

Make it Small

The Magic of Audio and Video Compression

Jack Moffitt
@metajack
<http://metajack.im>

**Compression is
everywhere.**

Texting:

“do u want 2 go b4 the end?”

Digitizing **sound** or **video**

iPods

Making the Internet
faster

Compression comes in two flavors:
lossless and **lossy**.

Lossless compression is
exactly
reversible.

Lossy compression is
mostly
reversible.

Lossless Compression

It compresses
textual data
extremely well.

Lossless image compression ratios are usually between **4 to 1** and **10 to 1**.

Specialty lossless audio codecs
can get near a
2 to 1 ratio.

Run Length Encoding (RLE)

Replace **consecutive** values with a **single value** and a **count**.

4 8 15 16 16 16 16 16 16

4 8 15 **7x16**

internationalization = i18n

andreesen horowitz = a16z.com

Lempel-Ziv-Welch (LZW)

Replace arbitrary **sequences** with a **code.**

The table of codes is built up from the **data.**

Lossy Compression

There's no need to store information
humans can't perceive.

Audiovisual data can
degrade in **perceived quality**
without affecting **comprehension**
too much.

Compression ratios are **much higher** than lossless codecs.

JPEG files can get from **10 to 1** to **100 to 1**.

Audio is often over **10 to 1**.

Auditory Oddities

Your brain is **magic**.

Listen to the **proof**.

<http://www.youtube.com/watch?v=0bG7EFhMw8w>

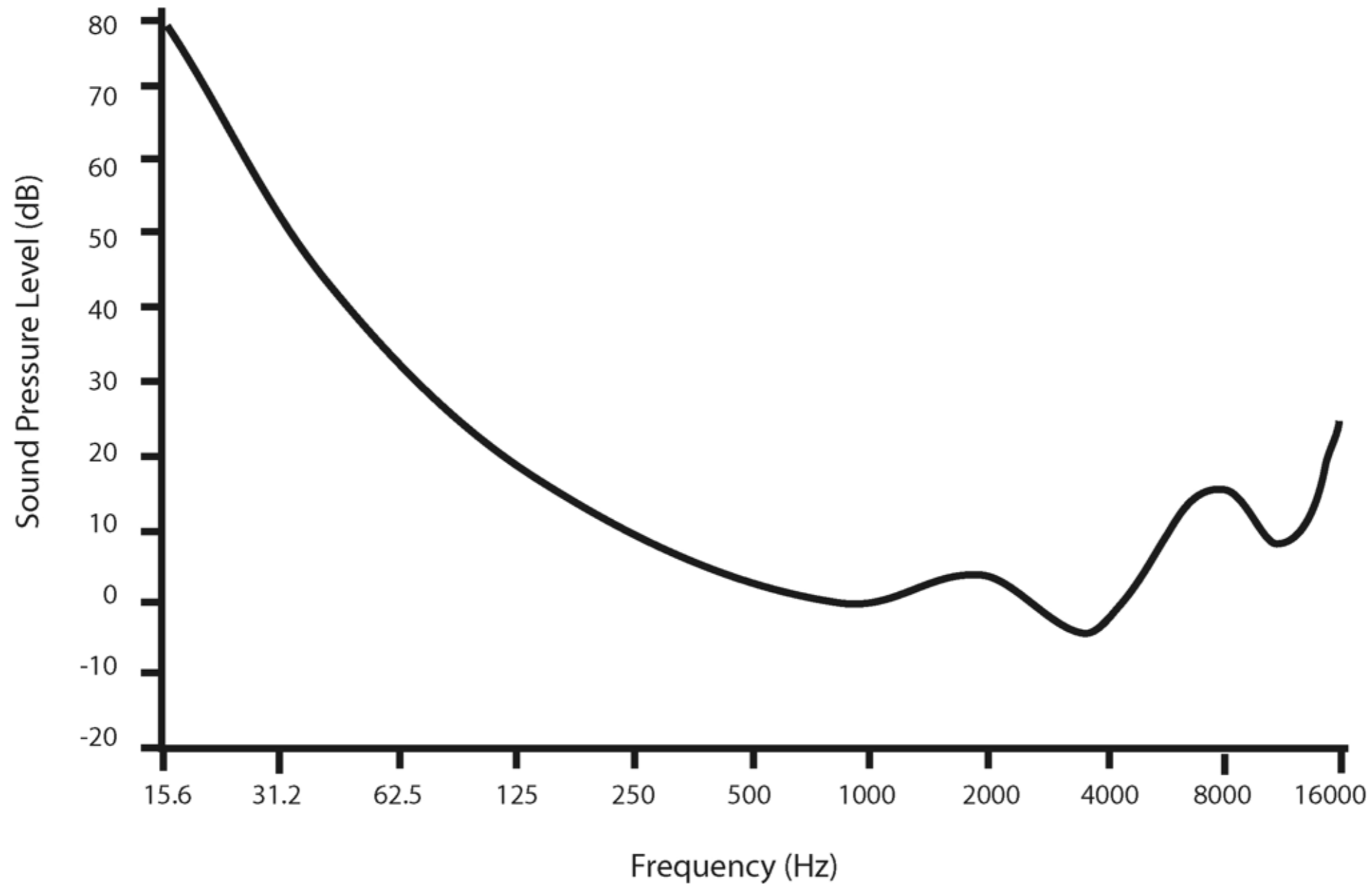
Losing Sound

Range of human hearing is just
20Hz to 20,000Hz.

Equal Loudness Contours

1933 - Fletcher and Munson

Perceived Human Hearing



Auditory Masking

Loud sounds block quiet sounds.

Temporal Masking

Loud sounds hide quiet sounds
you already heard and also sounds
you will hear soon.

Big changes are more important than small ones.

Lossless compression is used at the very end to pack all the information as tightly as possible.

Losing Pictures

JPEG Compression

Humans see
brightness
much better than
color.

Step 1:

Convert from RGB to YCbCr.

Y = brightness

Cb and Cr are color information

Step 2:

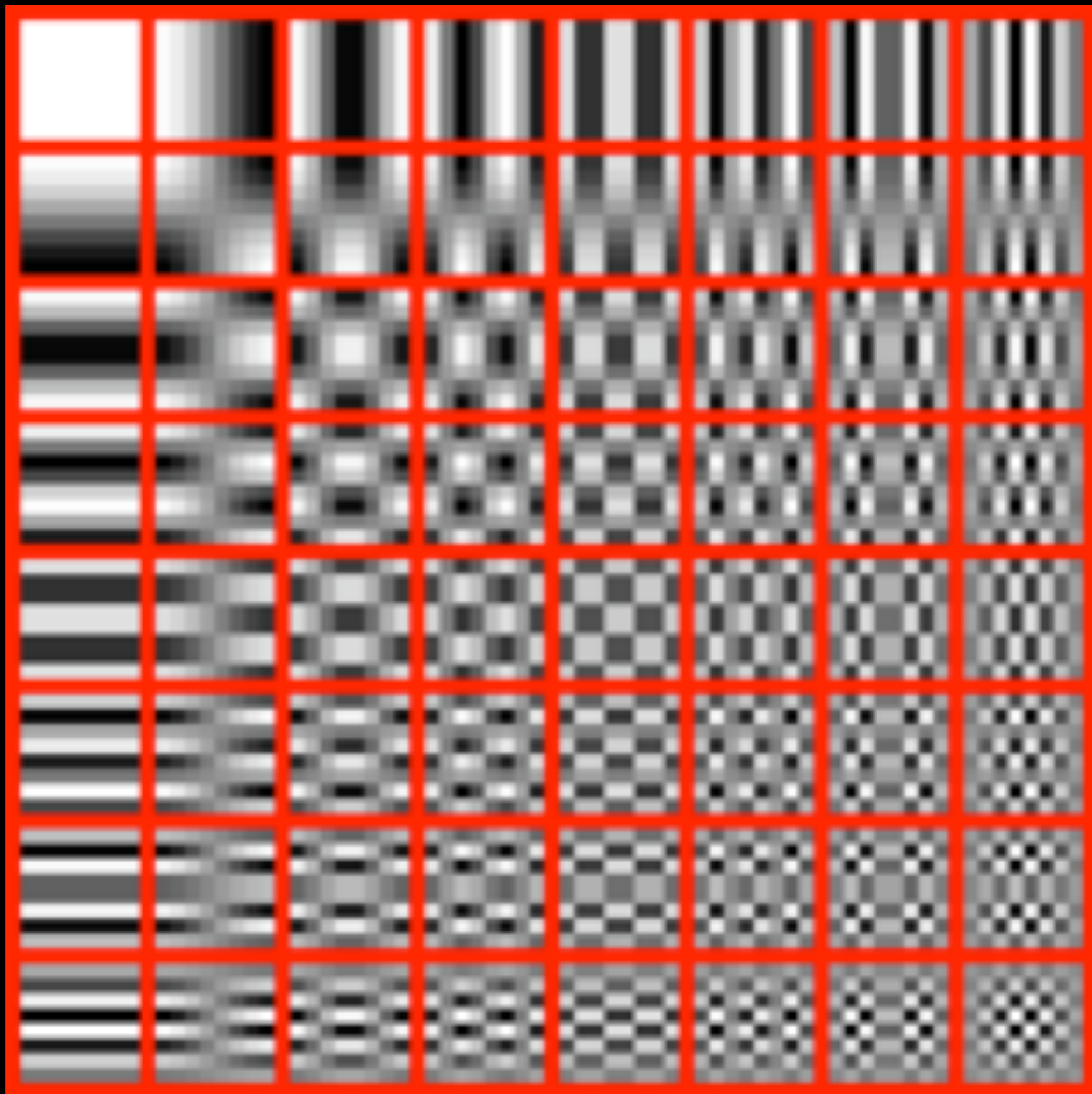
Reduce the resolution of the
color information.

Step 3:

Divide the image into 8x8 blocks.

Step 4:

Use the
Discrete Cosine Transform (DCT)
to split the block into
spatial frequencies.



Step 5:

Frequencies with
large magnitudes
are **most important.**

One of the **hardest parts** is smartly storing all these **extra numbers**.

Similar techniques to JPEG
compression
can be used for video using
3D wavelet transforms.

Losing Video

Video compression tends to use a
layered approach.

Codecs range in sophistication from
motion JPEG to **MPEG-4**.

Data from
previous frames
can be used by
later frames.

Instead of looking at the full frame,
you can analyze
only the differences.



Predicting motion
can also help quite a bit.





Support
Royalty Free
Open Standards!

WebM - <http://webmproject.com>

Ogg Vorbis, Ogg Theora, FLAC
<http://xiph.org>

HTML5 video is *good*, but could be **awesome**.

Questions?