Efficient Handling of Large Digital Images in Geographic Information Systems

Fayez S. Shahin

Abstract
There has been a tremendous demand for handling digital images in geographic information systems (GIS) and real time mapping applications, mainly due to interest in multi-media information systems. Due to the large amount of data involved in large digital images, efficient storage, management, and processing of such images in geographic information systems presents a challenge that needs to be addressed properly. This paper presents a new methodology for handling digital images in GIS. This methodology is based on partitioning large digital images into smaller tiles and using a hierarchical image compression algorithm.

Introduction
There is a growing list of applications, such as softcopy photogrammetry, real-time mapping, geographic information systems (GIS), and remote sensing of Earth resources, which has led to a tremendous demand for digital images. Digital images can be obtained directly by using solid state cameras and satellite sensors, or indirectly by scanning existing photographs or paper maps. Despite all the advantages of digital images over film and paper maps, the problem of efficient, economic storage and flexible retrieval and manipulation of vast amounts of pictorial information has become more important and requires careful consideration. Single SPOT panchromatic scenes contain 6,000 by 6,000 pixels, or 36 Mbytes; aerial images (9 by 9 inches) are often scanned with a resolution of 20 micrometres, which results in 10,000 by 10,000 pixels, or 100 Mbytes. Digital images generally contain a significant amount of redundant information; therefore, large storage requirements can be reduced by proper image compression.

Image compression is concerned with minimizing the number of bytes required to represent an image. Storage of large quantities of image data is required for archiving aerial photographs, satellite images, digital orthophotos, and digital images collected by solid-state cameras. Other applications of image compression include time critical data processing, in which case the speed of image analysis algorithms is improved by applying those algorithms to compressed data.

For the past few years the Joint Photographic Expert Group (JPEG) has been working towards establishing an international standard for still images, both gray scale and color (Rabbani *et al.*, 1990; Wallace, 1991; Sarjakoski *et al.*, 1995).

Image compression algorithms can be classified as lossless or lossy techniques. Lossless compression reduces the number of bits required to represent an image such that the reconstructed image is numerically identical to the original on a pixel-by-pixel basis. On the other hand, lossy compression schemes allow degradation in the reconstructed image in exchange for a reduced bit rate. These degradations may or may not be visually apparent, and greater compression can be achieved by allowing more degradation.

This paper deals with image compression and presents special methods for handling digital images in GIS. Finally, experiments are presented to show the advantages and disadvantages of image compression with respect to geometric quality, loss of information, compression ratio, and compression speed.

Image Compression
This section explains basic terms and techniques used in digital image compression. It deals with the definition of entropy and redundancy, the compression ratio, and quantization. It also describes a hierarchical approach for digital image compression and the method of tiling large digital images.

Terminology Review and Definitions
The general *compression ratio* is defined as the ratio of the number of bytes of the original image before compression to the number of bytes of the compressed image (Rabbani *et al.*, 1990).

A measure of the information content of an image is the *entropy* (Lynch, 1985). *Entropy* expresses the minimum number of bits necessary for the representation of an image without any loss of information. *Entropy* is also a global measure of the correlation between gray values of neighboring pixels. For an n-bit image with M gray levels ($M = 2^n$), the *entropy* H can be computed as

$$H = -\sum p(g_i) \log_2 p(g_i) \tag{1}$$

where $p(g_i)$ is the percentage of occurrence of each gray value, which can be obtained from the image histogram.

In most real-world images, the gray values of adjacent pixels are highly correlated. This means that a large amount of redundant information is available in these images. For a given digital image, the *redundancy* R is defined as follows (Lynch, 1985):

$$R = \log_2 M - H \tag{2}$$

A large portion of the compression achieved by lossy schemes is due to *quantization*. A *quantizer* is simply a staircase function that maps a large number of input values into a smaller, finite number of output levels. The type and degree of quantization impacts the quality of a lossy scheme (Rabbani *et al.*, 1990).

An-Najah National University, Nablus, West Bank, Palestine (Fayezsh@najah.edu).

Photogrammetric Engineering & Remote Sensing, Vol. 63, No. 10, October 1997, pp. 1181–1184.

0099-1112/97/6310-1181$3.00/0

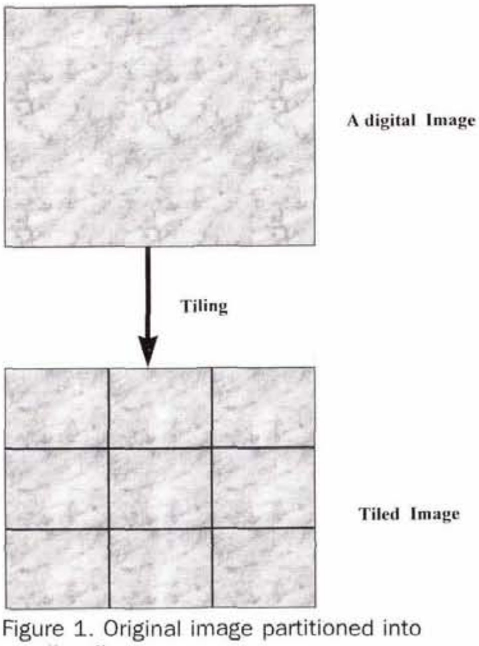

Figure 1. Original image partitioned into smaller tiles.

The Hierarchical Predictive Coding Scheme

Hierarchical Predictive Coding (HPC), as described in this paper, was developed based on a compression scheme for digital images in softcopy photogrammetric stations (Shahin, 1994). The algorithm starts by creating a scale level representation of the original image (image pyramid). It is derived by smoothing the gray values of the original image to a predefined lower resolution using a low-pass filter. For instance, a square image of 1024 by 1024 pixels is decomposed into the following five scale levels: 1024 by 1024, 512 by 512, 256 by 256, 128 by 128, and 64 by 64. The image pyramid is generated by convolving higher resolution images with a low-pass filter and reducing the pixel size by half.

Difference images are computed between all levels. These differences represent the high frequency components of the previous level of the image pyramid. The original image can be reconstructed from its representation at the coarsest level (e.g., 64 by 64 pixels), and all the difference images. The pixel values of the difference images are usually small. Therefore, they can be coded and stored with a lower number of bits. This would result in a lossless compression of the original image. However, in order to achieve higher compression ratios, the number of bits of the difference images are reduced by quantization, which leads to image degradation and lossy compression.

The HPC algorithm consists of the following steps:

(1) The image pyramid is created by smoothing the original image sequentially with a 3 by 3 Gaussian operator. This results in a scale level representation of the original image.
(2) The difference images at all levels are computed. For instance, an original image of 1024 by 1024 pixels is replaced by one 64- by 64-pixel image and four difference images.
(3) The values of the highest level image (64 by 64) and the first difference image (128/64) are stored as 8-bit integers.
(4) The values of the 256/128 difference image are stored and coded as 5-bit integers.
(5) The values of the 512/256 difference image are stored and coded as 3-bit integers.
(6) The values of the 1024/512 difference image are stored and coded as 1-bit integers.

Using this representation (HPC-8.5.3.1), a compression ratio of 3.6 can be achieved. For higher compression ratios, the difference images are represented by fewer bits. However, the degradation of the reconstructed image increases. In the experiment a compression ratio of 4.5 was reached by quantizing the difference images in the following way: 128/64: 8bit; 256/128: 3bit; 512/256: 3bit; 1024/512: 1bit. This will be referred as HPC-8.3.3.1.

Tiling Large Images

Compression and decompression of large digital images are time-consuming. Handling large digital images as a whole is rather cumbersome. Furthermore, there are many applications where only part of the image is required immediately. An approach in which the original image is partitioned into smaller compression units, called *tiles*, is very useful when handling large digital images (Figure 1). Any part of the image needed can be accessed easily and retrieved quickly through the tile system.

When using piecewise compression based on tiles, only the needed parts of the image are decompressed, and the time required to decompress a few tiles is moderate compared to decompressing the whole image. Therefore, piecewise compression is very suitable for real-time mapping systems, geographic information systems, and image information systems.

Methods for Handling Large Digital Images

Many current algorithms have been designed to effectively manage points, lines, surfaces, and attribute data. However, large spatial objects such as remotely sensed images, scanned maps, and digital elevation models are series of large arrays with various spatial, spectral, and temporal information. They often require long delayed processing to provide display, analysis, or query.

Image compression and tiling, as described in the previous section, will not solve all problems of handling large quantities of digital data. Databases, in the form of pictorial information systems, need to be developed to manage digital images in a GIS, and make them easily and quickly accessible for the user. In a GIS, photogrammetric images must be tied to a geographic coordinate system and should contain the full orientation information (interior and exterior orientation parameters), or they should be available as geocoded or ortho-rectified scenes which can be directly used as layers of a GIS.

The approaches that are used to handle large digital images in conjunction with image compression and tiling are data abstraction and indexing and spatial reasoning.

Data Indexing and Abstraction

Abstraction of data is the process by which new concepts can be derived from a given collection of concepts and is a method of generating indexes. Abstraction operations to generate indexes can be described as follows (Shi-Kuo, 1989; Zhou, 1994):

(1) A set of picture objects are given a common label (labeling). This can be accomplished by either (a) identifying common attributes with identical values or (b) extending the objects' attributes by a common labeling attribute.
(2) The picture objects having the same label are clustered together to form a new picture object class (clustering), with the common label as its type.
(3) A new relational object, called the index object, is created, which contains (a) a type which is the type of the picture object class created in step 2 and (b) other attributes extracted from the attributes of the picture objects in the same picture object class.

The abstraction operation can be applied recursively, so that the index objects can again be labeled, clustered, and in-

Table 1. Entropies and Redundancies of the Three Test Images (units: bits/pixel)

[bits/Pixel]	Aerial Image	Road Image	Target Image
Entropy	5.79	3.36	2.75
Redundancy	2.21	4.64	5.25

dexed. Using this abstraction and indexing procedure, answers to a query may be tiny and may be accessed very quickly. Therefore, the derived data can be stored in a GIS database for various applications. A major application of such data abstraction and indexing is the classification of multispectral images into several classes such as water, forest, land use, and vegetation. Another important application is the use of feature extraction algorithms to extract roads, streams, trees, or buildings from digital images.

Spatial Reasoning

Spatial reasoning is the process of inferring a consistent set of spatial relationships among the objects in a picture (Shi-Kuo, 1989). When retrieving pictures from a GIS, one of the most powerful methods for discriminating among the pictures is the perception of the relationships that exist between objects in the desired picture. This provides the ability to make queries on these relationships an important attribute of a GIS. The user of a GIS may ask the system such questions as "What objects are located between a river and the forest in this picture?" or "Is there a gas station west of the road in the picture?" This type of query might be of interest in cases where a large number of pictures are archived at a central site. By examining these relations, users at remote sites can retrieve those pictures of interest while operating on a small fraction of the information stored at the central site.

Experiments and Results

Both JPEG (baseline system) and HPC image compression and decompression software were implemented on a Sun SPARC station 2 to investigate the advantages and disadvantages of the two techniques. The quality of the reconstructed images after compression and decompression was evaluated by the following measures: (1) visual loss of information, (2) compression ratio, and (3) geometric and radiometric distortions. The algorithms were tried out on different levels of entropy and should, therefore, show the influence of image content on the compression results. The following three images were used in the tests:

(1) An 8-bit aerial image of the Chaco Canyon area. The size of the image is 1024 by 1024 pixels.
(2) An 8-bit road image captured by a digital camera mounted on a van. The size of the image is 1024 by 1280 pixels.
(3) An 8-bit target image captured during the field-test of a digital camera. The size of the image is 1024 by 1280 pixels.

Table 1 shows the entropy and redundancy of the three test images. The aerial image has the largest entropy, while the target image has the smallest entropy. Thus, the target image contains the most redundant information, while the aerial image contains the lowest amount of redundant information.

Table 2. Compression Ratios Dependent on Image Type and Quantization Value

Compression Type	Aerial Image	Road Image	Target Image
JPEG-100 (lossless)	1.3	2.4	2.9
JPEG-75	5.9	33.3	43.5
JPEG-40	12.5	40.0	62.5
HPC-8.5.3.1	3.6	3.6	3.6
HPC-8.3.3.1	4.5	4.5	4.5

Table 3. Radiometric Distortions of the Reconstructed Images (units: gray values)

Compression Type	Aerial Image	Road Image	Target Image
JPEG-100 (lossless)	3.06	1.52	0.54
JPEG-75	4.48	1.64	0.59
JPEG-40	9.8	2.36	1.98
HPC-8.5.3.1	1.81	1.78	1.79
HPC-8.3.3.1	1.89	1.80	1.80

The three images were compressed using the JPEG algorithm at different quantization levels and the HPC technique using two different quantization thresholds. Table 2 shows the resulting compression ratios.

These results clearly show that a large amount of compression can be achieved when using lossy compression schemes. The compression ratio obtained by using lossless schemes is limited by the entropy of the image. For example, in the lossless compression case of the aerial image, the compression ratio is 1.3, which is approximately 8 divided by the entropy of the image (5.79 bits per pixel).

It is also of great importance to notice the difference in compression ratios for different images at the same quantization value. For example, at JPEG quantization level 75 the aerial image was compressed only 5.9 times, while the target image was compressed approximately 43 times. This huge difference is due to the large amount of redundancy available in the target image and the large information content in the aerial image. As the HPC scheme is not based on transform coding, the compression ratio is independent of image content.

Quality of the Reconstructed Image

The reconstructed image, after being compressed by lossless schemes, is identical to the original image on a pixel-by-pixel basis. However, there is degradation in the reconstructed image from lossy schemes in exchange for greater compression as compared to lossless techniques. This degradation in the reconstructed image will eventually lead to radiometric and geometric distortions.

For investigating the radiometric distortions, the gray values of the reconstructed images were compared with the original image on a pixel-by-pixel basis. Table 3 presents the radiometric distortions of the images at different quantization levels.

These results show that the radiometric distortions depend to a limited degree on the level of quantization and, therefore, increase with the compression ratio. Thus, the radiometric quality of the reconstructed image deteriorates with larger compression. This is especially obvious for the JPEG algorithm; a block pattern appears in the reconstructed images if the quantization level was selected too low. The HPC scheme, on the other hand, causes a blur of the reconstructed image.

The geometric distortions are the most important because they change the locations of image points and, consequently, influence the coordinates and computations obtained by photogrammetric procedures. For the purpose of investigating geometric distortions of the reconstructed images, 2500 points were defined in the original images. In the reconstructed images after compression, the positions of these points were computed using the very accurate least-squares matching method. The differences between the coordinates in the original image and the computed coordinates in the reconstructed image indicate geometric distortions. Table 4 presents the geometric distortions in pixels for different quantization levels.

Table 4. Geometric Distortions of the Reconstructed Aerial Image (units: pixel)

Distortions	JPEG-75	JPEG-40	JPEG-5	HPC-8.5.3.1	HPC-8.3.3.1
Maximum x	0.292	2.654	17.036	0.308	1.303
Maximum y	0.353	0.815	12.972	0.671	2.075
RMSE in x	0.038	0.107	0.786	0.036	0.056
RMSE in y	0.042	0.102	0.807	0.031	0.07

By analyzing Table 4, one can conclude that geometric distortions for compression ratios below 5 are small and can be neglected. Thus, it is safe to use these images for precise photogrammetric analysis and point determination. Reconstructed images from compression levels above 10 can be used for tasks that do not require precise point determination, such as extraction and interpretation of features.

Computing Time

The time required to compress and decompress a 1024 by 1024 image on a Sun SPARC 2 station using the JPEG algorithm is about 20 seconds. The HPC algorithm needs 7 seconds for the same image. The JPEG technique is slower because it employs the discrete cosine transformation, which is very time consuming. The HPC scheme is faster because it is mostly based on convolutions. Therefore, it can be easily implemented on an array processor to perform real-time compression for digital image collection.

Conclusion

Image compression is an important topic in real-time mapping and GIS, because very large digital images are commonly used in these applications. Different compression schemes are already available in hardware and, as such, have become an integral part of real-time mapping and GIS applications. Most commercial compression hardware is based on the JPEG algorithm, which is the current industry standard for image compression.

Two compression schemes were used to evaluate the compression effects on digital images. The standard JPEG compression scheme was compared to the hierarchical predictive coding algorithm. The HPC technique is faster and maintains a good geometric and radiometric image quality. However, only small compression ratios were achieved. The JPEG technique, on the other hand, is rather slow if implemented in software, but it achieves high compression ratios. Both radiometric and geometric distortions were evaluated, and it was found that, for compression ratios smaller than 5, these distortions are smaller than 0.1 of a pixel and can be neglected. Larger compression ratios may distort the images considerably and are not recommended for applications that require precise spatial analysis. However, the reader should keep in mind that compression ratios obtained by JPEG depend heavily on the image content (entropy) and, therefore, larger compression ratios may be adequate for some applications where precise spatial analysis is not required.

Image compression alone, however, will not solve all problems of handling large quantities of digital data. Databases, in the form of pictorial information systems, need to be developed to manage digital images in a GIS, and to make them easily and quickly accessible for the user. In a GIS, images must be tied to a geographic coordinate system and should contain the full orientation information, or they should be available as orthophotos which can be directly used as layers of a GIS.

References

Lynch, T., 1985. *Data Compression Techniques and Applications*, Lifetime Learning Publications, Belmont, California.

Rabbani, M., and P. Jones, 1990. *Digital Image Compression Techniques*, SPIE Optical Engineering Press, Bellingham, Washington.

Sarjakoski, T., and J. Lammi, 1995. Image Compression by the JPEG Algorithm, *Photogrammetric Engineering & Remote Sensing*, 61(10):1261–1266.

Shahin, F.S., 1994. A Hierarchical Approach for Digital Image Compression, *Proceedings of the Annual ASPRS Convention*, 1:570–578.

Shi-Kuo, C., 1989. *Principles of Pictorial Information Systems Design*, Prentice Hall, Englewood Cliffs, New Jersey.

Wallace, G.K., 1991. The JPEG Still Picture Compression Standard, *Communications of the ACM*, 34(4):30–44.

Zhou, W., 1994. Large Spatial Object Handling in Geographic Information Systems, *International Archives of Photogrammetry and Remote Sensing*, 30(Part 2):212–219.