planar Points,* Computers & Geosciences , 17(3), pp. 431-443.
2. Brett J. Larkin, 1991. An ANSI C Routine to Determine if a Point is Within a Specified Convex Polygon in Logarithmic Time, Computers & Geosciences , 17(6), pp. 841-847.
3. Environment Systems Research Institute, Inc., 1989. ARC/INFO Programmers Manual Volume 1, Version 5, California:Environment Systems Research Institute Inc.
4. Environment Systems Research Institute, Inc., 1987. PC ARC/INFO Starter Kit Users Guide, Redlands, California:Environment Systems Research Institute, Inc.
5. ERDAS, Inc., 1991. ERDAS Field Guide, 2nd ed., Atlanta:ERDAS, Inc.
6. Egenhofer, Max J.and A.U Frank, 1988. Designing Object-Oriented Query Languages for GIS: Hunan Interface Aspect; Proceedings of the 3rd International Sysposium on Spatial Data Handling, pp. 79-96.
7. NCGIA, 1989. TECHNICAL ISSUES IN GIS, California, Santa Barbara : NCGIA.