Image and Movie Compression

Review

Last time we talked about data compression techniques.

Compression takes advantage of the difference between data (characters) and information (surprise).

We can remove or reduce redundancy in the data and send fewer 1s and 0s with the same information (lossless compression) or at least the most important information (lossy compression).

In our example we got fairly close to the theoretical best compression with a simple algorithm.

Image and Movie Compression

We said last time that we would expect images and movies to be particularly good candidates for data compression because there is a lot of redundancy. And important since these files are huge.

Remember you generally will not have a lot of change from one pixel to the next or one frame to the next. The amount of **surprise** is small. That means there is redundancy.

We also talked about the idea that images (unlike most other data) can be subjected to lossy compression safely.

We will see that the lost information is sometimes irrevelvant.

Even with substantial loss, there is no risk of mistaking the image for something else in most cases.

Image and Movie Files

Uncompressed images are huge files and movies are even bigger!

A 6 megapixel image, say 3000 x 2000 pixels, is relatively small these days. But with 3 bytes per pixel it's 18 million bytes!

A quite low res two hour movie, let's say 600 x 400 pixels and 24 frames per second, would have

600 x 400 x 3 bytes = 720,000 bytes PER FRAME

But there would be 24 frames per second for two hours so that's

24 x 2 x 3600 = 172,800 frames. So we would have

172,800 x 720,000 = about 124 x 10^9 bytes?!?! + sound data!

This just won't work.

Simple Compression Method: Run Length Encoding

Instead of storing each pixel, store the number of identical pixels (run length) and their colour e.g., 100 pixels all white.

Consider a document with text only. This is still a pretty common use of images.

Has large sections of white – whole lines in fact.

Get a dramatic reduction in file size for a text-only image.

This is also a component of data compression methods for more complex images, but by itself does not give much compression for more complex images.

Lossy Compression

A simple example of a lossy compression technique:

RLE works well only if there are lots of identical pixels.

In a typical image, adjacent pixels are probably close but not identical in value.

One can scan the image and replace pixels that differ by +/- 1 with identical pixels, and then use RLE.

JPGs-Lossless and Lossy

JPEG or Joint Photographic Experts Group

A standard that actually includes multiple lossless and lossy data compression routines that are adapted to the content of a particular image.

JPG-Lossy

Example: One can use something similar to a “Fourier Transform” to remove rapid variations from the image and then use a lossless compression method such as RLE or Universal Coding

Tradeoff: as you use more sophisticated compression routines you save space and bandwidth, but you increase the computing time.

File Types for Images

Image.bmp Bitmap Uncompressed
File size is very close to the number of pixels x 3 bytes per pixel.

Image.png Portable Network Graphics:
Lossless compression

Image.jpg Joint Photographic Experts Group
Lossless and lossy compression

Examples

A good case for compression: one colour

Yellow.bmp	300.1 kB	Image
Yellow.jpg	2.2 kB	Image
Yellow.png	691 bytes	Image

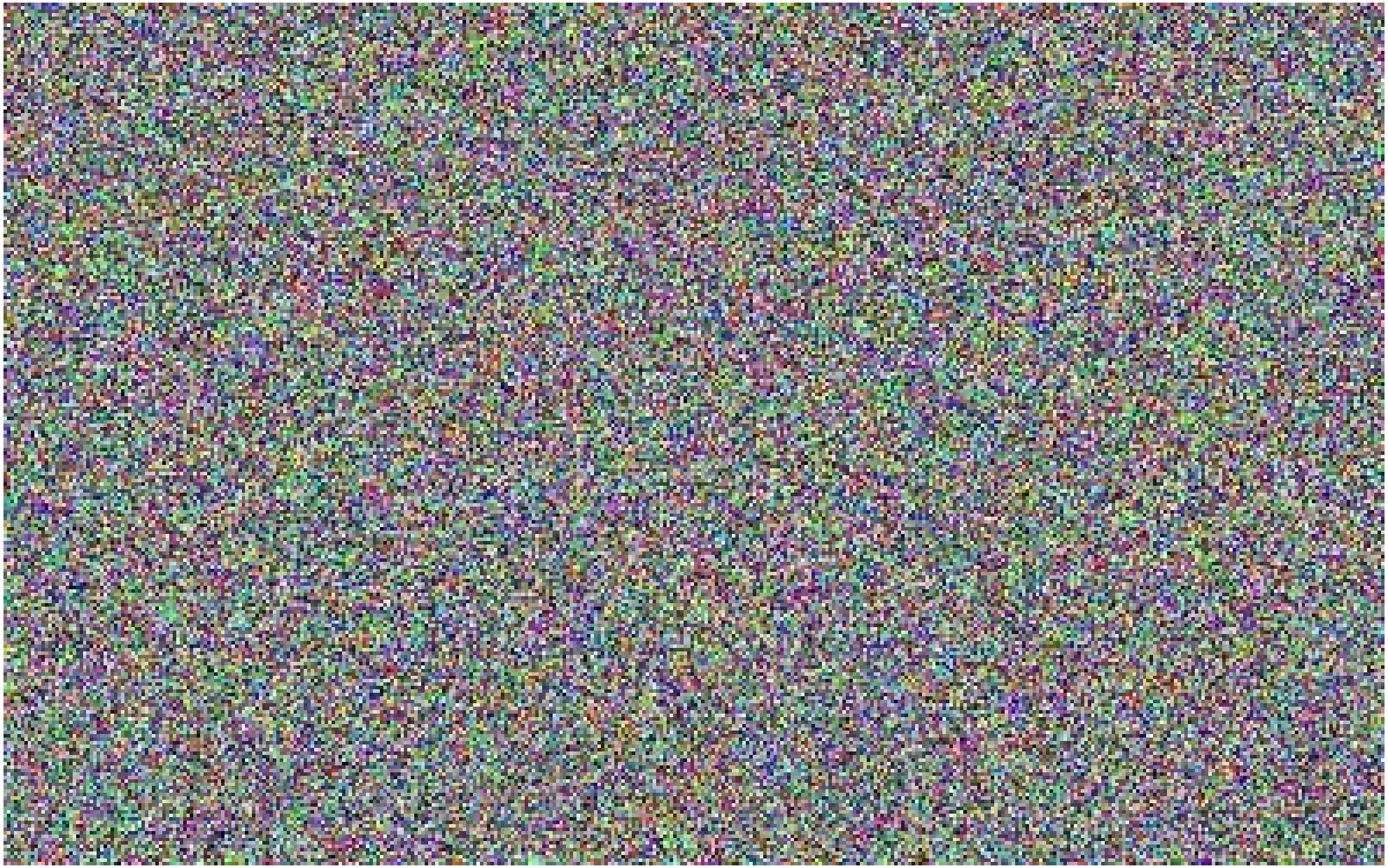
Note the jpg is larger than png!

This is due to the overhead associated with the fancy compression routines in a jpg.

For very simple images png is sometimes better.

Examples

Random Image



Examples

Random.bmp	300.1 kB	Image
Random.jpg	68.6 kB	Image
Random.png	300.5 kB	Image

This file has no redundancy: each pixel is a complete surprise. No lossless compression is possible.

Lossy compression does reduce the file size but a careful look at Random.jpg shows it is smeared.

Examples

Simple Image

Predict:

A. Lossless will be effective, lossy not so much

B. Lossy will be effective, lossless not so much

C. Both will be effective

D. Neither will be effective

Examples

Simple Image

Simple.bmp	300.1 kB	Image
Simple.jpg	22.5 kB	Image
Simple.png	144.8 kB	Image

A. Lossless will be effective, lossy not so much
B. Lossy will be effective, lossless not so much
C. Both will be effective
D. Neither will be effective

Examples

Complex Image

Predict:

A. Lossless will be effective, lossy not so much

B. Lossy will be effective, lossless not so much

C. Both will be effective

D. Neither will be effective

Examples

Complex Image

Complex.bmp	300.1 kB	Image
Complex.jpg	39.7 kB	Image
Complex.png	283.0 kB	Image

Predict:

A. Lossless will be effective, lossy not so much

B. Lossy will be effective, lossless not so much

C. Both will be effective

D. Neither will be effective

Movies

Movies offer lots of opportunities for compression.

Basic idea is that one frame in a movie cannot be very different from the previous frame.

And remember each frame is an image that can be compressed.

Result is extensive compression.

Questions

You are given a bmp file for a very high quality image.

You use a photo editing program to convert it to a PNG with lossless compression.

Unfortunately you lost the original bmp file.

Is it possible to recover the original bmp?

Questions

Your friend was given another high quality image as a bmp file.

He used a photo editing program to convert it to a jpg and lost the original.

Is it possible to recover the original bmp?

Questions

Sometimes a lossy compression routine will make a file bigger (oops). Under what circumstances is this likely to happen?

Summary

Images and movies contain extensive redundancy.

This makes them good candidates for both lossless and lossy data compression, and both are used extensively.

Both lossless and lossy compression is also used for sound files.