Principles of MPEG Compression

The human ear perceives sound in the frequency range of 20 Hz to 20 kHz, with the most sensitive range being between 2-4 kHz (note: a tuning fork vibrates at a frequency of 440 Hz). A key characteristic of human hearing is that when two sounds of similar frequencies are played simultaneously, the one with higher intensity masks the one with lower intensity. This masking effect also occurs if the lower-intensity sound starts 2 ms earlier and ends approximately 15 ms later than the higher-intensity sound of a nearby frequency.

Quantization should be performed with particular attention to the fact that frequencies between 2-4 kHz require higher precision than those in the 10-20 kHz range.

MPEG compression offers various methods for both video and audio. In video compression, MPEG takes into account changes between individual frames. It defines keyframes (I-frames) that compress the entire frame using the previously known JPEG method, while intermediate frames (P-frames, B-frames) contain information about changes. This is particularly important during fast-forward operations, as a full frame for a P-frame can only be reconstructed starting from the preceding I-frame since P-frames do not contain the complete image, only the changes.

In MPEG, it is crucial to keep audio and video synchronized over time, so both audio and video include timestamps.

Additional Information on MPEG Compression

MPEG (Moving Picture Experts Group) compression is a widely used standard for reducing the file size of video and audio content while maintaining an acceptable level of quality. It achieves high compression ratios by eliminating redundant information both spatially within a frame and temporally between consecutive frames.

Temporal Redundancy Removal: Temporal redundancy exists because consecutive video frames often have many similarities. MPEG compression uses motion compensation to identify and encode only the differences between frames, instead of storing every frame in full detail. This is done by using the following frame types:

1. I-frames (Intra-coded frames): These are complete images, compressed independently of other frames, much like a JPEG image.

1. P-frames (Predicted frames): These store only the changes from the preceding I-frame or P-frame, thus reducing the amount of data required.

1. B-frames (Bidirectional frames): These use data from both preceding and succeeding frames, which allows for even greater compression by predicting changes from both directions.

Spatial Redundancy Removal: Spatial redundancy within a single frame is reduced using techniques similar to JPEG compression, including Discrete Cosine Transform (DCT) to convert spatial domain data into frequency domain data, allowing for more efficient quantization and compression.

Audio Compression: MPEG audio compression, such as MPEG-1 Layer 3 (commonly known as MP3), also takes advantage of the human auditory system's limitations. By discarding inaudible parts of the sound, MPEG audio can significantly reduce the bit rate without a perceivable loss in quality.

Psychoacoustic models are used to identify which parts of the audio can be removed or more coarsely quantized without affecting the listening experience.

Applications: MPEG compression is used in various multimedia applications, including digital television, video streaming, DVD, and Blu-ray discs. Its ability to maintain synchronization between audio and video streams makes it suitable for both broadcasting and interactive media applications.

Container Types for MPEG Compression

MPEG compression is often used in conjunction with different container formats that store video, audio, subtitles, and other metadata. Some common container formats include:

From:
https://edu.iit.uni-miskolc.hu/ - **Institute of Information Science - University of Miskolc**

Permanent link:
https://edu.iit.uni-miskolc.hu/tanszek:oktatas:techcomm:mpeg_compression?rev=1728295675

Last update: **2024/10/07 10:07**