Framework for a Geographically Referenced Conservation Database: Case Study Nepal

*Sean C. Ahearn**
Remote Sensing Laboratory, University of Minnesota, 1530 N. Cleveland Ave., St. Paul, MN 55108
James L. David Smith
Department of Fisheries and Wildlife, University of Minnesota, 1980 Folwell Ave., St. Paul, MN 55108
Catherine Wee
Remote Sensing Laboratory, University of Minnesota, 1530 N. Cleveland Ave., St. Paul, MN 55108

ABSTRACT: Many national governments and most international development organizations are attempting to integrate conservation and sustainable development goals. To do so they need biological diversity information in a readily accessible form. Currently available conservation databases are constrained by reliance on paper map products as repositories for spatial information and an inability to quickly obtain information and integrate it with other map layers. A framework for developing a conservation database using the convergence of technological developments in digital imaging, remote sensing, micro-computers, database management, geographic information systems, global positioning systems, communications and user interface software is proposed. We also describe a pilot project in which we developed the software for a geographically referenced visual/tabular conservation database for the lowlands of Nepal.

INTRODUCTION

WORLDWIDE CONCERN FOR LESS OF BIOLOGICAL DIVERSITY has resulted in a variety of policy and program initiatives to slow the high rate of species loss. Miller *et al.* (1989) argue that it is time to shift from a defensive role of stemming the loss of species to an offensive position of integrating conservation goals with sustainable development and the needs of indigenous people. Current biological diversity information, that can be accessed easily, is needed by national governments and international development agencies as they attempt to integrate conservation and sustainable development goals. At present, conservation databases are limited because they lack geo-referenced data or the ability to rapidly access information and integrate it with other spatial data.

While movement is underway to automate conservation databases, a framework for doing so needs to be established. The convergence of technological advancements in digital imaging, remote sensing, micro-computers, database management, geographic information systems (GIS), global positioning systems (GPS), communications, and user interface software will have a profound effect on the nature of this framework. Recent developments in these technologies can help solve existing problems as well as provide new forms of information previously ignored. These technologies are impacting the development of conservation databases in the following ways: locational information is gathered more easily through GPS; data access is enhanced through the use of database management systems; geographically referenced visual catalogues of flora, fauna, and habitats are being developed through the use of imaging technologies (Hamilton *et al.*, 1989); and remote sensing enables the creation of image maps that will greatly enhance data input, interpretation of data, and data output products for decision makers.

In this paper we propose a framework for developing a conservation database founded on these new technologies. We also describe a pilot project in which we developed the software for a geographically referenced visual/tabular conservation database. It is not the intent of this paper to give a definitive statement or framework for a global conservation database; our objective is to begin development of the adumbrations for one in the context of today's relevant technological developments. The study site for demonstrating this software is the lowlands of Nepal.

CONSERVATION DATABASES

From the 1970s to the mid-1980s, the central clearing houses for endangered species information have been institutions such as the International Union for the Conservation of Nature's (IUCN) Survival Species Commission (SSC) which issues the Red Data Books, and The Nature Conservancy (TNC) which developed the Natural Heritage Data Base (NHDB). The SSC's Red Data Book has grown from a part-time effort relying on a network of volunteer wildlife experts to a separate institution, the World Conservation Monitoring Center (WCMC), Cambridge, which brought together bird, mammal, and protected area databases (Thornback and Jenkins, 1982). The present system has massive volumes of gray literature that are stored in files and periodically summarized in the Red Data Books or in specialized products based on individual requests. WCMC has recently undertaken a GIS project to develop an atlas of tropical forests (Mark Collins, pers. comm.) and is seeking input from the scientific and conservation community in developing new approaches to conservation databases. (Robin Pellew, pers. comm.).

The Nature Conservancy's Natural Heritage Data Base originated in 1974 as a cooperative effort between TNC and South Carolina to develop a state-wide conservation database (Jenkins, 1988). This original system has continually been improved, and the format has been adopted by all states. The key to its success is the highly standardized methodology that provides efficient communication and exchange of information as well as the economy of having TNC act as system administrator and developer. With the recent establishment of Latin American Conservation Data Centers in several South and Central American countries, the system is expanding throughout the western hemisphere. The data are used both internally by the TNC and state and national agencies and externally as products these agencies prepare for other organizations. Internal uses include products such as inventories and monitoring, setting of conservation priorities, preserve selection and design, and element management (an element is a species, a community type, or

*Presently at the Department of Geology and Geography, Hunter College - City University of New York, 695 Park Avenue, New York, NY 10021.

0099-1112/90/5611-1477$03.00/0

some other feature or phenomenon of special interest to conservationists)(Jenkins, 1988). In addition, external products are generated for other government agencies and research, conservation, and development institutions for development planning, environmental impact analysis, and predictive modeling (Jenkins, 1988).

Despite their tremendous strengths, the NHDB and the WCMC have the following weaknesses that derive from the fact that their maps are not digitally based:

- During data entry, obtaining geo-referencing is time consuming and prone to small scale generalizations as well as occasional large errors.
- For most users it is critical to know an element's location and spatial context. However, with the present systems it is difficult to generate map-based products with this information.
- These systems do not provide photographic or image products. The adage, a picture is worth a thousand words, is not trivial when trying to describe habitats or species or to create the emotional impact that motivates decision makers.
- It is not easy to create data layers that are needed for sophisticated spatial analysis in a GIS environment; the layers must be compatible in format, spatial content and quality.

FRAMEWORK FOR PROPOSED SYSTEM

The importance of conservation databases depends in part on how well they contribute to the making of informed decisions regarding the potential impact land uses may have on biological systems. With greater tendencies toward the use of GIS to make such decisions, it is critical that conservation databases be in a form that easily can be integrated into the GIS environment. However, when a GIS has not been developed for a region, the conservation database should stand on its own as a tool for sound environmental decision making. A conservation database can be thought of as a specialized subset of a generic GIS. The distinction between the systems concern their analytic capabilities and the data requirements of each system. The spatial operations necessary for a conservation database are confined largely to point in polygon and distance computations. It also is necessary to support combined spatial/attribute queries (i.e., find all tigers within a specified polygon in lowland forests. In contrast, GIS are multi-layered databases that have a wealth of spatial operations involving single and multiple layers.

It is also necessary for conservation databases to communicate more than just text information regarding elements. A visual component can turn the species cataloged in the database from an abstract to a tangible entity. An image provides habitat quality information that cannot be readily communicated in a text form. The technology is available to develop conservation databases to meet these requirements and a diversity of individuals and institutions (e.g., Conservation International; International Council for Bird Preservation; WCMC) have begun to develop such systems.

The wide range of computer expertise in the user community necessitates a system that is very simple for both data input and data access. For any database system the primary costs and impediments to a functional system are usually associated with data input. Therefore, it is necessary to minimize the time and expertise needed for the local input of data as well as to provide a means to regionalize and centralize the information.

DATA STRUCTURE

Conservation data sets do not form a continuous, topologically complete layer of geographic information but consist of a set of points or polygons that designate the location of elements. Associated with this spatial information is the attribute information that describes the nature and significance of the element. The attributes for a conservation database should include both structured and unstructured data. Both spatial and attribute information can be managed using a relational database. The cartographic base linked to the conservation data can be raster, vector, or both. While a raster system is easy to implement, it is limited in its ability to link the cartographic information with the conservation data. In contrast, a vector structure facilitates the linking of the cartographic data and its associated attributes with element data. For example, it might be desirable to find a specific element type within a particular political unit (i.e., a panchyat in Nepal). This could only be done with a topologically complete cartographic layer. Digital cartographic information is becoming more prevalent and is now available for the entire world at 1:1,000,000 (MundoCart, 1989). Digital line graphs (DLG) also are available for much of the United States at scales ranging from 1:24,000 to 1:2,000,000 (USGS, 1987). Digital images of the Earth's surface can be obtained at resolutions as low as 10 metres in Pancromatic mode to 20, 30, 80, and 1100 metres multispectral. These images can be geo-referenced and overlayed with cartographic information to create image maps (Colvocoresses, 1986). The advantage of the image map is the additional information it contains pertaining to cover types and topographic conditions. The image map is very useful for orientation purposes during data input and access as well as for data presentation. We suggest that the optimal spatial framework is a digital image map that maintains the cartographic data in vector form and images in raster. This linkage is essentially the approach used in TYDAC'S SPANS and in the ARCINFO/ERDAS Live Link. The approach has all the advantages of an image map while retaining the topological relationships necessary to access features within their spatial context. The types of spatial operations needed in a conservation database are relatively simple in comparison to a GIS. Spatial queries include proximity to a point, line, or polygon and point in polygon operations. These spatial queries are modified by searches for particular element types or sub-attributes of the elements.

DATA INPUT

A conservation database must be capable of ingesting both existing and new conservation data. Existing data in ASCII format can easily be incorporated by means of a data conversion routine. Written data sets need to be manually entered. An automated check on the validity of the geographic coordinates for all written and digital data sets is a requisite function of the system. This check is for major blunders such as the identification of data points that occur outside the state or region in which they are supposed to be contained. The specificity with which these checks can be made depends on the level of detail that is contained in the cartographic/image database used by the system. Manual checks of spatial location can also be made by the person entering the data. This is done by displaying the points or polygons on the cartographic/image base as a test for consistency with the element's location as defined by its political or geographic location. The coordinates for new elements can be derived from two possible sources: (1) GPS receivers or (2) images, maps, or image-maps in either hardcopy or digital form. GPS enables the recording of locational information while data are collected from the ground. Ground GPS systems have been found to be very useful even in the most demanding of conditions (Wilke, 1989) and can provide very accurate vertical and horizontal coordinates during field examination. Screen digitization on digital cartographic or image maps automatically records the geographic coordinates for the elements entered. Elements located on the paper products can be electronically digitized after map registration. The cover type information interpretable on an image-map can be used to guide the person to the location of a particular point or help identify regions that may contain elements that need to be recorded. This is of particular importance in areas with few cultural or natural features.

Data Retrieval

Retrieval of data will be required through several different avenues: by attributes, by geo-coordinates, by region or country, or by pointing to a location on the cartographic base and/or image map and invoking a search. The search region can be specified as a radial distance from a chosen point or it may be delimited by a user defined polygon. The search may be made for all elements in the search area; for a specific element type, or it might include any of the other attributes associated with that element.

Integration with a Geographic Information System

Any data for which the geographic coordinates are known can be integrated into a GIS. The spatial and non-spatial information associated with element occurrences in our proposed conservation database framework can easily be imported into a GIS. While most of these systems do not accommodate unstructured data, many of the analyses performed using the GIS would involve only the structured attribute data and its related spatial information.

PILOT SYSTEM

A pilot project was developed for Nepal's lowlands to (1) meet the needs of the Department of National Parks and Wildlife Conservation (DNPWC) that were not being addressed by their text conservation database, (2) provide government and international development organizations in Nepal with biological information (e.g., distributions of rare and endangered species and communities; locations of critical habitats), and (3) demonstrate to DNPWC and other departments how a nationwide conservation database can function to influence policy and land-use decisions in areas over which the DNPWC does not have direct jurisdiction.

The Director General of DNPWC felt that the existing text conservation database did not provide products and services to justify a full time wildlife officer staff position. The Department could not produce hand drawn map products rapidly enough to meet the needs of other departments and these agencies had little interest in the text reports that were generated. Our database will address these problems by providing instant access to and rapid output of information as maps with the appropriate information overlayed, images (e.g., species; habitats; land use), and text.

We focused on the conservation strategy for tigers (Smith *et al.*, 1987) to (1) demonstrate how our system can be used to provide information to other government agencies and (2) suggest a scenario for how this information can be used to influence land use outside DNPWC jurisdiction. The major conservation issue concerning tigers and other large animals in Nepal and elsewhere in Asia is that they need large areas to support the minimum population size required to assure population survival. The primary tiger populations in Nepal are centered around two reserves that have been expanded to encompass as much tiger habitat as possible. Nevertheless, these reserves contain only half to two thirds of the tigers in these populations. The remaining animals are distributed across a narrow belt of forest that once was continuous across the lowlands, but now is fragmented into a series of habitat islands. Heavy use of these forests for cattle grazing, fodder harvest, and fuel wood collection continues to fragment tiger habitat and threaten populations of this endangered species with extinction in the near future.

Data on the distribution of tigers were entered into our database. All tiger sites outside the protected area system were labeled as high priority areas for protection if habitat degradation could result in additional forest fragmentation. This information can now be produced rapidly and with emotional impact. For example, the World Bank has recently initiated a Lowland Community Forest Program to stem the degradation of natural forests by providing alternative sources of forest products. The ability to produce a map showing high priority tiger habitat sites, photographs depicting forest degradation, and a description of the importance of these sites to tiger conservation may influence decision makers to incorporate the conservation objective of buffering threatened tiger corridors when planning the locations of community forestry projects.

Components

The primary components of the software system are the database manager, the cartographic base, and a link between the two environments. The database system prototype stores geographically referenced text and images. Information is managed by a relational database. The relational database used in this case is called Structured Query Language (SQL) (Elmasri and Navathe, 1989). SQL has been available since 1974. The SQL we used is the IBM OS/2 Extended Edition Data Base Manager. SQL is based on the relational database mode, where a database is a collection of one or more named tables. Each table is an unordered collection of rows, each of which has a fixed number of columns. Each column has a heading (name) associated with some attribute of a certain data type. As in other database languages, SQL has a data definition language (DDL), with which one can define a database, and a data manipulation language (DML), with which one can manipulate the data in that database. Table 1 provides an example of how DDL is used to create tables in the database. In the tables, spatial information is combined with other structured information, such as habitat, species found, name of observer, landuse, and unstructured information such as text files and images in the relational database. Unstructured data are managed in their respective medium and columns in tables are reserved for linking structured information to unstructured information.

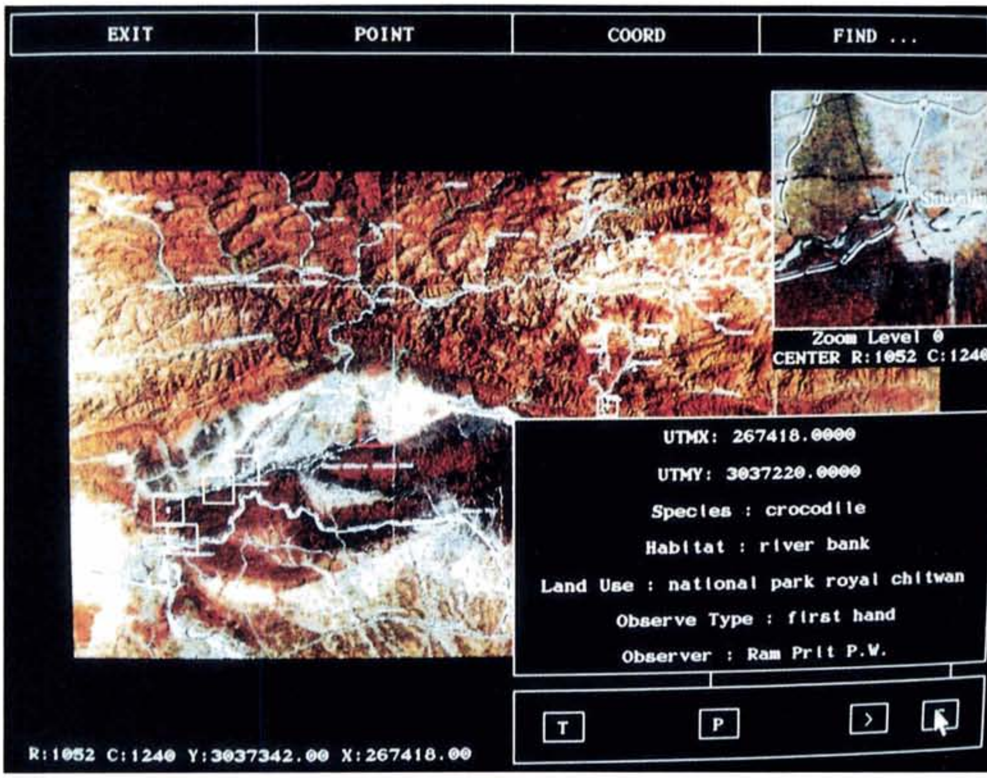
The user specifies information to be retrieved and them DML

Table 1 This Example Shows How Three tables Are Created. Table1 Contains Structured Data Such as Position, Index to Species Found, Index to Habitat, Name of Observer, and Index to Land Use. It also Contains Reference to Unstructured Data such as a Filename to a File Containing Additional Information (e.g., Notes, Scientific Research) about this Particular Location. Table2 Allows Reference Using Positional Information to Unstructured Data such as Images of the Element and its Habitat at that Location. Table3 Contains Species Information and its Corresponding Index. It also Allows Image of the Species to Be Referenced.

```
CREATE TABLE table1
(x_position decimal(13,4) NOT NULL,
y_position decimal(13,4) NOT NULL,
species_indx smallint NOT NULL,
observer_name char(80) NOT NULL,
observe_indx smallint NOT NULL,
habitat_indx smallint NOT NULL,
landuse_indx smallint NOT NULL,
text_filename char(80)).

CREATE TABLE table2
  (x_position decimal(13,4)NOT NULL,
  y_position decimal(13,4)NOT NULL,
  video_filename char(80)NOT NULL).

CREATE TABLE table3
  (species_indx # smallintNOT NULL,
species_name # char(40)NOT NULL,
video_filename char(80)NOT NULL).
```



(a)

(b)

PLATE 1. (a) Example of pointing and clicking at a gharial (fish-eating crocodile) location (red box) on an image-map of Nepal to display structured attribute data in the large window. Boxes below the attribute data select: T, free form text; V, images, P, previous screen; <, previous location; >, next location; E, exit. The small image window gives a close up of the location. (b) Gharial/habitat image associated with the above location.

is used to formulate a query to retrieve information from the database. Table 2 gives an example of how DML can retrieve all tiger locations in Nepal. Because the general purpose programming language used is C, all the above DML are used in conjunction with this general purpose programming language by imbedding SQL calls within the C language program.

For the prototype a raster image map referenced to the UTM coordinate system is used as the cartographic base on which new element locations can be designated for input into the database or existing elements retrieved.

Data Input and Retrieval

New data are input by entering geo-coordinates for the element occurrence or by digitizing its location on the digital image map which automatically places geo-coordinates in the appropriate field of the new element occurrence. A window will then open that contains the attribute template for the element. If free form text and pictorial image files also are to be included, the file names are entered in the appropriate fields. Input of text and image capture can occur by starting a new session in a separate window. Data can be retrieved by attribute for a given geographic location which can be typed in or selected on the image map. A radius about a point or a search within a defined polygon can be chosen. These searches result in the display of a summary data form stating the number of occurrences, the locations of all records that fit the search criteria as displayed on the image map, and a typical example (digital picture) of the element of interest. Information for a selected element is displayed in various windows. Icons indicate if additional free form text or images are associated with a given site. Map and image data for the given location can then be displayed or printed (Plate 1).

Table 2. In this Example, table3 Is Used to Obtain the Corresponding Index for Tiger. Using this Index, all Locations of Tiger Are Retrieved from table1. Using these Locations, the Image of these Locations Can Be Produced from table2.

```
EXEC SQL

DECLARE cursor1 CURSOR FOR
  SELECT SPECIES_INDX, VIDEO_FILENAME
    FROM table3
  WHERE SPECIES_NAME = "tiger".

EXEC SQL
  DECLARE cursor2 CURSOR FOR
    SELECT X_POSITION, Y_POSITION, OBSERVER_NAME,
      OBSERVE_INDX,HABITAT_INDX,LANDUSE_INDX,
      TEXT_FILENAME
    FROM table1
    WHERE SPECIES_INDX = :species_indx.

EXEC SQL
  DECLARE cursor3 CURSOR FOR
    SELECT VIDEO_FILENAME
      FROM table2
    WHERE (X_POSITION = :x_position and
        Y_POSITION = :y_position).
```

CONCLUSIONS

As conservation and development programs increasingly require decisions at local, regional, and global levels, the success of conservation efforts will depend on the establishment of an international network of conservation data centers. The WCMC and TNC recognize the need for these networks and currently are developing them. If we are going to martial the information in conservation databases to realistically integrate conservation and sustainable development goals, we need to combine databases such as those maintained by the WCMC and TNC to eliminate costly duplication of effort and maximize use of information stored in these systems.

The system described in this paper is a pilot system for concept exploration. Some advantages of the system are: it is geographically based, it is easy to input existing and new data, and it generates a wide array of output products. Because the system incorporates digital maps and requires and facilitates georeferencing, it provides spatial output products that currently are not readily available (e.g., hard copy maps, GIS data layers). Instant access to maps, visual images, and text make our system an excellent medium for communication.

ACKNOWLEDGMENTS

We thank P. Lau for his contribution to software development, F. Cuthbert, and two anonymous reviewers for comments on the manuscript and R. Jenkins and staff at TNC for discussions on conservation databases. The manuscript was developed while Ahearn and Smith were supported by funds from the Minnesota Agricultural Experiment Station.

REFERENCES

Colvocoresses, A.P., 1986. Image Mapping with the Thematic Mapper, *Photogrammetric Engineering & Remote Sensing*, Vol. 52, No. 9, pp. 1499–1505.

Elmasri, R., and S.B. Navathe, 1989. *Fundamentals of Database Systems*, The Benjamin/Cummings Publishing Company, Inc.

Hamilton, M.P., L.A. Salazar, and K.E. Palmer, 1989. Geographic information system: providing information for wildland fire planning. *Fire Technology*, Vol. 25, pp. 5–23.

Jenkins, R.E., 1988. Information Management for the Conservation of Biodiversity. *Biodiversity* (E.O. Wilson, ed.) National Academy Press, Washington, D.C., pp. 231–239.

Miller, K.R., W. Reid, and J. McNeely, 1989. A Global Strategy for Conserving Biodiversity. *Diversity*, Vol. 5, pp 4–7.

MundoCart, 1989. Chadwyck-Healey LTD, Cambridge Place, Cambridge, CB2 INR.

Smith, J.L.D., C. Wemmer, and H.R. Mishra, 1987. A tiger geographic information system: the first step in a global conservation strategy. *Tigers of the World: the Biology, Biopolitics, Management and Conservation of an Endangered Species* (R.L. Tilson and U.S. Seal, eds.). Noyes Publ., Park Ridge, N.J., pp. 464–474.

Thornback, J., and M. Jenkins, 1982. *The IUCN Mammal Red Data Book*

USGS, 1987. *United States Geological Survey, Digital Line Graph Users Guides*, Branch of Technical Management, USGS, 507 National Center, Reston, Virginia 22092

Wilke, D.S., 1989. Performance of a Backpack GPS in a Tropical Rain Forest. *Photogrammetric Engineering & Remote Sensing*. Vol. 55, No. 12, pp. 1669–1812.