IMPROVED REMOTE SENSING AND GIS RELIABILITY DIAGRAMS, IMAGE GENEALOGY DIAGRAMS, AND THEMATIC MAP LEGENDS TO ENHANCE COMMUNICATION

John R. Jensen
Sunil Narumalani
Department of Geography
University of South Carolina
Columbia, SC 29208, U.S.A.

ISPRS Commission VI

ABSTRACT

Students, scientists, and users viewing images or maps produced from remotely sensed imagery or GIS technology encounter a bewildering array of unstandardized, often uninterpretable map or image annotation. This research focuses on the development of improved image and thematic map annotation which is designed to enhance the reader's ability to extract information efficiently and accurately. The specific types of legend information under consideration include: 1) calibrated, gray-scale step-wedges surrounding images (and photomaps) to ensure correct exposure and visual presentation, 2) detailed image coordinate (image space) and map coordinate (map space) information including the map graticule, 3) improved thematic map legends for a) maps derived from individual dates of imagery, and b) change detection maps derived from multiple dates of imagery, 4) spatial reliability diagrams summarizing the thematic as well as geometric accuracy of sources of information used in the creation of the final image or thematic map, and 5) methods of storing and summarizing the lineage (genealogy) of each image or final thematic map. In addition, the process of developing such annotation should be facilitated by the incorporation of graphical user interfaces (GUIs) in all remote sensing and GIS software packages. It is believed that these improvements will increase the reader's ability to understand what is portrayed in a map or image map.

KEY WORDS: Remote sensing, User interface, Gray-scale step-wedges, Static/dynamic legends, Spatial reliability diagrams, Lineage, Genealogy.

INTRODUCTION

Remote sensing and geographic information systems (GIS) are rapidly growing technologies (Jordan, 1992). Applications of these technologies range from long term planning by federal, state and local governments to short term crisis management (e.g. Jensen *et al.*, 1990; Graff, 1991; Albers *et al.*, 1991; Narumalani, *et al.*, 1992). The number of users of products from these technologies is also increasing. Unfortunately, the users often have little or no experience with remote sensing and GIS, and are only exposed to the final product, which in many cases is a thematic map or image map. Poorly designed maps restrict the communication of information, or may convey false impressions (Weibel and Buttenfield, 1988). Therefore, it is essential that the remote sensing and GIS community make a concerted effort to provide thematic maps and/or image maps that are a) geometrically and thematically accurate, and b) annotated using correct cartographic principles which are conducive to easy interpretation and comprehension.

Lunetta *et al.* (1991) emphasize that decisions based on geometrically and thematically inaccurate maps and image maps increase the probability of implementing bad decisions. Many of the remote sensing and GIS products are often inaccurate because they do not meet basic *cartographic* standards. Cartographers emphasize the terms "clarity", "readability", and "aesthetics" in map design, construction, and reproduction (Crawford, 1971).

Unfortunately, remote sensing and GIS analysts often have little cartographic training. They must avail themselves of this body of knowledge to develop software and high quality output which incorporate correct cartographic principles. This paper focuses on the development of improved image and thematic map cartographic annotation which is designed to enhance the reader's ability to extract information efficiently and accurately.

BASIC CONCEPTS

There are six (6) basic considerations which, if properly implemented, can increase the probability of producing accurate and interpretable image maps or thematic maps. First, it is important that the image processing or GIS software being used has an effective user-interface which can facilitate the cartographic development of a final product. Second, there should be calibrated gray-scale step-wedges (or a color wheel or color bar) surrounding images (and photomaps) to ensure correct visual presentation on the CRT screen or 'hard copy' output. Third, the presence of image coordinate (image space) and map coordinate (map space) information is essential to make the user aware of the coordinate system to which the map has been transformed. Fourth, it is important to provide accurate and improved thematic map legends for a) static maps, (those derived from individual dates of imagery), and b) dynamic or change detection maps derived from multiple dates of imagery. Fifth, the user should be able to refer to "spatial reliability diagrams" that summarize the thematic as well as geometric accuracy of the sources used in the creation of the final map product. Finally, since it is inevitable that several iterations have been performed on the data, there is an urgent need to have methods of storing and summarizing the 'lineage' or 'genealogy' of each final image or thematic map product. It is instructive to review the nature and utility of each of these topics.

USER INTERFACE

The quality of the 'user interface' not only affects the ease-of-use of the software, but also contributes to the cartographic design of the final product (Driver and Liles, 1989). However, according to Barr (1986), and Cowen and Love (1988), effective user interfaces have not been strong points of GIS or remote sensing digital image processing systems, thus making such systems difficult to use for interactive map design.

Basically, there are two types of user interfaces: 1) command line interfaces, and 2) graphical user interfaces (GUIs). Historically, user interfaces were simple command line prompts, where input (an executable command followed by parameters) was expected from the user (Mark and Gould, 1991). An improvement over this system was the development of "fill-in forms", which prompted the user to add required parameters for executing a function. These interfaces are still evident in some of the most widely used image processing and GIS software packages. The major drawback of the command line interface is the need for a user to memorize or continuously refer to a manual in order to correctly execute a function.

The new GUIs allow more efficient and accurate human-computer interaction. The history of GUIs is chronicled by Seymore (1989) who describes their evolution from the first Xerox 'Star workstation' to Open Software Foundation's 'Motif'. GUIs use window-icon-mouse-pointer (WIMP) theory, where icons of familiar objects represent operational functions (e.g. a magnifying glass is used to zoom in). The advantage of such systems lies in the ease with which a user can initiate commands and manipulate the operating environment. In addition, 'multi-tasking' is facilitated, allowing the user to operate within two or more applications simultaneously.

Even the new WIMP user interfaces, however, suffer from the problems of overfilling the screen with icons, the creation of very long menus, the use of inappropriate metaphors, and the lack of 'activity indicators' on the status of an operation (Raper and Bundock, 1991). Also, the interface must progress and allow the user to manipulate objects that are meaningful in terms of the application, such as "sub-divide a parcel" instead of "split a polygon".

There are several new WIMP based GUIs in the marketplace including ERDAS 8.0, Arc-Info 6.0, Intergraph, and ER Mapper 3.0. Figure 1 depicts the ER Mapper 3.0 GUI for analyzing standard 1-8 band digital remote sensor data. It uses the WIMP point and click technology plus 'Macintosh' like "pull down menus" (ER Mapper, 1992). The GUI of the Spectral Image Processing System (SIPS) developed by the Center for the Study of Earth from Space (CSES) at the University of Colorado, Boulder is shown in Figure 2. This unique interface is designed to analyze "hyper-spectral" remote sensor data composed of up to 192 spectral bands. It is anticipated that such data will be common place in the Earth Observation System (EOS) era of the 21st century and will require such a user interface (Wickland, 1991). SIPS uses menus, buttons, and slider-bars along with a mouse and keyboard input to create a user-friendly interface (Kruse *et al.*, 1992). Basically, the user can move the cursor to any x, y coordinate in the scene and plot on the bottom graph the complete spectral signatures (e.g. .4-2.5 μm) for that pixel. This signature can be compared to a library of 'saved' spectra in an adjacent graph. Therefore, this system represents the first truly hyper-spectral, graphical user interface.

CALIBRATED GRAY-SCALE STEP-WEDGES COLOR-WHEELS, AND HISTOGRAM INFORMATION

One of the simplest and most useful tools for interpreting digitally processed images or image maps is the presence of calibrated gray-scale step-wedges for black and white images or a color wheel or color bar for color images. The basic function of such annotation on the screen or hard copy is to ensure correct exposure and correct visual presentation, and interactive color selection.

The concept of incorporating gray-scale step-wedges can be found in one of the earliest image processing systems -- the Video Information Communication and Retrieval (VICAR) system (Castleman, 1979). Figure 3 depicts a typical VICAR black and white 'mask' composed of systematic gray-scale step-wedges. The wedge is applied to all black and white images on the CRT and hard copy. If the exposure is correct all shades of gray level will be interpretable. However, if the image is over or underexposed, only a portion of the wedge will appear correctly and the user knows that some adjustments are required. Also present is the histogram which can be useful for communicating 'before' and 'after' image enhancement operations.

When working with color images and image maps, there are standardized color specifications which can be used to depict the exact nature of the colors used. For example, one digital image processing system allows the user to easily switch between any of the following color specification systems (Duotone, Indexed, RGB, CMYK, HSL, HSB, Multichannel) which may be displayed in a color wheel diagram (Adobe Systems Inc., 1991). However, the 'best' color wheel legend has yet to be determined.

IMAGE AND/OR MAP COORDINATES & ANNOTATION

Maps and image maps must be geo-referenced to a standard coordinate system and map projection to be truly useful. It is quite common for a final image map to be composed of data from various remote sensing systems (e.g. SPOT, Landsat TM merge) or for a final GIS map to be the product of data from very diverse source materials. Therefore, it is necessary for the data to be transformed to a single coordinate system, most commonly the Universal Transverse Mercator (UTM) or the State Plane Coordinate System (in the U.S.). The process involves the application of one of several rectification algorithms to transform the image to standardized planimetric basemap (Jensen, 1986). Once rectified, the image file contains both image coordinates (row and column) and map coordinates (e.g. UTM), and can be merged with other similarly geo-referenced GIS data. Figure 3 has an 'image space' grid superimposed on it. In addition, users of remotely sensed or GIS data must be provided with maps containing accurate map graticules whenever possible.

Another very important annotation which is often overlooked is the 'sector location diagram'. When a map sheet being displayed (e.g. sheet 3E) is but one of several other sheets in a region, a location diagram will allow the user to correctly identify which map or map image is currently being studied. Additional research on the design of these diagrams is required.

LEGENDS FOR STATIC AND DYNAMIC MAPS

Cartographers and remote sensing professionals have become quite adept at creating 'static' thematic map legends which depict the condition of the earth at a static instant in time. Ideally, good cartographic practice is followed using a relatively small number of classes (e.g. <16), logical class intervals, and cartographically correct use of colors or shades of gray. Plumb (1988) suggests that the class intervals for static maps be selected using a 'goodness of fit' index which will more accurately depict the data.

Many new products are based on the analysis of multiple dates of imagery. These 'dynamic' maps are very powerful but require new legends in order to communicate effectively. The ubiquitous change detection map is a good example of a 'dynamic' map. Monmonier (1992) suggests that an animated sequence of maps and their related statistical graphics could be used to study these "spatial-temporal" data. These methods would be useful for maps shown on CRTs, however, for 'hard-copy' output there is a need for more carefully designed dynamic map legends which depict change. New legends are required which depict the "from-to" information more efficiently and accurately.

SPATIAL (GEOMETRIC AND THEMATIC) RELIABILITY DIAGRAMS

Cartographers often use manually drawn 'reliability diagrams' to communicate the geometric and thematic reliability of their products and the source materials used (Robinson et al., 1984). This tradition should be continued in products derived from remote sensing and GIS technologies. Information on source material used and the accuracy of the material should be represented by digital *geometric* and *thematic* reliability diagrams (Lunetta *et al.*, 1991).

A geometric spatial reliability diagram should indicate the sources from which the final thematic map was compiled and which parts of the data can be considered reliable based on an established accuracy standard (e.g. National Map Accuracy Standards). For example, Figure 4 depicts a geometric reliability diagram where a thematic map has been compiled from SPOT panchromatic data, from USGS digital line graph (DLG) transportation data, and a USGS digital elevation model (DEM) containing "good" and "bad" data. It is evident that the geometric reliability of such data sources is clearly stated in the legend. The legend also identifies that the DLG vector data were converted to raster format and resampled to 10 x 10 m. Additional information such as the root mean square error (RMSE) associated with the resampling procedures of each data set can also be included. This type of annotation helps readers identify portions of the final thematic map which have reduced geometric reliability and can be useful for improved decision making. It need not be present on the map, but should be easily accessible on the system by the user.

Most modern mapping applications utilize thematic data obtained on different dates and/or at different minimum mapping units. Although a final map may look uniform in its accuracy, it is actually an assemblage of thematic information from diverse sources which vary in accuracy. Newcomer and Szajgin (1984) and Walsh *et al.* (1987) suggest that the highest accuracy of any GIS output product is only as accurate as the least accurate file used in its creation. It is important for the reader to know the source of the error by depicting them in a thematic reliability diagram. The thematic reliability diagram shown in Figure 5 identifies two sources of data used in a supervised classification of wetlands and the location of *in situ* samples used to assess map accuracy. Scientists who map wetlands might be concerned that only DLG wetland data were used. Also, the diagram reveals that the *in situ* sampling was spatially biased toward locations which were accessible only by boat. These two facts can help a reader to determine the value of a thematic map product derived from the application of various techniques.

When developing digital *geometric* and *thematic* reliability diagrams, there is a need to standardize their design and function. The most common questions pertain to the information content and the amount of detail presented on such diagrams. First, the diagrams should contain information on the data source (e.g. USGS 1:24,000 topographic quad). Second, the date of the original compilation of source data and the dates of subsequent updates should be included. Third, details on the spatial resolution to which the data may have been resampled (e.g. 10 x 10 m resampling of a Landsat TM scene) should be clearly stated. Fourth, the reliability diagrams must indicate the areas which would be considered "bad" data, or more specifically data that do not conform to some accepted accuracy standards. Fifth, if *in situ* data is used, then the bias or limitations in the acquisition of such measurements, such as the number of sample points used or the restricted access to parts of a study area should be shown.

By including this information in geometric and thematic reliability diagrams a reader is made aware of the overall accuracy of the final map. It will also limit the liability of the producers of such maps, and increase the public confidence in the integrity of products from the remote sensing and GIS community.

LINEAGE (GENEALOGY) OF THEMATIC MAPS AND IMAGE MAPS

It is important to identify the difference between lineage and spatial reliability diagrams. Lineage documentation records the entire history of all analytical operations performed on a dataset, and its resultant products. For example, Chrisman (1983) defined "lineage" as the documentation of data sources and transformations [iterations] applied to them. Conversely, the spatial reliability diagrams previously discussed provide details on the sources used in the compilation of the 'final product'.

Remote sensing and GIS final products are produced from basic source materials. Manual "book-keeping" of the processes used for deriving the final product is cumbersome and rarely performed. There are systems which provide automated methods such as 'history' or 'audit' files to keep track of the iterations and operations performed. However, none of these methods are capable of fulfilling the informational requirements of a true 'lineage' report which itemizes the characteristics of image and cartographic sources, the topological relationships between source, intermediate and final product layers, and the transformations applied to sources to derive the output products (Lanter, 1990).

The National Committee for Digital Cartographic Data Standards (NCDCDS) proposed that lineage information be included in every 'quality report' of a digital cartographic product (NCDCDS, 1988). The committee specified five requirements for the lineage criteria, including:

a) source material from which the data were derived;
b) methods of derivation, including transformations applied;
c) if data from different distinct sources are used, such sources must be identified;
d) include reference to specific control information used, e.g. National Geodetic Reference System or if other points are used then sufficient detail must be provided to allow recovery; and
e) description of the mathematical transformations of coordinates used in each step from source material to final product.

Lanter (1991) categorized geographic data layers into source layers, intermediate layers, and product layers. Lineage information on source layers should include the NCDCDS digital cartographic data standards, while intermediate layers require documentation on the nature of the transformations used in their derivation. Final product layers must be associated with information concerning their use, such as the users' role in decision making, release dates, and those responsible for product-layer maintenance (Lanter, 1990).

Lineage or genealogical documentation should, therefore, form an integral part of the annotation of remote sensing or GIS products. Software designed to document lineage must have the following components: 1) lineage tracing, 2) maintain data quality information, 3) automatic error detection, 4) rule building (i.e. flexibility to users on building their own rules into a knowledge base about how their GIS data should be handled), 5) data-driven user interface, and 6) project management (such as keeping track of times, dates, and user names to show who did what to the database and when) (Lanter, 1989). This will resolve data management problems by maintaining an automated, dynamic model of the database. In addition, the user will have information on cartographic materials used and a chronicle of the remote sensing or GIS transformations applied to derive the final products. In most cases it may only be necessary to explicitly state in a textual legend a) the name of the lineage file, e.g. Jensen.21092, and b) the cognizant scientist (and his/her address) who was responsible for creating the final product. The lineage file must then accompany the final product file.

CONCLUSION

Remote sensing and GIS products will be cartographically enhanced by adopting the five types of annotation discussed in this paper. It is also important for the image processing and GIS software to have a user interface which can facilitate the design of such maps. A schematic is presented for an "ideal" thematic or image map product (Figure 6). Such maps will benefit the user or the final decision maker, and increase the level of public confidence in remote sensing and GIS products. The suggestions presented here are not difficult to implement. However, since most computer programmers are not aware of the cartographic principles that are essential for good quality maps, it is up to the scientists and educators who are involved with the everyday use of image processing and GIS software to create an awareness for the incorporation of such annotation.

REFERENCES

Albers, B., J. Dobbin, and K. Buja, 1991. Planning in Paradise. Geo Info Systems, 1(10):26-38.

Barr, R., 1986. Automated Cartography and Geographical Information Systems. In: Advances in Computer Graphics II, Edited by Hopgood, F. R. A., Hubbold, R. J., and D. A. Duce, pp. 29-53.

Castleman, K. R., 1979. Digital Image Processing, Prentice-Hall, Inc., Englewood Cliffs, NJ, 429 p.

Chrisman, N. R., 1983. The Role of Quality Information in the Long-Term Functioning of a Geographic Information System. Proceedings, Sixth International Symposium on Automated Cartography, 1:303-312.

Cowen, D. J., and S. R. Love, 1988. A Hypercard Based Workstation for a Distributed GIS Network. Proceedings, GIS/LIS 1988, ACSM, 1:285-294.

Crawford, P. V., 1971. Perception of Grey Tone Symbols. Annals, Association of American Geographers, 61(4): 721-735.

Driver, B., and W. Liles, 1983. A Communication Model for the Design of a Computer Assisted Cartographic System. Proceedings, Fifth International Symposium on Cartography and Computing, ACSM, Falls Church, VA, 267-274.

Nixon, S., 1992. Personal Communication, May 12, 1992. ER Mapper, San Diego, CA.

Graff, S. 1991. The Political Process of Redistricting with GIS. GIS World, 4(4):64-70.

Hoffman, R. R., and J. Conway, 1989. Psychological Factors in Remote Sensing: A Review of Some Recent Research. GeoCarto International, 4(4): 3-21.

Jensen, J. R., 1986. Digital Image Processing: A Remote Sensing Perspective, Prentice-Hall, Englewood Cliffs, NJ, 286 p.

Jensen, J. R., E. W. Holmes, J. M. Michel, J. E. Savitsky, and B. A. Davis, 1990. Environmental Sensitivity Index (ESI) Mapping for Oil Spills Using Remote Sensing and Geographic Information System Technology. International Journal of Geographical Information Systems, 4(2):181-201.

Jordan, L. E., 1992. Image Processing - Map of the Future. Geo Info Systems, 2(5):75-77.

Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., and A. F. H. Goetz, 1992. The Spectral Image Processing System (SIPS) - Interactive Visualization and Analysis of Imaging Spectrometer Data. International Space Year Conference, Earth and Space Science Information Systems, Pasadena, CA, 10 p.

Lanter, D. P., 1989. Techniques and Method of Spatial Database Lineage Tracking, Ph.D. Dissertation, University of South Carolina.

Lanter, D. P., 1990. Lineage in GIS: The Problem and a Solution. Technical Paper, National Center for Geographic Information and Analysis (NCGIA) #90-6, 1-16.

Lanter, D. P., 1991. Design of a Lineage-Based Meta-Database for GIS. Cartography and Geographic Information Systems, 18(4):255-261.

Lunetta, R. S., Congalton, R. G., Fenstermaker, L. K., Jensen, J. R., McGwike, K. C., and L. R. Tinney, 1991. Remote Sensing and Geographic Information System Data Integration: Error Sources and Research Issues. Photogrammetric Engineering and Remote Sensing, 57(6):677-687.

Mark, D. M., and M. D. Gould, 1991. Interacting with Geographic Information: A Commentary. Photogrammetric Engineering and Remote Sensing, 57(11):1427-1430.

Monmonier, M., 1992. Summary Graphics for Integrated Visualization in Dynamic Cartography. Cartography and Geographic Information Systems, 19(1):23-36.

Narumalani, S., Jensen, J. R., Weatherbee, O., Murday, M., and W. J. Sexton, 1992. Coastal Sensitivity Mapping for Oil Spills in the United Arab Emirates Using Landsat Thematic Mapper Imagery and GIS Technology. Proceedings, American Society of Photogrammetry and Remote Sensing, Albuquerque, NM, 314-322.

NCDCDS, 1988. The Proposed Standard for Digital Cartographic Data. American Cartographer, 15(1):142 pp.

Newcomer, J. A., and J. Szajgin, 1984. Accumulation of Thematic Map Error in Digital Overlay Analysis. American Cartographer, 11(1):58-62.

Plumb, G., 1988. Displaying GIS Data Sets Using Cartographic Classification Techniques. Proceedings, GIS/LIS '88, San Antonio, TX, 1:340-349.

Raper, J. F., and M. S. Bundock, 1991. UGIX: A GIS Independent User Interface Environment. Technical Papers, ASPRS-ACSM Annual Convention, Baltimore, MD, 6:275-295.

Robinson, A. H., Sale, R. D., Morrison, J. L., and P. C. Muehrcke, 1984. Elements of Cartography, John Wiley & Sons, New York, NY, 544 p.

Seymore, J., 1989. The GUI: An Interface You Won't Outgrow. PC Magazine, 8(15):97-109.

Walsh, S. J., Lightfoot, D. R., and D. R. Butler, 1987. Recognition and Assessment of Error in Geographic Information Systems. Photogrammetric Engineering and Remote Sensing, 53(10):1423-1430.

Weibel, R., and B. P. Buttenfield, 1988. Map Design for Geographic Information Systems. Proceedings, GIS/LIS '88, San Antonio, TX, 1:350-359.

Wickland, D. E., 1991. Mission to Planet Earth: The Ecological Perspective. Ecology, 72(6): 1923-1933.

Figure 1. An example of a graphical user interface (GUI) which uses window-icon-mouse-pointer (WIMP) technology. A panchromatic aerial photograph has been scanned, rectified and overlayed with three types of GIS information (fences, contours, and traffic density) from three non-image sources (Arc-Info, Genamap, and Oracle). (Reproduced with permission of Stuart Nixon, ER Mapper, San Diego, CA).

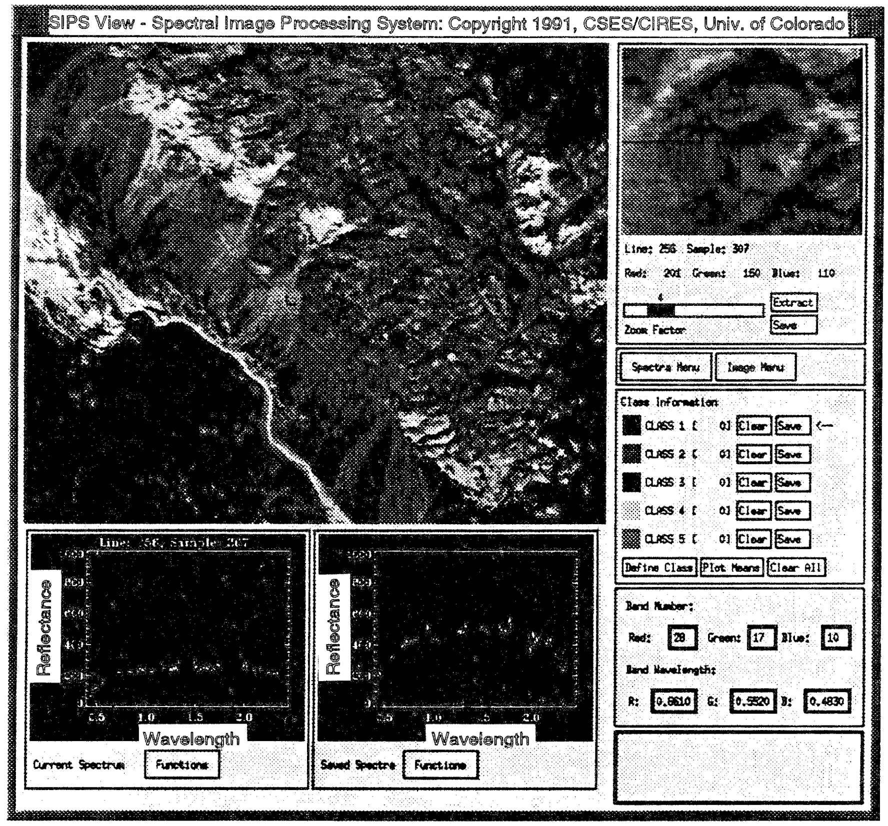

Figure 2. The Spectral Image Processing System (SIPS) graphical user interface is capable of displaying "hyper-spectral" remote sensor data (i.e. more than the standard 1-8 bands). In this example bands 28, 17 and 10 are used in the RGB color planes respectively. A comparative evaluation of the spectral signature of a pixel at an x, y location can be made with a saved spectral signature file. (Courtesy of Dr. Alexander Goetz, University of Colorado, Boulder, CO).

Figure 3. An example of a gray-scale step-wedge produced by the Video Information Communication And Retrieval (VICAR) digital image processing system. Also included is the 'image space' graticule and a histogram of the image which is so useful when performing image enhancement.

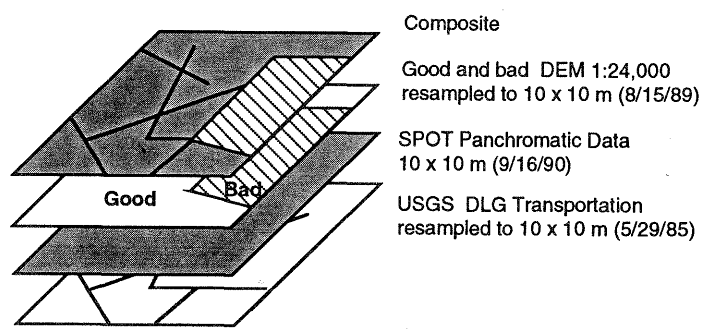

Figure 4. A geometric reliability diagram summarizing data sets and the degree of resampling.

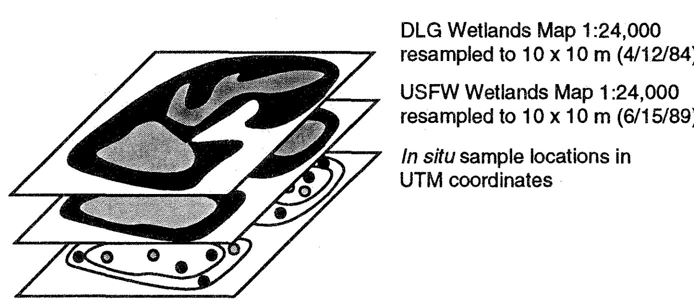

Figure 5. A thematic reliability diagram summarizing the sources used to educate a classifier and perform error evaluation.

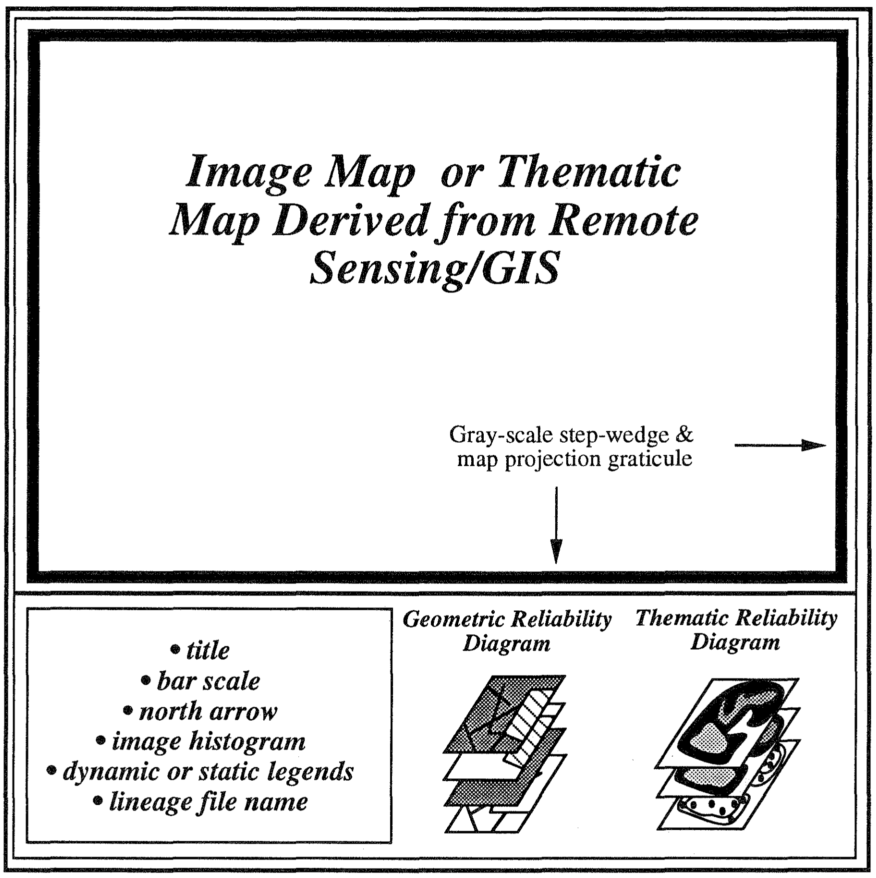

Figure 6. A schematic diagram of an "ideal" remote sensing and/or GIS derived thematic map or image map.