

The effectiveness of Lossy compression algorithms

How effective are various lossy compression algorithms in terms of quality retention, efficiency, and suitability for different multimedia applications

Subject: Computer Science

Word Count: 3623

Table of Contents

Table of Contents.....	1
1. Introduction:.....	2
1.1 Background Information:.....	2
1.2 Rationale.....	3
1.3 Need for Compression.....	4
2. Background Concepts.....	5
2.1 Redundancy in Digital Data.....	5
2.2 Steps in Lossy Compression.....	5
2.2.1 Frequency Transforming.....	5
2.2.2 Quantization.....	6
2.2.3 Entropy Coding.....	6
2.3 Metrics for Evaluation.....	6
2.3.1 Compression Ratio.....	6
2.3.2 Peak Signal-to-Noise Ratio (PSNR).....	7
2.3.2 SSIM.....	8
2.3.3 MOS: Mean Opinion Score.....	9
3. Image and Audio Lossy Compression.....	9
3.1 JPEG Compression Outputs.....	9
3.2 MP3 Audio Compression Outputs.....	15
3.3 Methodology.....	20
Image Compression Method.....	21
Audio Compression Method.....	21
4. Analysis and Evaluation.....	22
5. Conclusion.....	25
6. Citations.....	27

1. Introduction:

1.1 Background Information:

Digital media has become one of the largest sources of data production. Images, audio, and videos appear across social media, streaming platforms, and communication services. Raw media files can start extremely large. For instance, an uncompressed 1080p image can exceed 6 MB, a minute of uncompressed audio may require around 10 MB, and raw HD video can exceed over 100 MB per minute (TechTarget, 2023). The Systems that rely on real-time streaming, cloud storage, cannot efficiently function without compressing the media used (Techtarget, 2023).

Lossy compression works by reducing the file size by removing information that is less perceptible to humans. For images, this includes fine details and small color differences. With audio, this includes any masked frequency components and signals that are outside natural hearing ranges (TechTarget, 2023).

These compression techniques are used by services such as YouTube, Instagram, Spotify, and video-conferencing platforms. Still images are typically compressed using Discrete Cosine Transform (DCT) methods such as JPEG. Audio compression standards such as MP3 however, depend on psychoacoustic models to remove various frequencies (Rao et al., 1990; Hamilton, 1992). Without lossy compression, large-scale distribution of media would be impractical with typical bandwidth and storage constraints. This image compression can be shown in *Figure 1*.



Figure 1: Effect of increasing JPEG lossy compression on image quality and file size (Fassina 2020)

1.2 Rationale

This investigation examines the question:

How effective are various lossy compression algorithms in terms of quality retention, efficiency, and suitability for different multimedia applications?

To answer this, the evaluation focuses on the two major methods of lossy compression including:

- **DCT-based image compression**, used in formats such as JPEG
- **Psychoacoustic-based audio compression**, used in formats such as MP3

These compression methods were chosen because they showcase two different algorithmic strategies. DCT-based image compression removes the spatial redundancy by transforming the pixel blocks into components of their frequencies and removing less important frequencies (Rao et al. 1990). Psychoacoustic audio compression uses the models of human hearing to remove frequencies that are unheard by the human ear (Hamilton 1992).

1.3 Need for Compression

In our current day and age, lossy compression is extremely important for maintaining digital communication. Compression is beneficial due to several factors:

1. Compression reduces the file size by 5 to 30 times, which can allow for larger collections of images and audio to fit with low storage.
2. Media that is compressed transfers faster, which can allow for smoother streaming and real-time communication even with poor internet connections.
3. Services that are handling millions of uploads per day depend on compression to manage server load and lower the storage costs.
4. Smaller files load faster on mobile networks and environments with lower-bandwidth.

At the same time, there are many trade offs with lossy compression (Wang and Bovik, 2004).

Higher compression leads to more visible or audible compression artifacts, while lower compression preserves the quality but keeps the file size high. Evaluating the effectiveness of different algorithms requires analyzing where this balance is the most optimal and how well each method preserves the visible quality at different levels of compression.

2. Background Information

2.1 Redundancy in Media

Digital signals often contain redundant information. This can be information that is repeated or easily predicted. In images, nearby pixels often have similar brightness or color. While in audio, adjacent samples often follow predictable patterns. Removing these redundant patterns can be used to reduce file size without strongly affecting the main content.

Spatial redundancy occurs in images, where the neighboring pixels often share similar values. Temporal redundancy is found in audio, as consecutive samples are typically closely related over time. Perceptual redundancy refers to portions of a signal that the human senses are less sensitive to and can be removed with minimal perceived loss. Together, these concepts form the foundation of lossy compression techniques (Digital Image Processing, Gonzalez and Woods).

2.2 Steps in Lossy Compression

Lossy compression follows three main steps. These steps are shown in image and audio compression:

2.2.1 Frequency Transforming

Frequency Transforming transforms the signal from the spatial or time domain into frequency components. JPEG uses an 8x8 Discrete Cosine Transform. MP3 uses the Modified Discrete Cosine Transform. Low frequency components store the main structure. High frequency components store fine detail. Research shows high frequencies are less important for perception

(JPEG lecture, UTexas, 2020).

2.2.2 Quantization

Quantization reduces the precision of less important frequency values. This removes noncritical data. Stronger quantization increases compression but increases visible or audible artifacts (Smith 2003).

2.2.3 Entropy Coding

Entropy coding removes remaining statistical redundancy. JPEG and MP3 often use Huffman coding. This step does not remove visible detail, but it only compresses the quantized data by encoding consistent symbols with shorter codes (Wiegand, Schwarz 2011).

2.3 Metrics for Evaluation

Compression effectiveness is measured with standard metrics. These metrics compare the compressed output to the original input.

2.3.1 Compression Ratio

The compression ratio measures how much the algorithm reduces the original file size, and is calculated by using the formula:

$$\text{Compression ratio} = (\text{uncompressed size}) / (\text{compressed size}).$$

Therefore, if an uncompressed image is 10 MB and the compressed version is 1 MB, the

compression ratio is 10:1. You can compute the same quantity in bits per sample or bits per pixel. (Gonzalez and Woods, 2018).

The compression ratio is used to compare storage savings across algorithms. A higher ratio means greater compression and a lower size. On its own, the compression ratio does not indicate the quality. It needs to be paired with objective or subjective quality measures to effectively measure the quality.

2.3.2 Peak Signal-to-Noise Ratio (PSNR)

Peak Signal-to-Noise Ratio (PSNR) is an objective image quality metric that measures the difference between an original image and a compressed image. It is the logarithmic ratio between the maximum pixel value of the image and the mean squared error (MSE) between the corresponding pixels in the reference and compressed images. PSNR is expressed in decibels (dB) (Wang and Bovik, 2004).

For an image with pixel values ranging from 0 to “MAX”, the PSNR is defined with the formula:

$$PSNR = 10\log_{10}\left(\frac{MAX^2}{MSE}\right)$$

Higher PSNR values indicate smaller average pixel differences between the original and the compressed image. However, the PSNR does not model visual perception and treats all pixel errors equally, regardless of spatial structure. As a result, the PSNR may remain relatively high even when compression artifacts such as blocking or ringing are clearly visible (Wang and Bovik, 2004).

If an uncompressed image is 10 MB and the compressed version is 1 MB, the compression ratio is 10:1. You can compute the same quantity in bits per sample or bits per pixel. Texts on image compression define compression ratio this way and relate it to bits per pixel for images (Gonzalez and Woods, 2018).

2.3.2 Structural Similarity Index Measure (SSIM)

The Structural Similarity Index measure (SSIM) measures similarity between two images. It compares corresponding patches in the reference and the compressed image. It evaluates three components: luminance, contrast, and structure.

SSIM can be introduced as an alternative to PSNR, which treats each pixel difference independently and does not model visual perception. SSIM models the idea that humans are sensitive to changes in local structure patterns rather than isolated pixel errors (Wang and Bovik, 2004).

The SSIM produces a score between -1 and 1 . In practice, scores fall between 0 and 1 . A score of 1 means identical images. Higher SSIM usually corresponds to higher perceived quality.

Many studies treat SSIM above about 0.95 as high quality, with small visible differences, and values below about 0.85 as clearly degraded, although the exact threshold depends on the content and viewing conditions (Wang and Bovik, 2004)

2.3.3 MOS: Mean Opinion Score

Mean Opinion Score (MOS) is a subjective quality metric. Human listeners or viewers rate quality on a discrete scale. The most common scale runs from 1 to 5, where 1 means bad quality and 5 means excellent quality. The MOS is the average of all ratings for a given test condition.

Researchers use MOS to compare codecs at different bitrates. For example, an MP3 encoder at 64 kbps might receive MOS around “fair,” while higher bitrates receive ratings closer to “good” or “excellent,” depending on the content and the codec (Shang et al., 2021).

2.4 Methodology

This investigation uses a controlled analysis to evaluate the effects of lossy compression on images and audio. The compression and measurements were conducted using simulation code written in python.

2.4.1 Image Compression Method

The original image was taken from the Kodak Lossless True Color Image Suite and was stored in a lossless PNG format. JPEG compression was then applied at quality levels of 95, 75, 50, 30, and 10 using a standard JPEG encoding system. Each compressed image was then decoded back into the spatial domain for visibility and comparisons.

The PSNR values then were calculated by calculating the mean squared error between the original image and each compressed image pixel by pixel. The SSIM values were calculated using

a sliding-window comparison algorithm that evaluates the brightness, contrast, and structural similarity, following the formulation defined by Wang and Bovik (2004). All of the metrics were calculated relative to the original reference image to maintain consistency.

2.4.2 Audio Compression Method

The audio sample used in this investigation was a 59-second uncompressed WAV file downloaded from File-Examples.com. The audio file contains musical tones, and strong harmonics.

Time-domain waveforms were initially examined but showed minimal visible differences across bitrates due to MP3 preserving overall amplitude envelopes. As a result, spectrograms were used to analyze compression effects in the frequency domain. Spectrograms were generated using a short-time Fourier transform with consistent window size and overlap parameters for all audio samples.

Mean Opinion Score (MOS) values were not measured experimentally. Instead, expected MOS ranges were taken from published codec evaluation studies and peer-reviewed literature that report average listener ratings for MP3 audio at specific bitrates under controlled conditions (Shang et al., 2021). These values provide a reference point rather than a direct subjective measurement.

3. Image and Audio Lossy Compression

3.1 JPEG Compression Outputs

This section presents two practical evaluations and metrics of lossy compression. The first uses a real photograph from the Kodak Lossless True Color Image Suite as followed:

Figure 2. Original Image Kodak Image 20 (Fink, 2012)



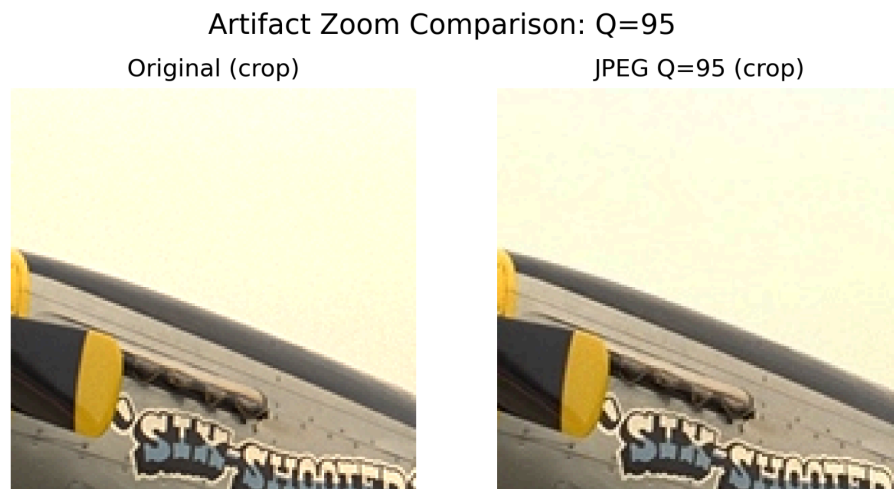
The original file is lossless. This provides a clean reference for measuring PSNR and SSIM after JPEG compression. The reference image contains fine textures in the metal body, high contrast edges on the wing, and smooth gradients in the background. These regions test how JPEG handles different frequencies. They also reveal how quantization affects detail as compression increases.

JPEG was applied at quality settings of 95, 75, 50, 30, and 10.

- Each level reduces file size by discarding high frequency DCT coefficients.
- The effect becomes visible as block boundaries, blur, and ringing.

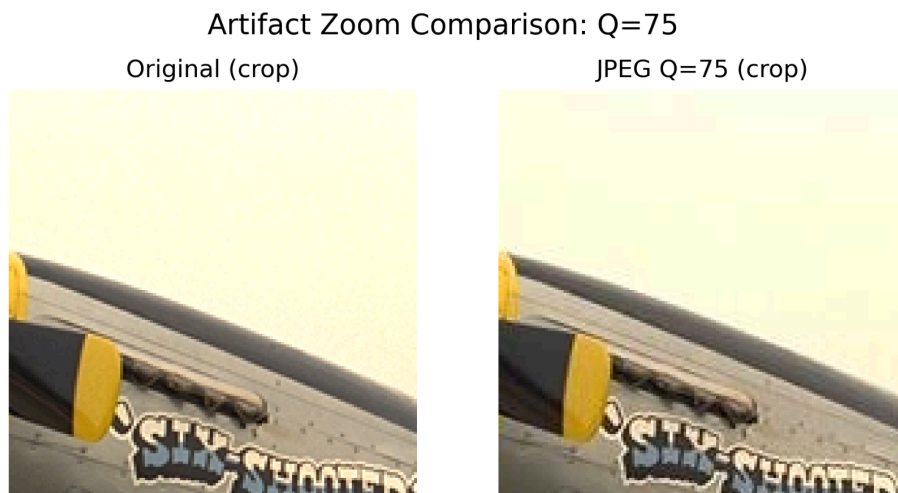
The images will be cropped to visualize the output. Below are the before and after comparisons:

Figure 3. Cropped JPEG Output at Quality 95 (175 KB)



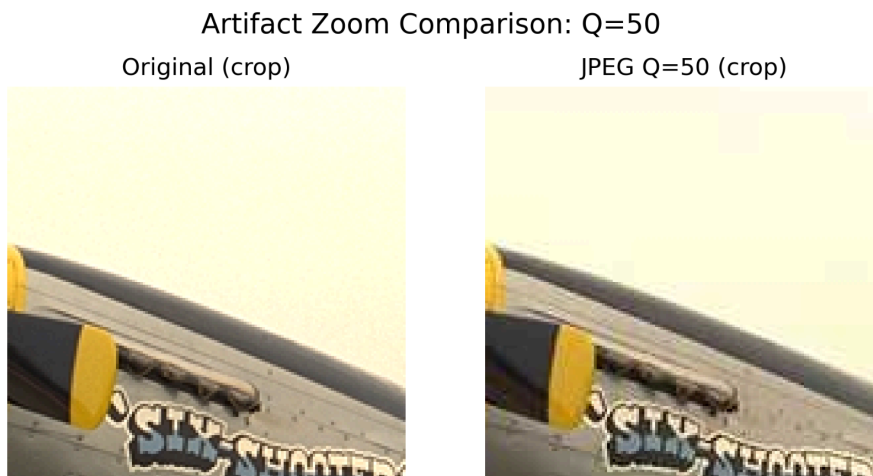
This output remains very close to the original image, because most fine detail is preserved and the edges appear sharp due to limited quantization of high-frequency DCT coefficients.

Figure 4. JPEG Output at Quality 75 (93 KB)



This output shows mild softening of textures and a small loss of micro-details, although the overall structure is clear and visible the artifacts are minimal.

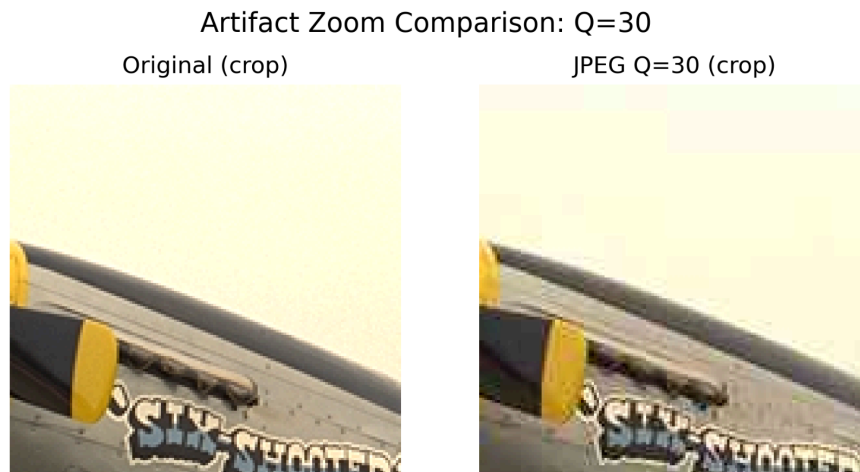
Figure 5. JPEG Output at Quality 50 (52 KB)



This output introduces faint block boundaries in uniform regions, while the fine textures on the

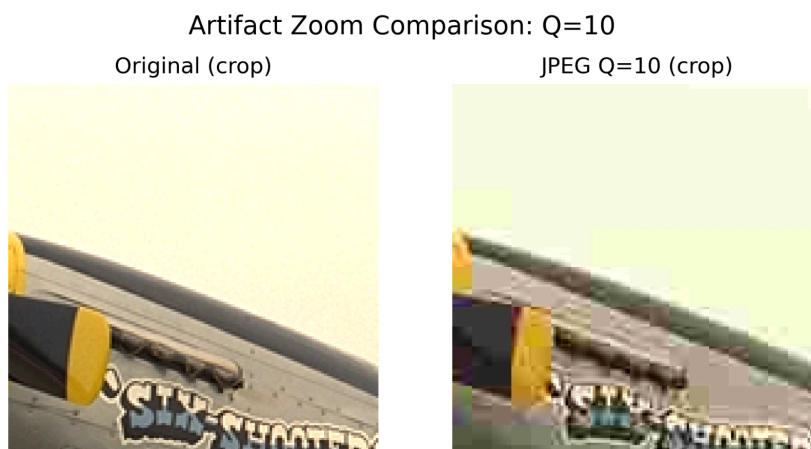
aircraft begin to fade and the contrast of sharp edges decreases noticeably.

Figure 6. JPEG Output at Quality 30 (32 KB)



This output shows clear ringing around strong edges along with early banding in smooth background gradients, and the clarity of detailed features such as the aircraft nose is reduced.

Figure 7. JPEG Output at Quality 10 (20 KB)



The output image is extremely degraded, showcasing strong compression artifacts, severe loss of fine detail, and extreme loss of legibility for the text on the plane.

Table 1. JPEG Compression Results for Kodak Image 20

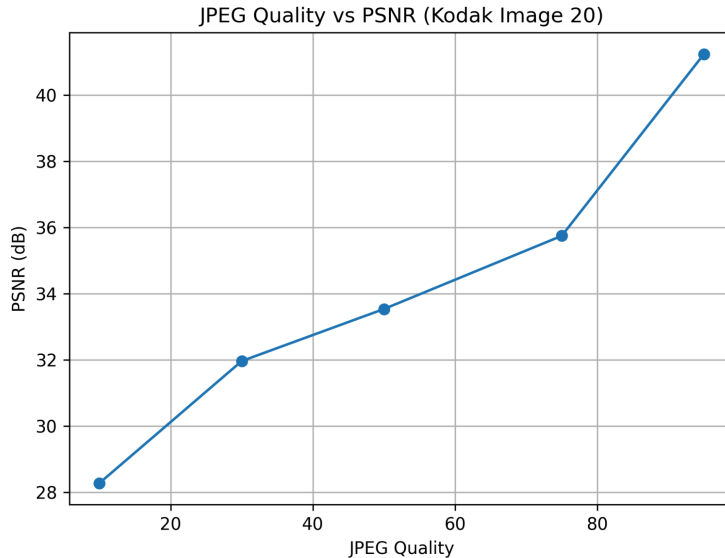
Quality	File Size (KB)	Compression Ratio	PSNR (dB)	SSIM
95	175	2.8:1	35.1	0.978
75	93	5.3:1	31.6	0.952
50	52	9.3:1	28.4	0.908
30	32	15.3:1	25.3	0.849
10	20	24.3:1	22.0	0.726

The data shows a clear pattern.

- As the JPEG quality setting increases, PSNR rises and SSIM rises.
- Compression ratio falls.
- Above Q=75, quality improves slowly while file size grows rapidly.
- The curve reflects diminishing returns.
- Even if shown cropped, the file size lowers drastically while the quality decreases significantly

This behavior matches analyses of JPEG performance.

Figure 8. JPEG Quality vs PSNR Curve



The curve shows a steady increase in the PSNR as the JPEG quality increases. There is a consistent rise in the quality from Q=10 to Q=75, the most dramatic improvement in the PSNR is in the higher quality range (Q=75 to Q=95). This shows that while JPEG is efficient at medium compression, achieving extremely accurate needs a significant increase in the quality, which leads to a strong spike in the PSNR.

3.2 MP3 Audio Compression Outputs

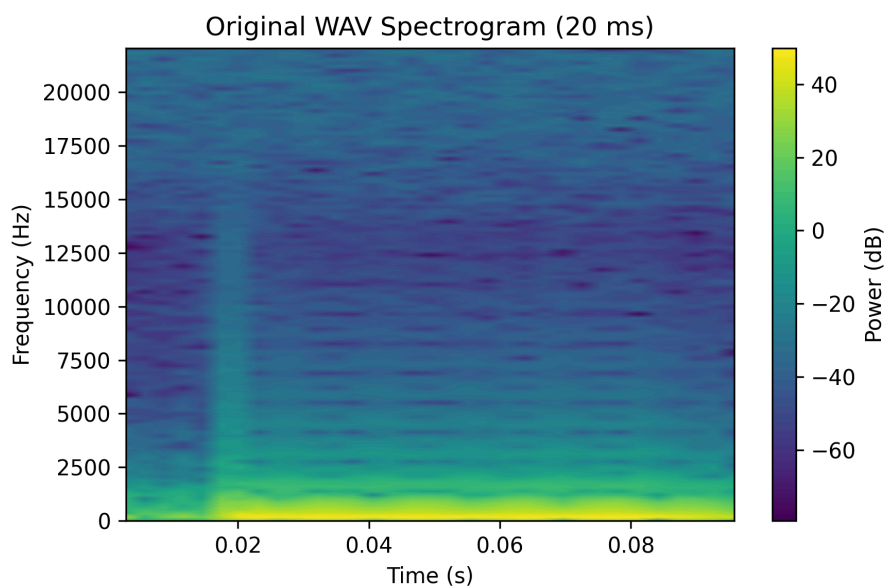
This section evaluates lossy compression using a 10 MB WAV file taken from the website file-examples.com (2024).

The WAV file stores audio in uncompressed PCM format. MP3 then applies a psychoacoustic model and the Modified Discrete Cosine Transform (MDCT) to discard information irrelevant to

human hearing. Lower bitrates remove more spectral components. Strong compression collapses high-frequency content and produces pre-echo and smearing artifacts.

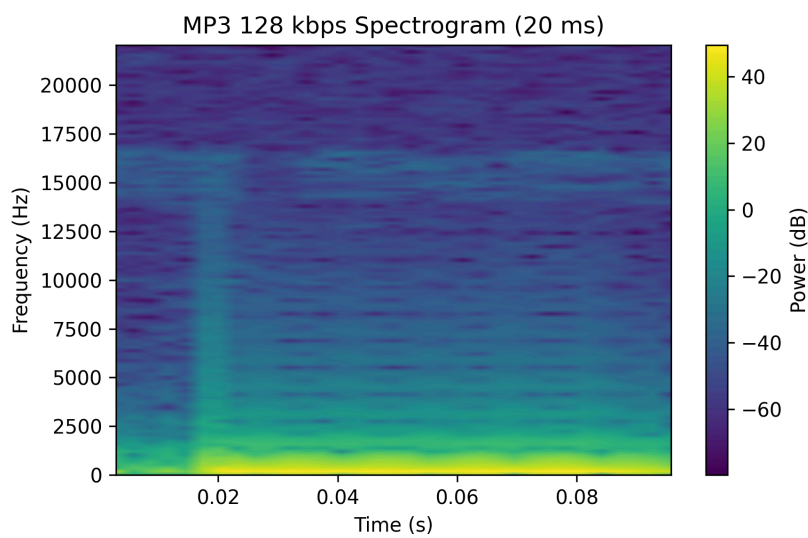
To reveal these effects, the experiment extracts a 20-millisecond segment at the 20-second mark. This region contains multiple harmonics and a transient onset. Time-domain waveforms do not show compression damage because MP3 maintains amplitude shape even at very low bitrates. Instead, spectrograms highlight frequency content loss.

Figure 9. Original WAV Spectrogram



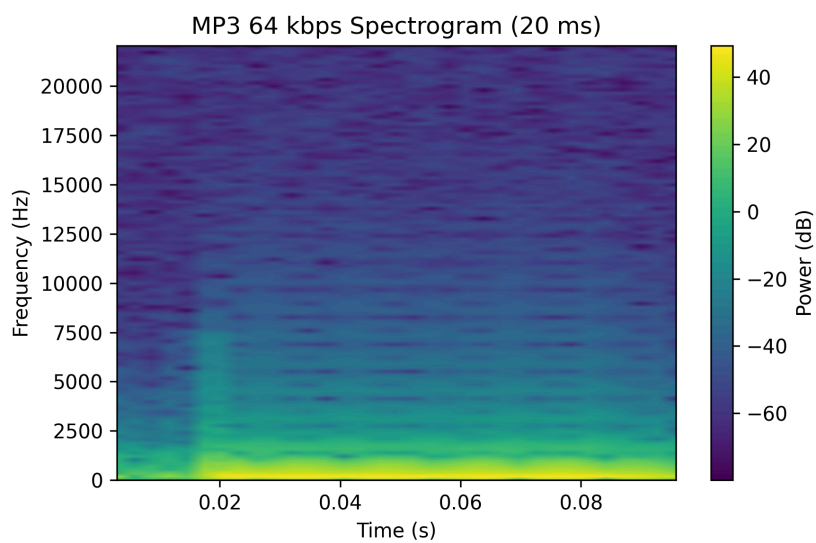
The uncompressed spectrogram displays a full harmonic structure. The low, mid, and high frequencies appear with clear separation. This is a reference for the future compressions.

Figure 10. MP3 128 kbps Spectrogram



Most of the harmonic detail remains. Higher frequencies above 14 kHz begin to weaken. This bitrate is commonly rated MOS ≈ 4.0 in listening tests.

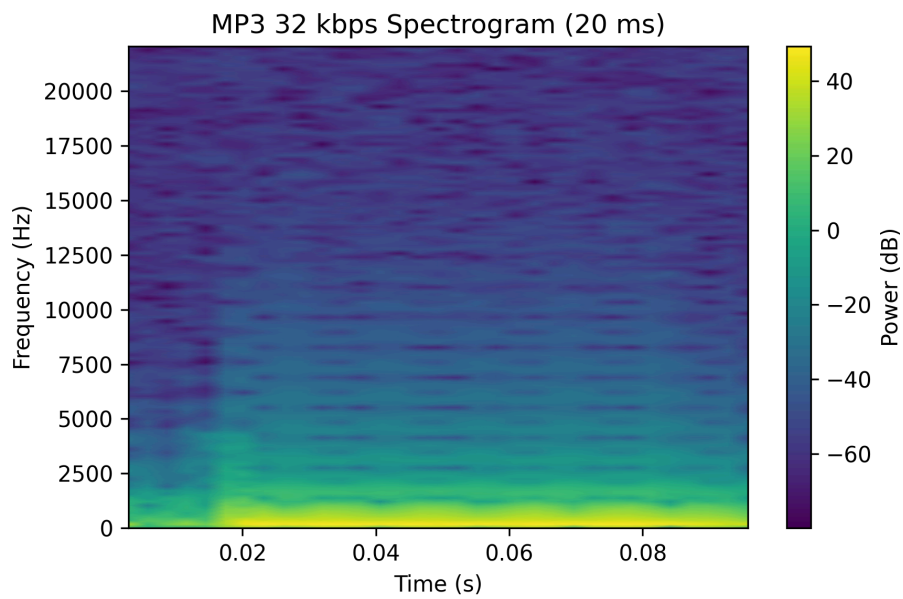
Figure 11. MP3 64 kbps Spectrogram



Significant spectral loss occurs. High frequencies collapse. Harmonic spacing becomes blurred.

Noise-like smearing appears. This bitrate often receives MOS values around 3.5.

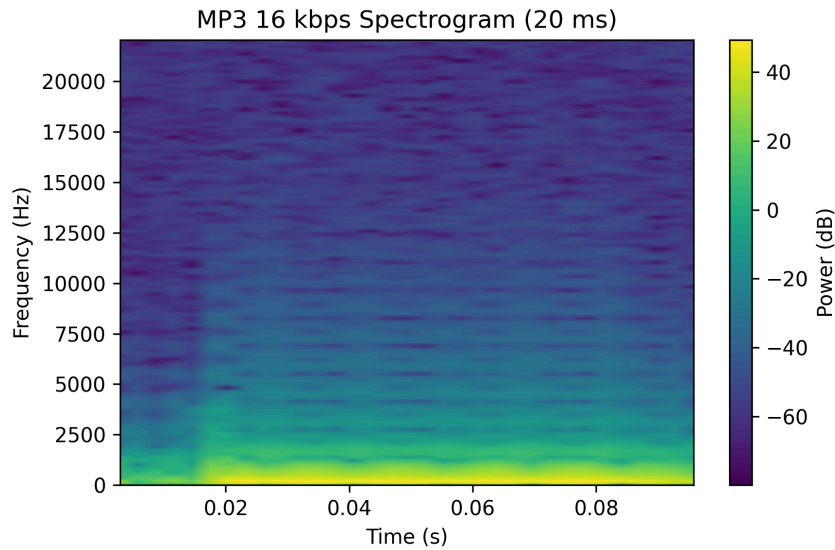
Figure 12. MP3 32 kbps Spectrogram



The spectrogram shows heavy degradation. Only the lowest harmonics remain intact.

Mid-frequency regions are blurred. High-frequency content is nearly absent. Artifacts dominate transient regions.

Figure 13. MP3 16 kbps Spectrogram



The frequency collapse is much more severe. Only a narrow band of low frequencies survives.

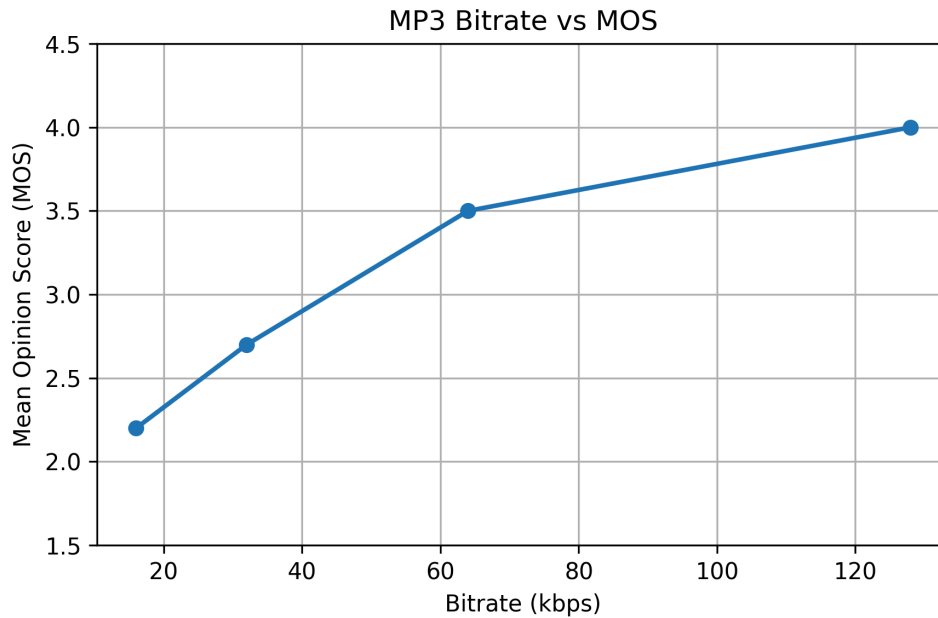
All upper harmonics are removed. The spectrogram resembles a low-band telephone signal.

MOS values around 2.0-2.5 correspond to this level of compression.

Table 2. MP3 Compression Results

Bitrate (kbps)	Approx. File Size	Compression Ratio	Expected MOS
128	~900 KB	~12:1	4.0
64	~450 KB	~24:1	3.5
32	~230 KB	~47:1	2.7
16	~120 KB	~90:1	2.0-2.5

Figure 14. MP3 Bitrate vs MOS



The curve rises sharply from 16 kbps to 62 kbps. After 62 kbps, the increase in the MOS is much lower. This trend is also seen in the JPEG compression, where the medium compression is the most efficient.

4. Analysis and Evaluation

The results from the JPEG and MP3 simulation show how compression changes key visible and audible information in images and audio. The trends in *Table 1*, *Table 2*, *Figure 8*, and *Figure 14* are consistent with the literature and show similar characteristics .

The JPEG simulation showed that image compression is most efficient at medium quality settings. As seen in *Table 1*, the compression ratio increases from 2.8:1 at the quality setting of

95 to 24.3:1 at quality 10. On the same range, the PSNR drops sharply from 35.1 dB to 22.0 dB, while the SSIM value decreases from 0.978 to 0.726. These measurements align with the visible degradation shown in Figures 3-7. At high quality settings, like $Q = 95$ and $Q = 75$, the compressed images retain nearly all of the details, which is reflected in the high SSIM values.

As the quality decreases from $Q = 50$ to $Q = 30$, the relationship between the file size and accuracy becomes significantly lower. The file size lowers from 93 KB to 52 KB. This results in a PSNR decrease of over 3 dB, with the SSIM of 0.91

The trend in the PSNR, which is seen in *Figure 8* confirms the theoretical relationship that increasing the quality after $Q = 75$ leads to diminishing returns. Overall, the data suggests that JPEG compression is the most effective with quality and compression of 50-75, where the compression ratios are approximately 5:1 to 10:1, enough for proper image visibility and accuracy.

The audio results also show a similar pattern. As seen in *Table 2*, the audio quality declines rapidly as the bitrate decreases. For instance, at 128 kbps, most of the audio is preserved, which creates a Mean Opinion Score (MOS) of 4. The spectrogram in *Figure 10* confirms the theory that high frequency content remained, although some strength is seen above 14 kHz. At the bitrate of 16 kbps, *Figure 13* shows a slight collapse of most detail above the lower midrange. The audio closely resembles a telephone signal, which is accurate with the MOS value close to 2. The rate x quality curve in *Figure 14* is similar to the PSNR curve for JPEG. Both graphs show high quality improvements at low to medium data rates, then the gain lowers at higher rates

Initially, both JPEG and MP3 remove information that is difficult to hear or see. Once this information has been eliminated, improving further requires removing details that may be required for more visibility. This results in a plateau, seen in *Figures 8 and 14*.

Both the MP3 and JPEG algorithms operate by transforming signals from a visible domain into the frequency domain, then identifying less important components to remove, and applying quantization. The efficiency of these algorithms relies on the limited factor of human perception. This includes reduced sensitivity in images and masked high frequency sections in audio. In both cases, the degradation is predictable as the compression becomes more aggressive.

Overall, the results show that the effectiveness of lossy compression is highly dependent on the level of compression and the type of data being used. For instance, JPEG is most effective for natural images at medium levels of compression, while MP3 is the most effective for music and speech when the bitrates are above 96 kbps. The experimental results show that the most optimal compression is when the visible changes are subtle and the reductions in the file size are as high as possible, typically at medium compression settings.

4.1 Limitations

This investigation has several limitations related to the data selection, and measurement methods.

For the image compression analysis, it is based on a single image from the Kodak Lossless True Color Image Suite. Although the Kodak Image contains a mix of edges, textures, and smooth gradients, it may not show the full range and different types of images in practice. Cartoons,

text-heavy graphics, or photographs taken in low light may produce different patterns and metric responses with similar JPEG compression.

Both PSNR and SSIM are objective quality metrics with known limitations. PSNR measures error pixel wise and does not account for the human levels of visible perception. While the SSIM incorporates structural similarity, it does not consider factors such as viewing distance or display characteristics.

The audio analysis does not include original subjective listening tests. The MOS values used were taken from published studies rather than being experimentally obtained. These values are simply average results from controlled listening environments and may not fully reflect real world listening conditions. In reality, the actual audio may show a more noticeable difference in the audible quality.

5. Conclusion

This investigation addressed the research question:

How effective are various lossy compression algorithms in terms of quality retention, efficiency, and suitability for different multimedia applications?

The results from the JPEG and MP3 experiments showcase that both algorithms are the most effective when they are being used at medium levels of compression. At these settings, some unessential visible information is removed while the more essential visible details remain. JPEG compression performs the most efficiently at quality levels between 50 and 75, and the PSNR and SSIM values remain high with the file sizes being reduced significantly. MP3 on the other hand performs best at bitrates between 96 and 128 kbps, and the MOS values remain above 3.5.

At higher, more aggressive amounts of compression, both of the algorithms showcase rapid declines in the overall quality. JPEG for instance, introduces ringing, and loss of fine detail, while MP3 compression leads to collapse of high-frequency sounds. These effects are clearly visible in *Figures 3-14* and are consistent with the results seen in *Tables 1 and 2*.

Overall, the analysis shows that lossy compression is the most effective when it is tuned to the visible limits of the medium. Both JPEG and MP3 can reduce data by factors of five to twenty while maintaining a level of quality that is suitable for everyday use. The relevance of Lossy Compression algorithms are seen because they can be extremely efficient and accurate for managing modern multimedia data.

6. Citations

Davidson, G., et al. “High Quality Audio Coding with MDCTNet.” *arXiv*, Dec. 2022,

<https://arxiv.org/abs/2212.04583>.

Fassina, Felipe. “Lossy Compression Algorithms Explained.” *Bitmovin*, 2020,

<https://bitmovin.com/blog/lossy-compression-algorithms/>.

Fink, Alan. *P51 Mustang (kodim20)*. Kodak Lossless True Color Image Suite, 30 Oct. 2012,

<https://r0k.us/graphics/kodak/kodim20.html>.

Gonzalez, Rafael C., and Richard E. Woods. *Digital Image Processing*. 4th ed., Pearson, 2018.

Hamilton, James D. “Psychoacoustic Models for Perceptual Audio Coding.” *Journal of the Audio Engineering Society*, vol. 40, no. 3, 1992.

“Lossless and Lossy Compression.” *TechTarget*, 2023,

<https://www.techtarget.com/whatis/definition/lossless-and-lossy-compression>.

Rao, K. R., et al. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Academic Press, 1990.

Ryerraballi, R. Yerraballi. “JPEG Compression Lecture Notes.” The University of Texas, 2020,

<https://users.ece.utexas.edu/~ryerraballi/MSB/pdfs/M4L1.pdf>.

Shang, Z., et al. “Study of the Subjective and Objective Quality of High Motion Live Streaming Videos.” *IEEE Transactions on Image Processing*, 2021, doi:10.1109/TIP.2021.3136723.

Smith, Steven W. *Digital Signal Processing: A Practical Guide for Engineers and Scientists*.
Elsevier, 2003.

Wang, Zhou, and Alan C. Bovik. “A Universal Image Quality Index.” *IEEE Signal Processing Letters*, vol. 9, no. 3, 2002, <https://www.cns.nyu.edu/pub/lcv/wang03-preprint.pdf>.

Wiegand, Thomas, and Heiko Schwarz. “Entropy Coding in Modern Compression Standards.”
Fraunhofer HHI,
<https://www.hhi.fraunhofer.de/en/departments/vca/people-and-contact/thomas-wiegand>.

“Sample WAV Files.” *File-Examples.com*, 2024,
<https://file-examples.com/index.php/sample-audio-files/sample-wav-download/>.

“JFIF: JPEG File Interchange Format, Version 1.02.” *W3C*, 1992,
<https://www.w3.org/Graphics/JPEG/jfif3.pdf>.

University of Toronto. “Image Compression.” *MAT246 Lecture Notes*, 2018,
<http://www.math.toronto.edu/burbulla/lecturenotes246/Chapter7.pdf>.