Compression Techniques for Distributed use of 2D and 3D Data on Internet

By

P.Lalitha

Indus Biotherapeutics Ltd,

INTAS Pharmaceuticals

Plot No. 423 P/A/GIDC

Sarkhej-Bavla Highway

Ahmedabad – 382 210

E-mail: *librarian_biotech@intaspharma.com*

T.A.V. Murthy

Director

INFLIBNET Centre

Gujarat University Campus, Post Box No. 4116

Navrangpura

Ahmedabad – 380 009

E-mail: *tav@inflibnet.ac.in*

ABSTRACT

Internet has become a forum for human communication in a wide variety of disciplines. The 2D and 3D data occupies large chunks of storage space, when converted from their natural analog form to digitized form. Hence there is a need to compress the data for storage space, fast processing and transmission. The present paper gives a view on need for compression of 2D and 3D data, different compression techniques. It also emphasizes on user query process.

KEY WORDS: Internet, 2D and 3D data, compression techniques, query process, artifacts.

0. INTRODUCTION

Present era of 21st century is flooded with abundant information across the world. The attitude of the people is rapidly changing towards automation and modernization in seeking information in a shortest possible way. At this juncture, **Internet** plays a vital role as it has resulted in an unparalleled degree of communication, collaboration, resource sharing and information access. It has become a forum for human communication in a wide variety of disciplines ranging form computers, sciences and social sciences to art, music and other recreations. The availability of information ranges from simple text to documents to images, graphics, audio and multimedia information. The most complex feature to be added up to the textual information is multimedia based resources like manuscripts, photographs, multidimensional images, paintings, artifacts etc., When the textual information supplemented with pictures, sounds, animation etc., it makes more meaningful and understandable. But due to this series of obstacles occur; prominent among them is digital obesity. Audio, graphics and animation etc., when converted from their natural analog form to digitized form for manipulation and retrieval,

shares large chunks of storage space [2]. Hence, it is necessary to compress the data for the optimization of storage space, fast processing and transmission over the network.

1. NEED FOR COMPRESSION OF 2D AND 3D DATA

Obtaining 2D and 3D data on networks have a wide range of application in a variety of scientific, engineering, industrial, and even in cultural disciplines. The 2D data comprises manuscripts, photographs etc., and the 3D data comprises natural and sculptured objects, architectural and heritage designs, engineering models etc., Some times the same item may be stored in several digital formats. These formats are exactly equivalent and it is possible to convert from one to the other. (e.g., an uncompressed image and the same image stored with a loss-less compression). The availability of improved software, scanners and hardware helped the accessibility of 2D and 3D data on networks. The compression is the process of reducing the file size by abbreviating the repetitive information. The compression can be done by a set of compression or mathematical algorithms. There are two types of compressions.

1.1 Loss less Compression

In this process of compression no information is lost. In this method repeated information can be converted into a mathematical algorithm that decompresses with out loosing any details of the original file.

1.2 Lossy compression

Lossy compressions are not the exact replica of the original file. In this type the compression discards or averages details that are least significant. It will not have a distortable effect on the quality of image.

2. COMPRESSION PROTOCOLS FOR 2D

The following are the commonly used protocols for bitonal, gray scale or coloured compressions.

2.1 ITU-G4

This standard is developed by the International Telecommunication Union (ITU), which is considered as the de facto standard compression scheme for storing black & white bitonal images.

An image created as a Tagged Information File Format (TIFF) and compressed using ITU-G4 compression technique is called a Group-4 TIFF or TIFF G-4 [1].It is a lossless compression scheme. Joint Bi-level Image Group (JBIG) (ISO – 11544) is another standard compression technique for bitonal images. TIFF is used to exchange files between applications and computer platform. It is a flexible bitmap image format supported by virtually all paint, image-editing, and page-layout applications [5].

2.2 Joint Photographic Expert Group (JPEG)

It is an ISO – 10918-I compression protocol that works by finding areas of image that have same tone, shade, colour or other features representing this area by a code. It is a lossy compression.

2.3 LZW (Lenpel-Ziv Welch)

This technique uses a table based lookup algorithm invented by Abraham Lenpel, Jacob Ziv, and Terry Welch. Two commonly used file formats in which LZW compression is used are the graphic interchange format (GIF) and the Tag Image File Format (TIFF). This compression is also suitable for compressing text files. A particular LZW compression algorithm takes each input sequence of binary digit of a given length (e.g 12 bits) and creates an entry in a table (sometimes called a dictionary or code book) for that particular bit pattern consisting of the pattern itself and a shorter code. When images are compressed, TIFF file size but increases the time required to open and save the file.

3. COMPRESSION OF 3D MODELS

Polygon Mesh Model

The one of the recent compression techniques used for the 3D data compression. It is based on Geometric compression. The strategies for 3D polygon mesh models have three components.

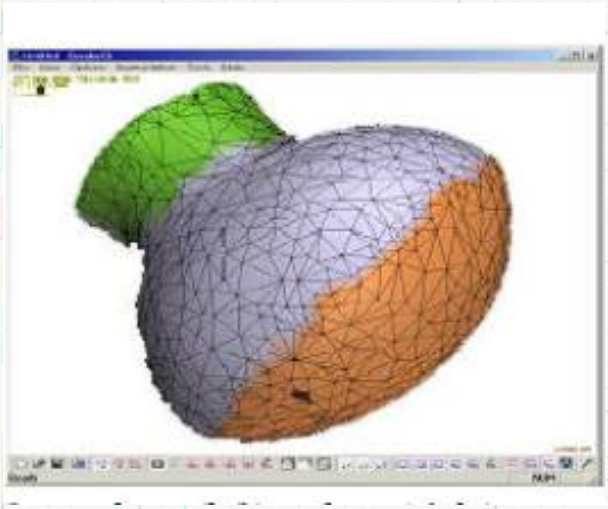
a) Compression of connectivity information

b) Compression of geometric information

c) Compression of attributes

a) Compression of connectivity information

In a triangle mesh, the number of triangles in a mesh is roughly twice the number of vertices and each vertex is referenced in 5 to 7 triangles, which indicate the scope for reducing redundancies in the description of connectivity among the vertices. Connectivity compression techniques encode triangle /polygon connectivity is lossless manner. High compression ratios have been achieved for connectivity encoding, typically a few bits per vertex.



Triangulated surface meshes from a pot segmented using the watershed algorithm based on absolute curvature [6].

b) Geometric Data Compression

Quantization of coordinates and colour values leads to an indiscriminate truncation of the accuracy regardless of the features in the geometric model. The Spectral Compression effort of Karni and Gotsman and Wavelet based technique of Khodakovsky et.al. are examples of this approach.

Wavelet compression

It is the method of mathematical modeling of images that breaks through the image down into small waves of that represents the frequency analysis of a function. The shapes patterns in an image are identified and then described using mathematical functions or formulae.

c) Compression of Attributes

Attributes such as colour, normal, texture mapping coordinates and material properties defined at vertices, faces, or corners of faces are compressed using specialized encoding schemes. The compression of this information involves compact encoding of two kinds of information: The property values and property mapping.

Beyond this type of compression a lot of research is taking place and other new techniques have been researched and research is still going on. Some of the new techniques are:

i) Advancing Fan Front (AFF)

It is an algorithm developed using a group of triangles, a fan as the unit of encoding. The repeating feature detection based compression technique takes this even further, by using components and even groups of components as the units of encoding. It is a connectivity type of compression [4].

ii) Compression using Automatic Discovery of Repeating Features

Its innovativeness lies in the fact that it is based on automatic detection of repeating features. It is applicable to large assemblies of components and is ideally suited for compressing engineering type models [4].

3. QUERY PROCESS

There need to be a very strong submission of queries to the database, which reaches the server. Excellent Java programmes matched up with Dublin Core cataloguing are available for the normal text material. However in the case of the visual databases application, programmes need to be developed in Java or Visual ++ for addressing the queries of the client in relation to the data stored. It is also expected that the user can select on of the thumbnail mechanism along with the descriptive data. Research is still grappling with how to bookmark or store searches, which result to create a traceable multiple searches [3]. Interactions such as queries require that information in digital library be organized effectively. The scanned digital material is frequently related to other materials by relationships such as part or whole sequence etc.

A single computer programmed is assembled from many files both source and binary with complex rules of inclusion. These items may be stored in several digital formats, which are some times exactly equivalent to convert from one to another. The manner in which the user wishes to access material may depend upon the characteristics of computer systems and network and the size of the material. The information architecture should consist of general approach to organizing the material with in the digital library in such a manner that the computers can understand the structure of the material and carry out the interactions that the user wishes. Normally the information architecture is motivated by the principles of application programmes with flexibility. Further it must recognize that information is valuable subject to terms and conditions and is transmitted over insecure networks that cross national boundaries. These factors underline the framework of the architecture. The information architecture is based on three simple concepts like data types, structural Meta data and Meta objects [7].

4. CONCLUSION

Extensive research is being carried out on 3D and 2D images using computers for building realistic models at Arizona State University. The experiment carried over there strictly confines to Native Americans artifacts in particular anthropological artifacts. India being one of the ancient civilizations is endowed with abundant wealth of artifacts ranging from uni-dimensional to poly and hectagonal objects of different make colour belonging to the different periods etc., and this complexity would provide a challenge to the researchers who are attempting to capture, maintain accurate data in the form of digitization i.e., adopting compression technique. The further challenge that lies ahead is to embark upon web based interactive models across platforms, ultrasound data linking up of these images with the text etc., These will also encourage artists, historians, computer specialist to enter into new vistas of comparative studies. This may even lead to discovering of new tools and techniques while highlighting the imitation of the artifacts loss of access etc., The techniques would also inbuilt scalable properties as a proof of concept which are inconformity with Dublin Core cataloguing structure.

REFERENCES

[1] Arora, Jagdish. “Building Digital Libraries: An Overview”, *DESIDOC Bulletin of Information Technology*, 21(6), (November 2001), pp-3-24.

[2] Padhi, Biswaranjan. “Data Compression Technologies in Multimedia CDs.” In Access to Electronic Information; edited by Mahapatra, M and Ramesh, D.B, 21-24.Bhubaneswar Chapter: SIS.

[3] Rowe Jeremy, Razdan Anshuman, Collins D and Panchanathan S. “A 3D Digital Library System: Capture Analysis, Query, and Display.” In *Digital Libraries; Dynamic Landscapes for Knowledge Creation, Access and Management*, edited by Urs, Shalini.R, Rajasekhar, T.B. and Raghavan, K.S, 149-159. Bangalore: University of Mysore and IIIT, 2001.

[4] Shikare Dinesh, Venkata Babji S, and Mudur S.P. “Compression Techniques for Distributed Use of 3D Data- an emerging media type on the Internet.” In *Proceedings of the ICCC 2002*, edited by Raghavan, S.V. and Mudur, S.P, 679-696. Mumbai: NCST, 2002.

[5] Xplora Design skool course material

[6] Schurmans, Utsav. “Advances in Geometric Modeling and Feature Extraction on Pots, Rocks and Bones for Representation and Query via the Internet.” 2001.

<http://3dk.asu.edu/archives/publications/g3dk/caa2001.pdf. (10 November 2002)

[7] Arms, William Y, Blanchi, Christophe and Overly, Edward A. “An Architecture for Information in Digital Libraries”, [*D-LibMagazine]*, February 1997; available from

<http://www.dlib.org/dlib/february97/cnri/02arms1.html>

BRIEF BIOGRAPHY OF AUTHORS

P. Lalitha is Information Resource Officer at Indus Biotherapeutics Pvt Ltd., Ahmedabad She holds M L I Sc and registered for Ph.D at Bundelkhand University in LIS. Her area of interest is digitization of artifacts related to Indian Heritage and Culture

Dr. T.A.V. Murthy is currently the Director of INFLIBNET and President of SIS. He holds B Sc, M L I Sc, M S L S (USA) and Ph.D. He carries with him a rich experience and expertise of having worked in managerial level at a number of libraries in many prestigious institutions in India including National Library, IGNCA, IARI, Univ of Hyderabad, ASC, CIEFL etc and Catholic Univ and Casewestern Reserve Univ in USA. His highly noticeable contributions include KALANIDHI at IGNCA, Digital Laboratory at CIEFL etc. He has been associated with number of universities in the country and has guided number of Ph.Ds and actively associated with the national and international professional associations, expert committees and has published good number of research papers. He visited several countries and organized several national and international conferences and programmes