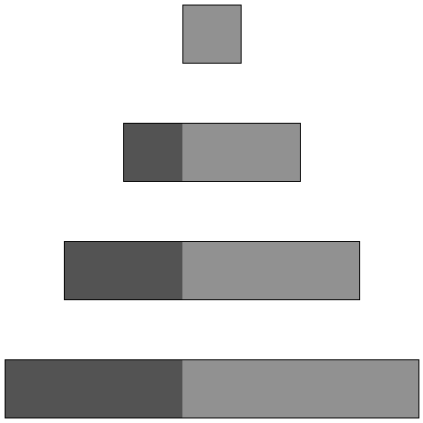
DIGITAL SUBBAND VIDEO

ENCODER INFORMATION

April, 2024



2023-2024 EMMIR
ENVEL GRAPHICS

Contact Information:

GitHub : https://github.com/LMP88959
YouTube: https://www.youtube.com/@EMMIR_KC/videos
Discord: https://discord.com/invite/hdYctSmyQJ

This document describes design decisions and features of the Envel Graphics DSV1 encoder.

A. Motion Estimation

Full-Pixel Motion

The encoder uses hierarchical motion estimation (HME) to derive full-pixel motion vectors. HME is the "best of both worlds" in terms of performance and accuracy. It effectively captures high and low motion like a full search yet performs nearly as fast as a traditional speedy search like the three-step-search.

Half-Pixel Motion

Half-pixel motion vectors are derived using a fixed size half-pixel block search for speed. The search is a full search around the pixel's eight neighbors but only checks a fixed size square in the center of the test block.

Intra Block Determination

Effective intra block determination is absolutely crucial to getting good looking video at low to medium bitrates. The encoder computes simple Human Visual System (HVS) based statistics on the block's luma and chroma to determine if it should be an intra block or an inter block.
The following statistics are computed:

- Squared Luma Variance
  - Max Chroma Variance - max between U and V
- Luma Texture - average of horizontal and vertical differences in the image
- Luma Average

The HVS tells us that error is mostly visible in flat/low texture regions, and so that is what intra block determination seeks to remove.

Intra Sub-block Determination

The DSV1 specification allows for an intra block to be split into 4 equal sub-blocks where each sub-block is either intra coded or inter coded with a zero motion vector. This is advantageous for scenes where there are 2D overlays on the video, like the score overlay in a football match. In cases where there is very high motion in the scene, these sub-blocks can be used to maintain a clear image of the score at the cost of accuracy around the edges of the overlay, which are generally not too visually important. The encoder computes an "intra metric" to help capture these static portions of the video. This metric is essentially a combination of texture and block similarity. A good candidate would be a sub-block that has high texture and a reference block that gives extremely accurate prediction for a majority of the sub-block.

B. Stability Blocks

The encoder keeps track of the average motion of every block over a certain refresh period. Blocks who have an average of approximately zero over the period are marked as stable. Another factor that can mark a block as stable is if it is determined to be "high detail." The encoder marks any block that has sufficient texture and variance as high detail. The threshold for the texture and variance increases if the left, top, and top left neighbors of the block were also high detail. This is also HVS based since detailed objects are more noticeable if they are surrounded by low detail regions and less noticeable if they are lost in a sea of similarly high detail blocks. If the block is low variance or low texture the stability tracking for the block gets reset in order to avoid wasting bits trying to keep low importance blocks high quality.

C. Scene Change Detection

The encoder uses a simple method of scene change detection to help reduce visual artifacts when the video seemingly changes contents greatly from frame to frame. The encoder simply computes the average luma of the frame and compares it to the average luma of the previous frame. If the difference between the two is large enough, an intra frame is inserted instead of an inter frame.

D. Rate Control

For single pass average bitrate, the encoder knows how many bytes each frame needs on average to satisfy the given rate. It uses a simple control loop to ensure the video approximately hits its bitrate target without heavy oscillation or instability.
The loop utilizes a few pieces of information:

- ***avgbpf*** = average bytes per frame (BPF) so far
- ***isP*** = whether the current frame is intra or inter
- ***forced*** = whether the current frame was forced to be intra and not just part of the GOP cycle
- ***over*** = whether the previous inter frame was significantly over the target BPF
- ***under*** = whether the previous inter frame was significantly under the target BPF

The control loop finds how far (as a fraction called delta) the current ***avgbpf*** is from the target BPF and adjust the quality by delta. If the rc_hmnudge option is enabled, the ***over*** and ***under*** metrics are used to push the delta value harder to compensate. If the intra frame was ***forced*** then the quality potentially gets bumped a little to allow a new scene to start out with more detail. If the frame ***isP*** then its lower quality bound is determined by the average quality of the inter frames minus four percent. This helps prevent the bit starvation that intra frames tend to impose on subsequent inter frames.