

Multiple Browsing Levels in Digital Libraries

Bala Srinivasan*, Santosh Kulkarni*
PD Le*

Abstract

Multimedia digital libraries involve different types of data objects such as text, audio, video and images. Most of these data objects are very large in size and accessing them in a distributed environment causes a transmission delay due to the vast amount of network traffic. Compressing these data objects before transmission, can reduce the response time, although it would mean a reduction in the quality of the output data. If the application doesn't demand a high quality output, data compression can be an acceptable means of reducing transmission time over the internet or other distributed environments. The loss of quality would be proportional to the amount of compression applied to individual data objects. Therefore different quality levels (browsing levels) can be achieved depending on how the data will be used. A lower quality level could be used for general browsing of data whereas a higher quality level could be used where the output data has to be further processed and analysed. The behaviour of images when compressed using various compression techniques was studied and it was observed that different images require different amounts of compression to reach the same quality level. This result allowed us to classify images into different classes, based on their compression behaviour. In this paper, we identify a set of rules to calculate a near optimal compression ratio to achieve a given level of image quality. We also explain how this set of rules can be incorporated into digital libraries with multiple levels of browsing, to achieve a faster response time.

1. EVOLUTION OF DIGITAL LIBRARIES

Digital libraries have been in use for the last three decades. Over this time there has been a constant improvement in the technology used by digital libraries. In the 1970's, digital libraries were based on minicomputers and they provided basic services of remote access and online search and retrieval. By the 1980's,

information science had progressed to the point where sophisticated information storage and retrieval systems were in operation. Digital libraries made use of techniques to share bibliographic records and link different remote systems.

In the 1990's, there has been a revolution in digital library systems. Today's technology has made it possible for digital libraries to include *different media objects such as text, image, audio and video*. In particular, visual information systems are getting more popular as compared to text based information systems. Therefore,

*Faculty of Information Technology,
Monash University, Australia.
email: srini@monash.edu.au

digital libraries are becoming more graphical in nature. In addition, the concept of hypertext has been introduced in digital libraries. Traditional library systems cannot handle backward tracing of references. Following a referential trail is also a problem with traditional on line systems. Hypertext technology solves these problems and makes it more convenient to access digital libraries. Linking references makes it easy to follow them forwards and backwards. It is therefore possible to structure and organise the same material in a variety of ways, so that it can serve multiple functions.

2. DIGITAL LIBRARIES AND MULTIMEDIA

The growing use of multimedia data has introduced new challenges in storage and communication of diverse multimedia objects like video, audio and images. Internet applications such as video conferencing, video on demand and GIS require communication of image and video data in a distributed environment. The World Wide Web (WWW) makes it easier to transfer such information over national and international networks. Image and video data are very large in size, of the order of gigabytes. As the internet applications mentioned above become more popular and the number of users on the WWW increase, network bandwidth will be a serious problem. This could lead to an increase in the response time as the communication time goes up. Considerable research has been conducted to increase the network bandwidth by developing new technologies such as ATM, FDDI and DQDB.

Even though these advanced technologies offer a very high bandwidth, there is still a need to find alternate ways to reduce the network traffic. If image and video data are compressed before transmission, the transmission time can be reduced substantially. The next section explains the need for image compression and introduces some of the techniques used.

3. IMAGE COMPRESSION

According to Rabbini¹⁶, three types of redundancies can be identified in digital images.

- ⊛ Spatial redundancy, which is due to the correlation between neighbouring pixel values.
- ⊛ Spectral redundancy, which is due to the correlation between different colour planes or spectral bands.
- ⊛ Temporal redundancy, which is due to the correlation between different frames in a sequence of images.

The goal of image compression techniques is to reduce the number of bits required to represent an image by removing these redundancies. There are many approaches to image compression, however they can be classified into two main groups: lossless and lossy compression.

In lossless compression, the reconstructed image is numerically identical to the original image on a pixel by pixel basis. Therefore, there is no loss of information in lossless compression but the amount of compression achieved is low. In lossy compression, the reconstructed image is not identical to the original image. As a result, much higher compression can be achieved as compared to lossless compression. In this paper, we analyse the behaviour of images when compressed using lossy compression techniques and introduce the notion of browsing levels.

3.1 Lossy Compression

The general framework for a lossy compression scheme is shown in Figure 1. It

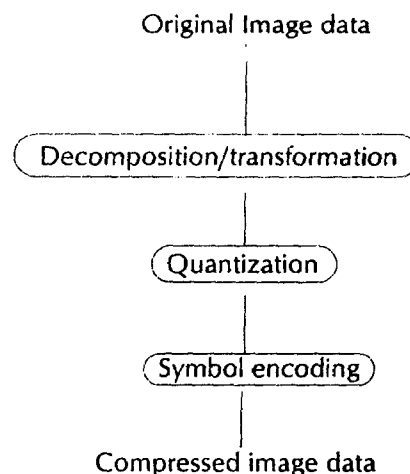


Figure 1: Components of Lossy Compression

includes three components: image decomposition, quantisation, and symbol encoding. Image decomposition or transformation is performed to eliminate redundant information and to provide a representation that can be coded more efficiently. Quantisation involves reducing the number of output symbols. The type and degree of quantisation has a large impact on the quality of lossy scheme. Symbolic encoding includes the use of techniques such as Huffman coding or arithmetic coding as a means of achieving rates close to the entropy of the quantised symbol source. Some of the most popular lossy compression techniques are JPEG coding²⁶, vector quantisation²¹, wavelet coding,^{15, 24} fractal compression^{10, 18}, sub band coding and hierarchical coding.

4. IMAGE QUALITY

One of the most important aspects of image communications is the quality of the images. There is no standard definition for image quality, but it can be loosely defined as the accuracy with which a compressed image can be regenerated. The basic arbiter of image quality is the human visual system. Image quality as observed by an human eye can vary from one person to another, thus making the judgement of image quality a very difficult task.

In order to standardise and automate the evaluation of image quality, several mathematical measures have been defined. Due to the complexity of the human visual system, it is hard to find a measure which as accurately reflects the quality of an image. One of the most common measures is the Root Mean Squared Error (RMSE)²⁸ between the original image and the regenerated image. The RMS model assumes that all errors of equal magnitude are equally visible and that errors combine in a quadratic fashion.

Another measure based on the colour values of an image has been proposed by Bhargava¹. This measure uses the colour histogram of an image to evaluate the quality of that image. It has been argued that the colour histogram of an image provides a more accurate measure as opposed to pixel to pixel matching used by the RMS model. PSNR²⁸ is another variation of

RMS which measures image quality as a ratio of the original image signal to the amount of noise present in the regenerated image. Several other measures discussed by Pratt²³ use different features of the image to evaluate image quality. Any one measure doesn't give an accurate reflection of image quality. If a number of different measures are combined together, we can approximate values for image quality. This means that measuring image quality is truly a multidimensional problem. We now discuss some of the image quality measures which have been used in our experimental calculations.

4.1 Colour Histogram based Technique

This technique, proposed by Bhargava¹, is based on the colour histogram of an image. A colour histogram of an image is a series of bars representing the different colours in the image. Each bar shows the number of pixels of a particular colour present in the image. For example, an image with the three colours red, green, and blue and a resolution 4 x 4, would have a colour histogram as follows: red: 9, blue:3, green: 4. This technique uses the colour histograms of the original and the regenerated image to calculate the colour difference between the corresponding bars.

4.1.1 Methodology

Given an image X, its loss induced version Y, and n, the total number of colours in X. The colour difference is calculated as:

$$\text{Col-diff} = \frac{\sqrt{\sum_{i=1}^n (\text{hist}_i^X - \text{hist}_i^Y)^2}}{\sum_{i=1}^n \text{hist}_i^n * n}$$

where: hist^X is the colour histogram for X and hist^Y is the colour histogram for Y

Based on the col-diff value, you can determine the quality of the regenerated image. A smaller value indicates better quality and vice versa. This measure tells you how to calculate the difference in the quality of the image but doesn't say how to produce an image of a given colour difference

value. Hence this method is not suitable for producing an image of a desired quality.

4.2 Root Mean Squared Error(RMSE)

This is the most commonly used quality measure. Before defining RMSE, we first define the mean absolute error(MAE), or per-pixel error. If the original image is of resolution ($m \times n$), then the mean absolute error is defined as:

mean absolute error

$$= \frac{1}{m} \frac{1}{n} \sum_{x=1}^m \sum_{y=1}^n |f(x, y) - g(x, y)|$$

where $f(x, y)$ and $g(x, y)$ represent the original and compressed images respectively. In contrast, the mean square error(MSE) averages the squares of pixel differences.

mean square error

$$= \frac{1}{m} \frac{1}{n} \sum_{x=1}^m \sum_{y=1}^n [f(x, y) - g(x, y)]^2$$

Root Mean Squared Error is the square root of Mean Square Error.

RMSE =

$$\sqrt{\frac{1}{m} \frac{1}{n} \sum_{x=1}^m \sum_{y=1}^n [f(x, y) - g(x, y)]^2}$$

RMSE is a reasonable measure for image quality but it has a few drawbacks. For example, if we shift the pixels of the original image to the right by 10, the reconstructed image would still be very close to the original but the RMSE value would be large indicating a large amount of distortion. Therefore systematic errors introduced into the image are not adequately handled by RMSE.

4.3 Peak Signal to Noise Ratio(PSNR)

Signal to Noise Ratio (SNR) doesn't disregard systematic errors but does attempt to improve on RMSE by taking into account the intensity of the original image. This is done by dividing the total image power with the total error power. It is calculated as

$$SNR_{ms} = \frac{\sum_{x=1}^m \sum_{y=1}^n f(x, y)^2}{\sum_{x=1}^m \sum_{y=1}^n [f(x, y) - g(x, y)]^2}$$

Where ms stands for mean square. As this measure is inversely proportional to RMSE, it is directly proportional to image quality. Therefore, higher the SNR value, better the image quality. Taking the square root, gives us the root signal to noise ratio, which is:

SNR_{rms}

$$= \left(\frac{\sum_{x=1}^m \sum_{y=1}^n f(x, y)^2}{\sum_{x=1}^m \sum_{y=1}^n [f(x, y) - g(x, y)]^2} \right)^{\frac{1}{2}}$$

Another option is to take the logarithm,

SNR_{log}

$$= 10 \log_{10} \frac{\sum_{x=1}^m \sum_{y=1}^n f(x, y)^2}{\sum_{x=1}^m \sum_{y=1}^n [f(x, y) - g(x, y)]^2}$$

Peak signal to noise ratio, which is the most commonly used measure is defined as follows:

PSNR

$$= 10 \log_{10} \left(\frac{[\max f(x, y) - \min f(x, y)]^2}{\frac{1}{m} \frac{1}{n} \sum_{x=1}^m \sum_{y=1}^n [f(x, y) - g(x, y)]^2} \right)$$

5. PREDICTING IMAGE QUALITY

The quality of an image can play an important role in image communications. Higher quality images take longer to transmit than lower quality images. All the measures discussed in the previous section are used to evaluate the quality of a compressed or rendered image. Consider a mechanism which does the reverse operation. Given a quality level for an image, we need to find the amount of compression possible. This would mean predicting the compression behaviour of an image based on the physical properties of the

original image. This mechanism, if possible, would be of enormous use in image communications. One of the most important applications would be a World Wide Web browser with multiple levels of browsing. For example, consider a web browser with five levels of browsing. Level one would mean browsing at the highest quality possible. All the images browsed would be of high quality without any quality loss. On the other hand, level five would be the lowest possible level, with the images being compressed with a substantial loss of quality. Level one would be ideal for scientists and engineers who need to further analyse and process the images obtained. Level five would be suitable for general browsing of the web, as the users only need to view the images for a short time and the image content is more important than the quality of the image. Depending on the intended use of the images, a suitable browsing level could be selected to achieve the desired level of quality.

Predicting the compression behaviour of an image is a complicated task. We have studied the behaviour of several images when compressed at various compression levels and observed that different images behave differently when compressed at the same compression level. This implies that images can be categorised into separate classes based on their compression behaviour, such that all the images in a particular class would have similar compression behaviour. Therefore, investigating one image in any class could tell us the compression behaviour for the rest of the images in that class. In the next section, we present an approach for predicting the compression behaviour and explain the results obtained.

6. EXPERIMENTAL EVALUATION

We assume that PSNR works fairly well with JPEG compression. A set of images were compressed using JPEG lossy compression at different compression ratios. The loss in quality was measured for each set of images using the different measures mentioned in section 3. As these mathematical measures do not reflect accurate values for image quality, a human survey was conducted to evaluate the quality of

the compressed images. Users were asked to rank the compressed images according to the quality they perceived on a scale of 1-5, 1 being the highest quality (indistinguishable from the original) and 5 being the lowest quality (unrecognisable). It was seen that all the images compressed up to 40% are indistinguishable from the original. As the compression ratio increases, the various images behave differently. The images labelled 'mongolia', 'trinidad' and 'india' do not lose any quality when compressed up to 70% and suffer a linear loss in quality thereafter. Similar behaviour is observed in the PSNR values for these three images. It was seen that the PSNR value drops steeply up to 70% compression and then follows a gradual slope. The images labelled 'morocco' and 'vietnam' have a similar behaviour as reflected in both the user and PSNR quality measurements. There is no loss in quality until 50% compression but, if compressed further, there is a sudden loss in quality. These results indicate that the colour distribution and homogeneity of an image plays a major role in determining its compression behaviour. The results imply that all images can be compressed to at least 40% of the original size without any loss in visual quality, and furthermore that certain images can be compressed at a higher compression ratio without any loss in visual quality. Many other images were tested and showed similar compression behaviour.

The second part of the experiment is aimed at predicting the compression behaviour of different images. In this paper, we focus on a fixed class of images which follow a certain compression behaviour. Since the PSNR values correspond more closely to the user quality judgements than the co-diff values, we use PSNR as a measure of quality in our experiments. Our goal is to predict the quality of an image by initially analysing the image.

First we compute the slopes for different segments of each image (m_1, m_2, \dots), using the values in Figure 1, and then calculate the average slope for each image (\bar{m}). We then compute the absolute difference between the average slope (\bar{m}) and the slope for the first segment of the image (m_1).

It can be observed that the magnitude of the average slope in most of the images was more

values with the actual values. It can be seen from the table that the error in the PSNR values

Image	m_1	\bar{m}	$ \bar{m}-m_1 $	Average $ \bar{m}-m_1 $
sweden.jpg	-0.815877	-0.94372	0.1278	
argentina.jpg	-0.83512	-1.12523	0.29011	
greece.jpg	-0.40668	-0.5859	0.1793	
swiss.jpg	-0.72800	-0.674	0.054	0.1495
morocco.jpg	-0.10277	-0.40768	0.30492	
panama.jpg	-0.6941	-0.6242	0.0699	
trinidad.jpg	-0.62636	-0.54307	0.073	
vietnam.jpg	-0.58762	-0.62342	0.036	
india.jpg	-0.4564	-0.673	0.22	
mongolia.jpg	-0.61601	-0.47724	0.14	

Table1: Computing the Average Slope and the Difference wrt m_1

than the slope of the first segment (m_1). This implies that the rate of drop in quality increases as the images are compressed more and more. Also the difference between \bar{m} and m_1 was within the minimum threshold value. Therefore, by adding an appropriate value to m_1 , we can obtain a value close to the average slope of the image. This value can be used to compute the PSNR values of the image at different compression levels. As the average difference between m_1 and \bar{m} is 0.1495, we use this value to evaluate the PSNR values for the images at different compression levels and compare these

calculated using the average slope was within acceptable limits. Therefore, using this approach, we can predict the approximate quality of an image at different compression levels.

Table 2 shows the original size of the image and size of the image at browsing level 2. The average reduction in the size of the images is 65.5%. This indicates that browsing at a slightly lower level could reduce the network traffic substantially and expedite the browsing process.

Image	m_1	$\bar{m} = m_1 + 0.1495$	PSNR calculated using \bar{m}	Actual PSNR
sweden.jpg	0.81588	0.96538	25.896	26.33
argentina.jpg	0.83512	0.98462	26.855	24.04
greece.jpg	0.4067	0.5562	33.347	31.977
swiss.jpg	0.10280	0.2523	45.21	40.23
panama.jpg	0.6941	0.8436	36.44	40.22
trinidad.jpg	0.6164	0.7659	40.83	44.42
vietnam.jpg	0.5876	0.7371	37.01	41.21
india.jpg	0.4564	0.6059	38.23	39.33
mongolia.jpg	0.6160	0.7655	36.33	41.35

Table 2: PSNR Values At 50% Compression

Image	Original size of image	Size of image at browsing level 2	% Reduction in the size of image
argentina.jpg	10438	5161	51
greece.jpg	17488	6912	61
morocco.jpg	14202	7115	50
panama.jpg	17844	5326	71
trinidad.jpg	34439	6885	80
vietnam.jpg	14937	5982	60
india.jpg	12642	2499	81
mongolia.jpg	8586	2579	70
Average reduction in size of images			65.5

Table 3: Image Size At Browsing Level 2

If every image on the internet is compressed before transmission, there would be a substantial reduction in the network traffic and response time would improve considerably. There could be a further drop in response time if the user chooses a higher level of browsing as the images could be compressed further.

Implementing multiple levels of browsing would mean assigning a fixed level of quality to each browsing level. Every image on the internet would be measured with respect to these quality levels. The amount of compression possible for each image in order to achieve these quality levels would be computed based on a fixed amount of preprocessing on the original image and then applying the appropriate mathematical function.

7. MULTIPLE BROWSING LEVELS AND DIGITAL LIBRARIES

The concept of multiple levels of browsing could be extended to multimedia digital libraries which have objects distributed over a wide area network. Each object (image, audio, video, etc) can be compressed using different compression ratios and assigned a quality level depending on the compression ratio. Users of this digital library at any site would be given a list of quality levels to choose from. Users can choose a quality level depending on their need for the data. Once a quality level is selected, the appropriate compressed file can be transferred to the user's host machine. This ensures that a user requesting lower quality data has a better response time

than a user requesting higher quality data. This would also reduce the network traffic substantially and lead to a fast and more effective usage of distributed digital libraries.

8. CONCLUSION

This paper presents an approach to classifying images into different classes based on their compression behaviour. Different features of an image (colour, texture, etc) play a significant role in determining the quality of an image. All these features of an image could be extracted and measured using suitable metrics. Based on these values, images could be grouped into different classes. Once the class of an image is identified, it would be easier to predict the compression behaviour of an image. Different compression levels could therefore be associated with corresponding browsing levels of a multi resolution browser.

The method presents a simplistic approach to achieve browsing levels. In order to improve efficiency and ease of use, further work needs to be done in the following areas:

- ⊕ Classifying images based on the original image.
- ⊕ Understanding the relationship between the quality measures and human perception.

We have introduced a new concept of browsing levels to reduce the traffic on the internet and make network applications faster. This feature could be incorporated in a web browser, to make browsing faster and more efficient. Other

multimedia network applications could also benefit by the use of this technique. In summary, multilevel compression would lead to better and more efficient usage of network resources.

REFERENCES

1. Annamalai, M, Sundaram A, BB. A color based technique for measuring visible loss in image data communication. *Database and Expert Systems Applications (DEXA)*, 1996, 39-48.
2. Baird, H, and Bunke H, YK. Structured document image analysis. Springer verlag, 1992.
3. Bates, R, and McDonnell, M. Image restoration and reconstruction. Oxford University Press, New York, 1986.
4. Baxes, GA. Digital Image Processing. John Wiley, 1994.
5. Bhargava, B, and Annamalai, M. Communication costs in digital library database. *Database and Expert Systems Applications (DEXA)*, 1995, 1-13.
6. Gonzales, R, and Woods, R. Digital image processing. Addison Wesley Publishing, USA, 1992.
7. Grosky, William I and Jain, Ramesh MR (Eds). The handbook of multimedia information management. Prentice Hall, New Jersey.
8. Hodge, W. Interactive television. McGraw-Hill, Inc, 1995.
9. Jabbari, B (Ed). Worldwide advances in communication networks. Plenum Press, 1994.
10. Jacquin, AE. Image coding based on a fractal theory of iterated contractive image transformations. *IEEE Transactions on Image Processing*, 1992, 1, 18-30.
11. Jain, A. Fundamentals of digital imaging. Prentice Hall, 1989.
12. Janhe, B. Digital image processing- concepts, algorithms, and scientific applications. Springer-Verlag, 1991.
13. Kessler, J. Internet digital libraries. Artech House, 1996.
14. Liu, L-G. The internet, and library and information services. Greenwood Press, 1995.
15. Antonini, M, Barlaud, PM, and Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2), 1992, 205-220.
16. Majid Rabbani, PWJ. Digital image compression techniques. SPIE Optical Engineering Press, 1991.
17. Marimon, D. Hypertext/hypermedia for libraries. *Advances in Library Automation and Networking*, 1991, 121-148.
18. Monro, D and Dudbridge, F. Fractal approximation of image blocks. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1992, 485-488.
19. Morley, J and Gelber, S. The emerging digital future. Boyd & Fraser Publishing Company, 1996.
20. Nabil R Adam, Bharat Bhargava, YY (Eds). *Digital Libraries*, 1994.
21. P Westerink, J Biemond, DB. and Woods, JW. Subband coding of images using vector quantisation. *IEEE Transactions on Communications* 1988, 36(6), 713-719.
22. Palmer, AW. and Kellerman, AS. Distributed multimedia. Addison-Wesley, 1996.
23. Pratt, W. Digital image processing. John Wiley & Sons, New York, 1978.
24. RA DeVore, ABJ, and Lucier, BJ. Image compression through wavelet transform. *IEEE Transactions on Information Theory*, 1992, 38(2), 719-746.
25. Russ, J. The image processing handbook. CRC Press, 1995.
26. Wallace, GK. The JPEG still compression standard. *Communications of the ACM*, 1991, 34(4), 31-34.
27. Warren Hicks, AT. Developing multimedia libraries. RR Bowker, 1970.
28. Watson, AB. Digital images and human vision. The MIT Press, 1993.

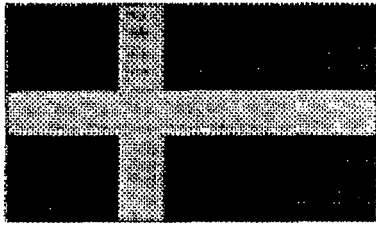


Figure 5: sweden.jpg

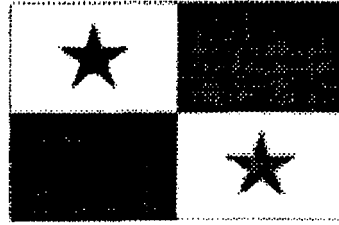


Figure 10: panama.jpg

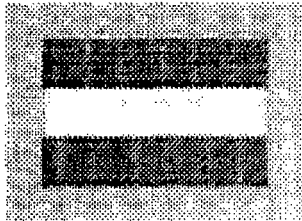


Figure 6: argentina.jpg



Figure 11: trinidad.jpg

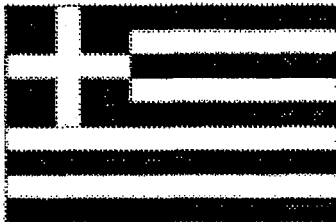


Figure 7: grecco.jpg

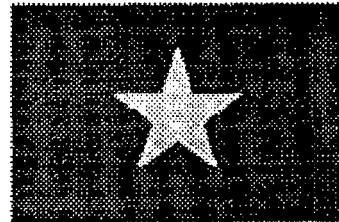


Figure 12: vietnam.jpg

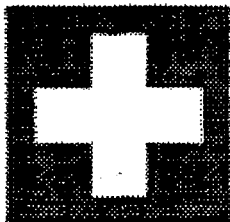


Figure 8: swiss.jpg

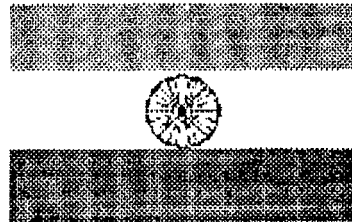


Figure 13: india.jpg

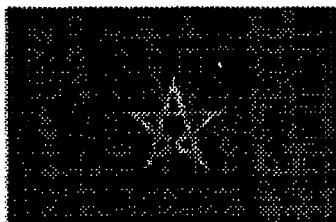


Figure 9: morocco.jpg

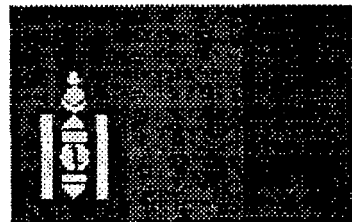


Figure 14: mongolia.jpg