Subjective HEVC Transcoding Quality

A comparison of:

Intel SVT-HEVC

NGCodec FPGA-based HEVC encoder (acquired by Xilinx)

NETINT Codensity T408 Transcoder

x.265 Medium and Veryfast Presets

By:

Jan Ozer
Streaming Learning Center

Report date: December 12, 2019

Executive Summary:

Using a service called Subjectify, we crowd-sourced subjective video comparisons of output from the NETINT Codensity T408 Transcoder, Intel's SVT-HEVC codec; NGCodec's FPGA-based HEVC encoder; and x265 using the medium and veryfast presets. The results below reflect 2508 pairwise judgments within which the NETINT (ni265) clip was compared to the other technologies 1179 times. With four videos and four competitors, this means that the NETINT technology was compared to the others about 74 times per clip, per competitor. As shown in Figure 1, NETINT (ni265) was preferred by a significant margin over all other codecs.

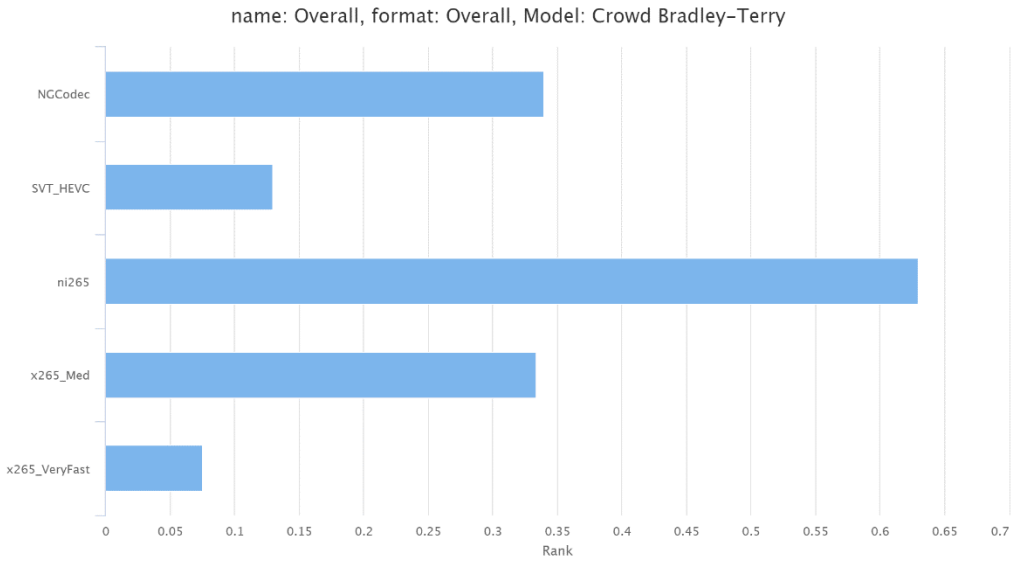

Figure 1: Overall results using the Crowd Bradley-Terry model.

Overview:

On September 11, 2019, *Streaming Media Magazine* published an article entitled Hardware-Based Transcoding Solutions Roundup: Testing Performance (bit.ly/hw_transcode) which compared various H.264 and HEVC encoders using objective and subjective testing. The HEVC encoders tested included Intel's SVT (Scalable Video Technology)-HEVC, a software-based codec; NGCodec's FPGA-based HEVC encoder; and x265 using the medium and veryfast presets.

The article details the scripts used for testing and other test conditions. Scripts and testbeds for the Intel and NGCodec products were supplied by both companies after full disclosure of the FFmpeg-based x265 scripts. The specific versions tested were not recorded but were the most current versions available on or around June 1, 2019.

The tests involved four test clips, Dinnerscene (Netflix), Football (Harmonic), GTAV (Twitch), and Meridian (Netflix), all 2-minutes in length and tested at 1080p60. The HEVC comparisons were output at 1 Mbps, 2 Mbps, 3 Mbps, and 4 Mbps, with subjective testing performed by Subjectify using the first 20-seconds of the 3 Mbps of each clip.

After the *Streaming Media* article published, NETINT approached the author of the article, Jan Ozer, and asked him to compare output from its Codensity T408 Transcoder to the three HEVC codecs tested in the article. NETINT supplied a testbed and created the appropriate scripts following the

procedures outlined in the article and Ozer ran the tests. NETINT tested firmware version 1.2 of their encoder.

This paper presents the subjective comparisons from this analysis.

Subjectify Overview

Subjectify (http://www.subjectify.us/) is a web platform for crowd sourcing subjective still image and video comparisons. The site was created by Moscow State University to support its codec comparisons and comparison tools. In terms of process, you supply Subjectify with comparison videos and Subjectify recruits and compensates viewers on the web. Each test session involves no more than ten A:B comparisons with verification checks inserted to ensure that the remote testers are paying attention.

During the NETINT testing, Subjectify ran 2508 pairwise judgments within which the NETINT (ni265) clip was compared to the other technologies 1179 times. With four videos and four competitors, this means that the NETINT technology was compared to the others about 74 times per clip, per competitor.

This report presents the results by clip with a short summary at the end.

Subjectify Results

Dinnerscene (Figure 2) is a low motion clip of a couple eating dinner with their child or grandchild. There is some slow camera panning and cuts between camera angles, but overall this is an easy-to-encode clip.

Figure 2. **The Dinnerscene test clip from Netflix.**

Figure 3 shows the results of the round-robin comparison of the Dinnerscene clip using the Crowd Bradley-Terry model, which, according to this paper, "extends the widely used Bradley-Terry model by explicitly incorporating the quality of contributions provided by different annotators" (bit.ly/crowd_BT). The NETINT (ni265) clip was preferred overall by all testers by a very significant margin over NGCodec (now Xilinx) and SVT_HEVC.

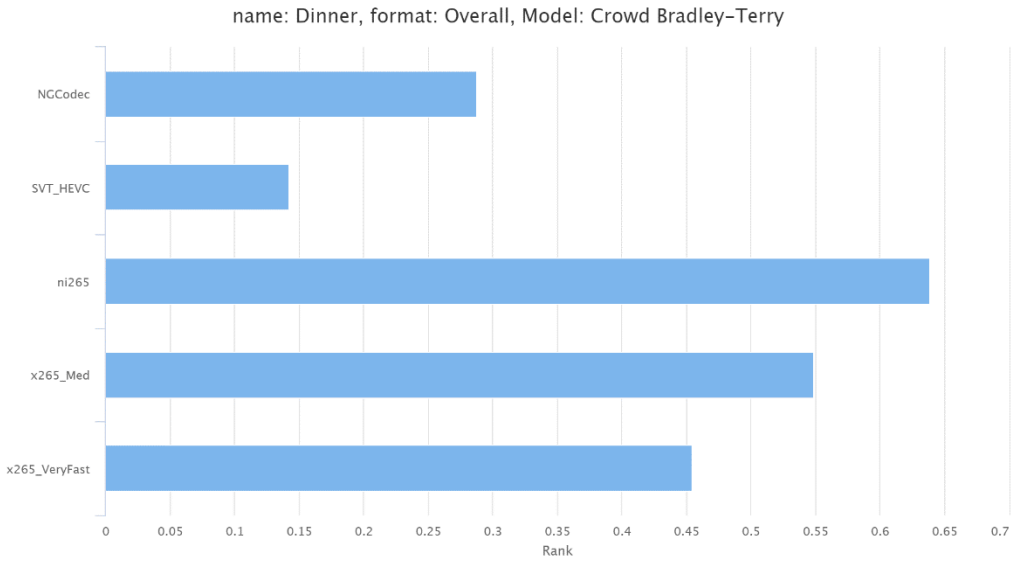

Figure 3: Results from the Dinnerscene round-robin comparison using the Crowd Bradley-Terry model.

Football is a very high-motion clip of a football game shot with multiple cameras. Views include full formation views at the start of the play and with the camera zooming in to follow the action. The players are never more than 100 - 120 pixels high so this clip contains both high-motion and significant detail (Figure 4).

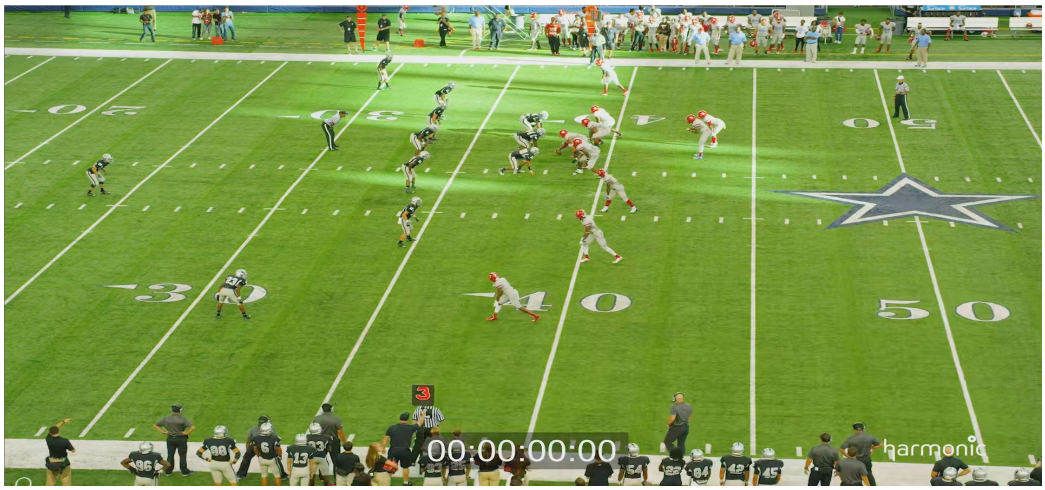

Figure 4. The Football test clip from Harmonic.

Figure 5 shows the results of the round-robin comparison of the Football clip with the NETINT (ni265) clip preferred overall by all testers by a very significant margin over NGCodec (now Xilinx) and SVT_HEVC. Given the challenging nature of the content these results are particularly impressive.

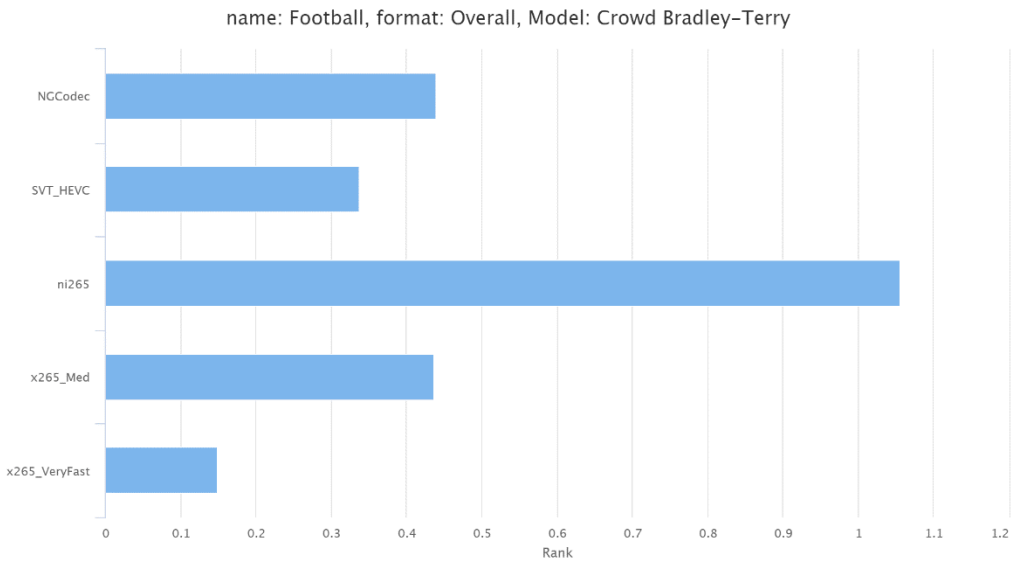

Figure 5: Results from the Football round-robin comparison using the Crowd Bradley-Terry model.

GTAV is a clip from computer game Grand Theft Audio which is very high motion (Figure 6).

Figure 6. The GTAV test clip from Twitch.

Figure 7 shows the results of the round-robin comparison of the GTAV clip where the NGCodec edged NETINT (ni265) slightly with the other codecs well behind.

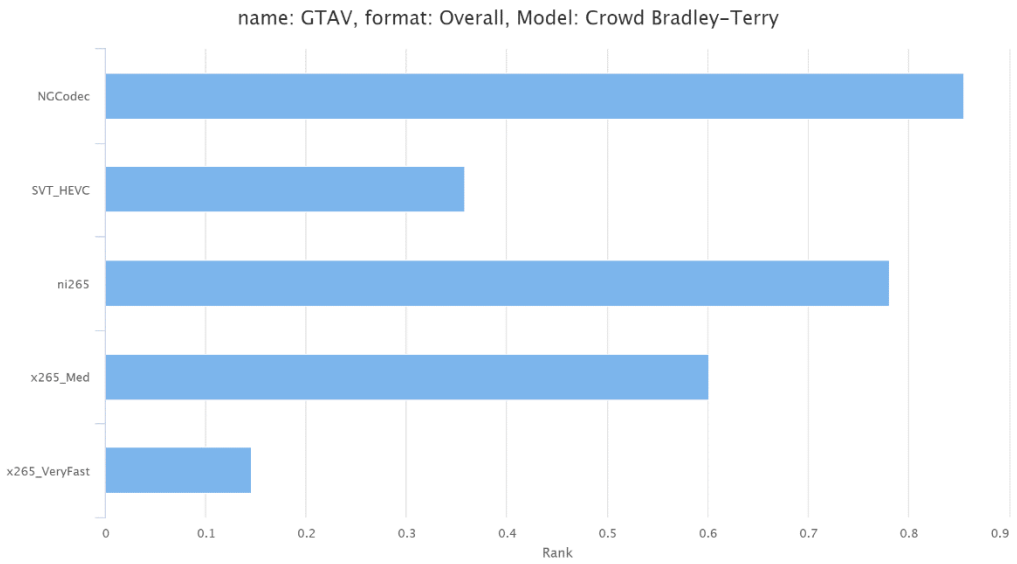

Figure 7: Results from the Twitch round-robin comparison using the Crowd Bradley-Terry model.

Meridian is a mixed motion test clip from Netflix that simulates a film noir detective movie shot in the '40s. The portion tested is relatively simple to encode, though there is somewhat complex lighting as shown in Figure 8.

Figure 8: The Meridian test clip from Netflix.

Figure 9 shows the results of the round-robin comparison of the Meridian clip with the NETINT (ni265) clip again preferred overall.

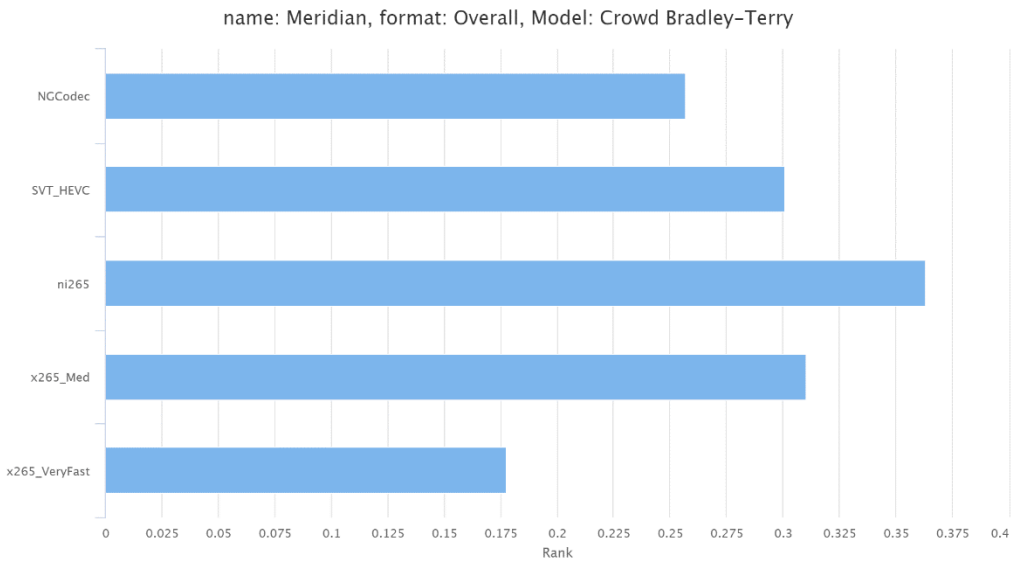

Figure 9: Results from the Meridian round-robin comparison using the Crowd Bradley-Terry model.

Figure 10 shows the overall rankings for all clips. Not surprisingly given the individual results, NETINC was preferred by a wide margin.

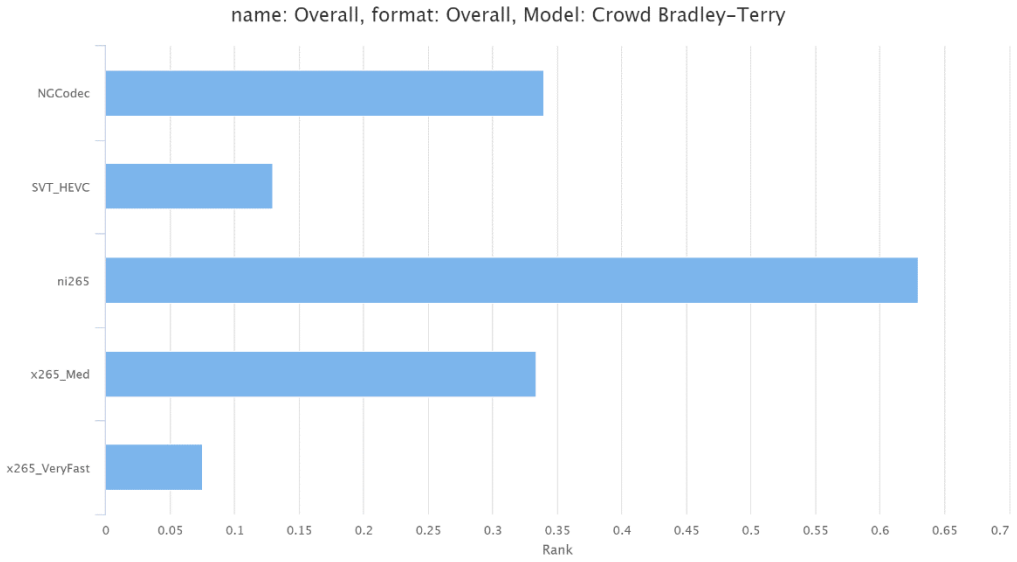

Figure 10: Overall results using the Crowd Bradley-Terry model.

What This Means

Most large-scale transcoding shops need a dense hardware solution which takes software-based x265 encoding off the table. Intel's SVT-HEVC codec is also software-based and performed poorly here, so currently doesn't appear to be a viable option.

Accordingly, it comes down to a two-horse race between NETINT and the FPGA-based NGCodec solution, now owned by Xilinx. While choosing between the two will involve many factors, from a quality perspective, most viewers preferred NETINT over NGCodec/Xilinx. The only category that

Xilinx won was computer gaming, perhaps not surprising given that one of NGCodec's most significant targets and design wins was Twitch, though this was for the VP9 codec not HEVC. Otherwise, NETINT prevailed over NGCodec by a significant margin, with particularly impressive results in the hardest to encode Football clip.