Beamr Video Technology Assessment

Jan Ozer
jozer@mindspring.com
276-238-9135
Doceo Publishing

WHITE PAPER: BEAMR VIDEO TECHNOLOGY ASSESSMENT

This document presents my assessment of Beamr Video, an automated video optimization technology. It starts with an overview of the technology and my analysis, shows the data rate reductions the technology achieved, and then discusses output quality.

Section 4 discusses technical and workflow considerations, followed by a short conclusion.

Contents

TECHNOLOGY ASSESSMENT – OVERVIEW

Introduction

Reducing video bandwidth provides many benefits, including cost reductions, the ability to send higher quality video to devices on lower connection speeds, and to pack more streams into a fixed bandwidth pipe. However, reducing file bandwidth is complicated because all videos are different and compress differently. Video encoding parameters that produce the optimum blend of quality and data rate with one video could produce a blocky mess with another.

In a perfect world, video producers could reduce bandwidth costs by optimizing encoding parameters for each video. Unfortunately, the costs of the time involved would likely consume any associated bandwidth savings.

Beamr Video is designed to automate that process, specifically to output videos with the optimum blend of data rate efficiency and video quality, or, in marketing speak, videos that look great at the lowest possible data rate. Technically, Beamer Video is a post-encoding process performed on previously encoded files; first you encode, then you optimize in Beamr Video, a command-line driven process that runs on Ubuntu 12.04 and RHEL 6.x, which outputs industry standard H.264 files.

From a quality perspective, Beamr Video can run in two modes, High and Best. Here are the descriptions according to the user manual.

- **Quality Best** – Provides the best output video quality possible. In this mode, the output video stream will be perceptually identical the input video stream, even when viewed by an expert viewer.

- **Quality High** -- Produces an output video stream with a lower bitrate than quality Best, which is perceptually identical to the input video stream when viewed by an average user.

In essence, this technology assessment is designed to prove or disprove these claims. By design, this assessment uses a blend of objective and subjective comparisons to achieve this goal, as designed in the following section.

Test Description

This test uses two files encoded to six different profiles using two different encoding tools. These processed were then processed by Beamr Video using the Best and High options.

TECHNOLOGY ASSESSMENT – OVERVIEW

The files were:

- **Walk** – The movie trailer for the movie Walk among the Tombstones.
- **New** – A short test clip comprised of various test sequences produced by Jan Ozer for these tests.

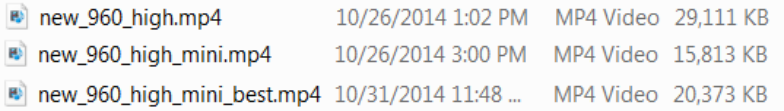
Encoding Parameters

Both clips were encoded to six different configurations using standard configurations supplied by Ozer, which were derived from Apple Tech Note TN2224. These configurations were:

- 1920 resolution @ 7 Mbps (1920_high)
- 1920 resolution @ 5 Mbps (1920_low)
- 1280 resolution @ 4.4 Mbps (1920_high)
- 1280 resolution @ 2.9 Mbps (1920_low)
- 960 resolution @ 2.2 Mbps (1920_high)
- 960 resolution @ 1.6 Mbps (1920_low)

The Walk clips were encoded by Telestream Vantage, while the New clips were encoded by Sorenson Squeeze, in both cases using the x264 codec.

After encoding, the clips were optimized with Beamr Video. Files encoded using the High quality settings have the _mini extension, while files encoded using Best quality have the _mini_best extension. The test file name was appended to the front to create the following three files for each test clip and configuration.



After processing, all files were analyzed using the Moscow State University Video Quality Measurement Tool (VQMT), which was developed in conjunction with Apple, Intel, Disney, Cisco, and Fraunhofer, (among other companies) and produces a range of quality-related metrics like Peak Signal to Noise Ratio (PSNR), Structured Similarity Index (SSIM) and the Video Quality Metric (VQM). While the differences in PSNR and SSIM were insignificant in most tests, the VQM scores did the best job identifying subjective differences between the files.

OBJECTIFYING THE SUBJECTIVE?

Which leads us discussing how to measure the "perceptually identical" standard set up in the product manual, particularly as it relates to the "average" and "expert" viewers. For insight on this issue, the author researched for articles defining the most common video artifacts, and differentiating those most apparent to average and expert viewers. Most relevant was, Video Quality Impairments 101 for MOS's, by Daniel Howard, then of VQLink, Inc (bit.ly/vid_art2). In the article, Howard differentiates between artifacts seen by expert and untrained viewers. A rough taxonomy would be:

- **Untrained viewers** – Blockiness, choppiness.

- **Expert viewers** – blockiness, choppiness, blurriness, mosquito noise, ringing, pulsing (noticeable shift on key frames which he calls background breathing) and interlacing artifacts which are not relevant here.

Another article, Can Small Be Beautiful? Assessing Image Resolution Requirements for Mobile TV (multiple authors, bit.ly/vid_art3), discusses that average viewers complain most about the loss of detail in text, probably because it's the most apparent, with loss of object and facial detail also common complaints.

Another article, Compression Artifacts in Modern Video Coding and State-of-the-Art Means of Compensation, by Andreas Underweger (bit.ly/vid_art1), added the familiar Stair Case, or jaggies artifact.

Other common compression artifacts include:

- **Bleeding** - where highly saturated colors bleed from the source object to the background, though in my experience, these artifacts are distracting and apparent to most viewers.

- **Banding** – where colors form visible bands in the video, usually in the sky or other smooth regions.

- **Background distortion** – either moving or still image distortion obvious in smooth backgrounds.

The author divided these artifacts into those seen by average and expert viewers in the table on the next page. To assess Beamr's perceptually identical claim, the author played the videos at their original resolution and at full speed.

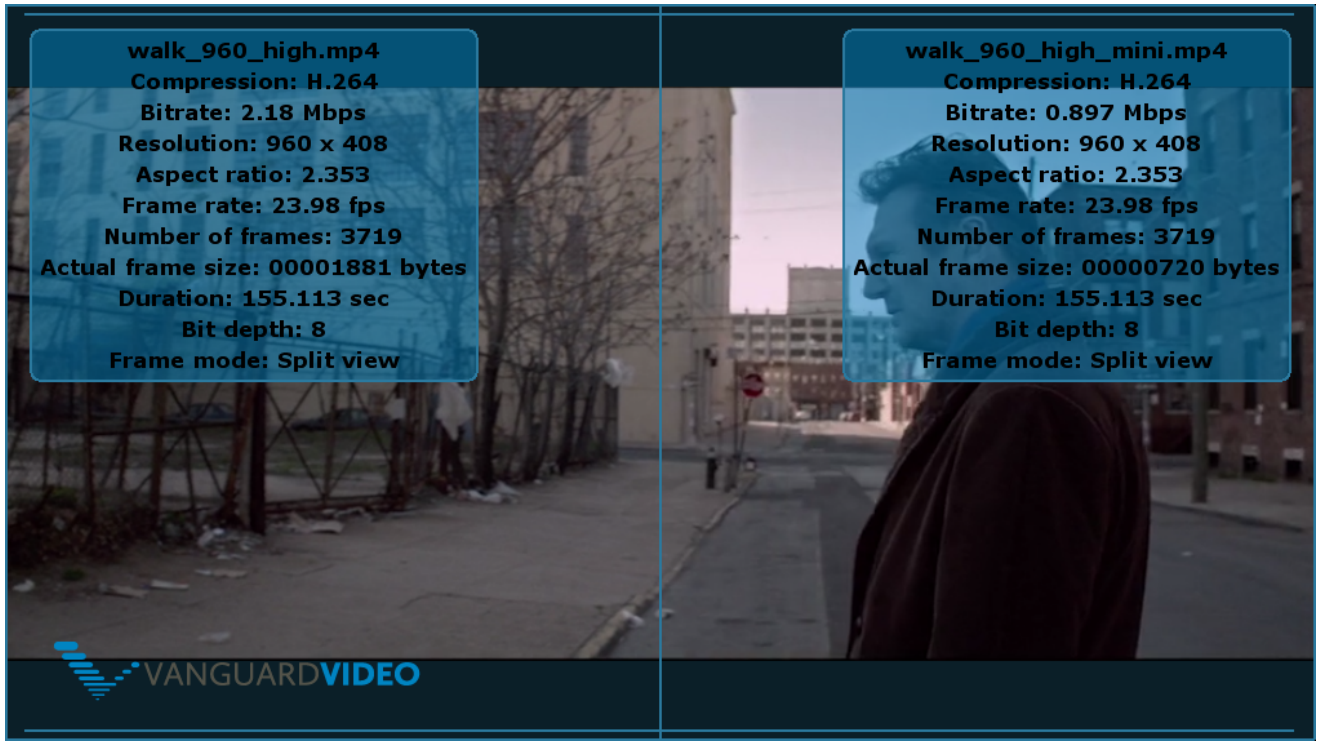
OBJECTIFYING THE SUBJECTIVE

With the Beamr videos processed using the High setting, the author noted any artifacts in the Average column. In Beamr videos processed using the Best setting, the author noted any artifacts in the Expert column.

In both cases, when artifacts were found, the author viewed the compressed source clip to determine if the artifact originated in that clip. If not present in the source, this was noted as an artifact.

Artifact	Average	Expert
Blockiness	X	X
Choppiness	X	X
Loss of text detail	X	X
Loss of object/facial detail	X	X
Jaggies	X	X
Bleeding	X	X
Background distortion	X	X
Mosquitoes		X
Ringing		X
Pulsing		X
Banding		X

ABOUT SUBJECTIVE TESTING



I performed all subjective tests on an HP Z800 workstation using a 31" HP LP3065 monitor driven by an NVIDIA Quadro FX 4800 graphics card. Monitor resolution was adjusted to maximize video display for each clip as follows:

- 960 video resolution – 1280 x 800 monitor resolution
- 1280 video resolution – 1680 x 1050 monitor resolution
- 1920 video resolution – 1920 x 1250 monitor resolution

I played the videos in the Vanguard Video Visual Comparison Tool. As shown above, this allowed me to load two videos into the display, and drag the middle line to view all or a portion of either clip. For example, if I discovered an artifact in the Beamr High clip above (walk_960_high_mini.mp4), I could drag the line to the right to see if the artifact existed in the compressed source (walk_960_high.mp4).

I performed all subjective tests at full resolution and at 100% playback speed.

DATA RATE REDUCTIONS

Let's start with a quick look at the data rate reductions delivered by Beamr in High and Best modes, which are fairly uniform. As a reminder, the Walk clip is a movie trailer with a very selective depth of field, and therefore limited detail, and the New clip is a general-purpose test clip with multiple scenes containing varying amounts of motion and detail, with all detail being in focus in each scene.

As shown on the right, Beamr produced significant data rate reductions in both clips. Specifically, at the High settings, the reductions were 50% for Walk and New. At the Best settings, the reductions were 33% for Walk, and 35% for New. As you would expect, reductions were greater in the higher data rate versions of each clip, with larger resolutions producing greater efficiencies.

In the Walk clip, the Best setting produced files that were 36% larger than the High setting, while the New clip produced a 32% differential.

With this as background, let's jump into the clip by clip analysis, starting with A Walk Among the Tombstones trailer.

		Quality High		Quality Best	
Walk	**Original**	**Beamr**	**Reduction**	**Beamr**	**Reduction**
960 low (1,600 kbps)	1,586	811	49%	1,086	32%
960 high (2,200 kbps)	2,183	897	59%	1,259	42%
1280 low (2,900 kbps)	2,878	1,544	46%	2,096	27%
1280 high (4,400 kbps)	4,365	1,834	58%	2,588	41%
1920 low (5,000 kbps)	4,953	2,910	41%	3,794	23%
1920 high (7,000 kbps)	6,934	3,527	49%	4,808	31%
Average	**22,899**	**11,523**	**50%**	**15,631**	**33%**
High over Best			**36%**		

		Quality High		Quality Best	
New	**Original**	**Beamr**	**Reduction**	**Beamr**	**Reduction**
960 low (1,600 kbps)	1,579	1,021	35%	1,247	21%
960 high (2,200 kbps)	2,379	1,243	48%	1,632	31%
1280 low (2,900 kbps)	2,846	1,559	45%	1,971	31%
1280 high (4,400 kbps)	4,702	1,674	64%	2,377	49%
1920 low (5,000 kbps)	5,020	2,566	49%	3,396	32%
1920 high (7,000 kbps)	6,936	2,986	57%	3,907	44%
Average	**23,462**	**11,049**	**50%**	**14,530**	**35%**
High over Best			**32%**		

A WALK AMONG THE TOMBSTONES

Overview

The Walk Among the Tombstones trailer is comprised primarily of scenes from the movie, which was shot with a large CCD camera that allowed the director to narrowly define the depth of field. As shown above, the film's star, Liam Neeson, is in focus, while the rest of the frame is not. This lack of detail makes movies, and movie trailers, comparatively easy to compress (and optimize) as compared to the typical news, sports and corporate videos where the entire frame is in focus.

In addition, the trailer is fast paced, with camera angles switching every 2-4 seconds, which makes spotting video artifacts very difficult. For these reasons, you can't generalize the results of movie trailers to general purpose clips.

As with each clip, first I'll discuss the objective results, then the subjective ratings.

Analysis

1. *Objective results*
2. *Subjective results*
3. *Summary*

OBJECTIVE RATINGS – COMPARED TO COMPRESSED SOURCE

This table compares the Beamr clips to **the compressed file output by Vantage,** which was, of course, their source clip. The three columns show the metric ratings for Beamr using the High and Best setting, and then the percentage quality improvement produced by the Best setting. The orange box identifies the clip with the lowest (and best) score in each metric.

The VQM metric proved much more sensitive than either SSIM and PSNR, with differentials ranging from 14%-19%, compared to under 1% for SSIM and 3%-4% for PSNR.

Walk Tombstones Trailer

VQM (lower is better)			
Walk Tombstones	**Beamr High**	**Beamr Best**	**Best > High**
960 low	0.47	0.39	17%
960 high	0.39	0.33	16%
1280 low	0.44	0.35	19%
1280 high	0.44	0.37	15%
1920 low	0.46	0.37	19%
1920 high	0.46	0.39	14%

SSIM (higher is better)			
Walk Tombstones	**Beamr High**	**Beamr Best**	**Best > High**
960 low	0.99	0.99	0.3%
960 high	0.99	0.99	0.3%
1280 low	0.99	0.99	0.3%
1280 high	0.99	0.99	0.3%
1920 low	0.99	0.99	0.3%
1920 high	0.99	0.99	0.3%

PSNR (higher is better)			
Walk Tombstones	**Beamr High**	**Beamr Best**	**Best > High**
960 low	44.80	46.51	4%
960 high	44.65	46.29	4%
1280 low	45.72	47.52	4%
1280 high	45.82	47.23	3%
1920 low	45.68	47.17	3%
1920 high	45.68	47.01	3%

A WALK AMONG THE TOMBSTONES – SUBJECTIVE RATINGS

	Average								Expert				
Walk - High	Blocks	Choppiness	Text Detail	Other detail	Jaggies	Bleeding	Background Distortion	Total Average	Mosquitoes	Ringing	Pulsing	Banding	Overall Total
960 low							1	1	2	1			4
960 high							1	1	1	1			3
1280 low							1	1	1	1			3
1280 high							1	1	1	1			3
1920 low								0	1				1
1920 high								0	0				0

	Average								Expert				
Walk - Best	Blocks	Choppiness	Text Detail	Other detail	Jaggies	Bleeding	Background Distortion	Total Average	Mosquitoes	Ringing	Pulsing	Banding	Overall Total
960 low							1	1					1
960 high							1	1					1
1280 low								0					0
1280 high								0					0
1920 low								0					0
1920 high								0					0

Artifacts perceived by average viewers impact the "perceptually identical" claim for clips encoded using the High setting, which is why the Total Average box is highlighted up top. Expert issues reported in those clips are reported for completeness. Technically, Beamr didn't meet the standard for four of six High-encoded clips, and four of six Best clips.

That said, subjective ratings for the Walk clip were extremely good, and you had to look really hard to see any artifacts.

Subjectively, the background distortion in the Average category was a very minor issue, as were the mosquitoes that appeared in several clips. **The reality was that for 99.99% of the expert and average viewers, the Beamr-processed clips would be perceptually identical to their source.**

Summary – Walk Clip

Beamr performed well in this analysis, but the cards were stacked in Beamr's favor. That is:

- There was limited depth of field, so large regions of most frames were out of focus (so easy to compress)
- Scenes switched very frequently, making artifacts hard to spot
- Many scenes were very dark, which also complicated finding artifacts

This makes Walk a great demo clip, but experienced compressionists who aren't in the movie business will know that the results may not translate to their own videos. This is why I included the New clip in the analysis.

NEW CLIP

Overview

The new test clip contains 11 unrelated scenes of original 4K footage. The footage differs from the Walk clip in many fundamental ways, including:

- By design, the entire frame is in focus at all times
- Camera switching is relatively slow, with some scenes persisting for ten seconds or longer
- All scenes are brightly lit and well exposed

Though there is no "one size fits all" test clip, the New clip is much closer to the typical corporate, news, sports, event or similar videos than the Walk test.

Analysis

1. *Objective results*
2. *Subjective results*
3. *Summary*

OBJECTIVE RATINGS – COMPARED TO COMPRESSED SOURCE

This table compares the Beamr clips to their source file, which was the compressed file output by Squeeze. The three columns show the metric ratings for Beamr using the High and Best setting, and then the percentage quality improvement produced by the Best setting. The orange box identifies the clip with the worst score in each metric.

Again, the VQM metric proved much more sensitive than either SSIM and PSNR, with differentials ranging from 16%-32%, compared to under 1% for SSIM and 4%-7% for PSNR.

New Test Video

VQM (lower is better)			
New	Beamr High	Beamr Best	Best > High
960 low	0.591	0.401	32%
960 high	0.733	0.581	21%
1280 low	0.774	0.535	31%
1280 high	0.802	0.669	17%
1920 low	0.745	0.540	28%
1920 high	0.739	0.619	16%

SSIM (higher is better)			
New	Beamr High	Beamr Best	Best > High
960 low	0.98	0.99	0.6%
960 high	0.98	0.99	0.6%
1280 low	0.98	0.99	1.0%
1280 high	0.97	0.98	0.6%
1920 low	0.98	0.99	0.8%
1920 high	0.98	0.98	0.5%

PSNR (higher is better)			
New	Beamr High	Beamr Best	Best > High
960 low	41.55	44.37	7%
960 high	40.42	42.23	4%
1280 low	40.24	43.09	7%
1280 high	39.95	41.48	4%
1920 low	41.30	43.78	6%
1920 high	41.19	42.69	4%

NEW CLIP – SUBJECTIVE RATINGS

	Average								Expert				
Walk - High	Blocks	Choppiness	Text Detail	Other detail	Jaggies	Bleeding	Background Distortion	Total Average	Mosquitoes	Ringing	Pulsing	Banding	Overall Total
960 low								0					0
960 high				2		1		3	3				6
1280 low						1		1					1
1280 high								0					0
1920 low								0					0
1920 high								0					0

	Average								Expert				
Walk - Best	Blocks	Choppiness	Text Detail	Other detail	Jaggies	Bleeding	Background Distortion	Total Average	Mosquitoes	Ringing	Pulsing	Banding	Overall Total
960 low								0					0
960 high								0	1				1
1280 low								0	1				1
1280 high								0	1				1
1920 low								0					0
1920 high								0					0

Beamr didn't meet the "perceptually identical" standard for two of six High-encoded clips, and three of six Best clips. That said, there weren't a lot of problem areas for Beamr, and beyond the 960 high clip, most test clips came very close to the perceptually identical standard. Given the real world nature of the clips, the results were impressive.

If you review the 1280 clips, you'll note multiple aliasing/jaggie artifacts. Note that Squeeze introduced these into the 1280 clips that it produced, so they did not give rise to point reductions in the Beamr clips.

Summary – New Clip

Though technically Beamr didn't meet the "perceptually identical" standard on multiple clips, performance in this real world clip was very good, with all problems localized in very small areas in the video frames.

I added this clip to the analysis because I felt that the Walk movie trailer wasn't representative of real world performance. It turns out that the differences between the Walk clip and this one were very minor.

OTHER ISSUES

Beyond the “perceptually identical” claim discussed on the previous pages, there are two other issues worth briefly noting. I cover these in this final section.

Issues

1. *Data rate control*
2. *Adaptive groups*

DATA RATE CONTROL/ADAPTIVE STREAMING

Data Rate Control

When producing a clip for adaptive streaming, much care and consideration goes into bitrate control. Most conservative pundits recommend CBR, while more aggressive consultants recommend constrained VBR, usually 200% or less. The theory is that variances larger than these could produce varying segment sizes that interrupt the smooth operation of the adaptive streaming mechanisms. In addition, when producing HLS files for distribution to Apple apps, streams that exceed the stated data rate by more than 10% can disqualify the app.

Intuitively, however, when it comes to video data rates, less is always more, since less data is always easier to retrieve. Other than the Apple concern about streams exceeding the specified data rate (which can be handled in the manifest file), there appear to be no hard restrictions by any technology regarding data rate control within a file or files within an adaptive group. So it doesn't appear that the lack of bitrate control within Beamer will "break" any adaptive technology.

Adaptive Streaming

Walk	Original	Beamr
960 low (1,600 kbps)	1,586	811
960 high (2,200 kbps)	2,183	897
1280 low (2,900 kbps)	2,878	1,544
1280 high (4,400 kbps)	4,365	1,834
1920 low (5,000 kbps)	4,953	2,910
1920 high (7,000 kbps)	6,934	3,528

Adaptive groups are formulated to space the alternatives with a couple of thoughts in mind. First, they shouldn't be too close in terms of data rate; otherwise, the quality difference wouldn't be meaningful and stream switching would occur too frequently. Second, you only want as many streams as necessary to create meaningful coverage for the compressed streams. That is, you want to provide logical stops along the quality continuum to best serve those watching on all possible connections.

Beamr can disrupt both goals as the table above shows. For example, the two 960 resolution streams in the Walk video start 600 kbps apart, which represents a meaningful difference in quality and bandwidth.

ADAPTIVE GROUP WORKFLOWS

After Beamr, the two files are less than 100 kbps apart, which could promote unnecessary stream switching. The difference between the two 1280 streams shrinks from 1.5 mbps to around 300 kbps, and from 2 mbps for the 1920 streams to around 600 kbps.

This might not be a bad thing, since you could advise the customer to eliminate three of the streams, cutting encoding and storage costs. A simple algorithm using some or all of the following concepts should be able to address the problem.

1. Retain at least one iteration for each resolution, no matter how close the data rates. Reasoning: Streams are often matched to display window sizes and you can't eliminate a stream because then the playback computer will have to scale or squeeze to window size, which is inefficient.

2. Don't eliminate any streams with different frame rates (usually mobile only). The reasoning is that the benefit in the lowest quality streams won't be that great, and trying to create an algorithm you could easily apply would be very complicated.

3. For streams with identical resolutions, start with the highest data rate stream. Eliminate any lower quality stream within 25% of the data rate of that stream. Restart the analysis for each stream that is preserved, so eliminate any streams that are within 25% of its data rate (this won't happen that often because few groups start with more than two or three files for each resolution).

The reasoning here relates to research I performed regarding the adaptive encoding practices used or recommended by technology companies like Adobe, Apple, Kaltura, Ooyala, Turner Broadcasting and others. In all cases but one, the lowest data rate differential between streams with identical resolutions files was 26.19%. This seems to represent a shared view that a data rate differential of less than this amount does not represent a meaningful difference in quality. For this reason, these streams can be safely eliminated.

Summary – Overall

Overall, the Beamr optimization process proved impressive in both test clips. I think most producers would be pleased with your results.

As mentioned, workflow integration, particularly within an adaptive streaming group, is a function that needs to be better addressed, and if possible, automated, a task that should be relatively easy to achieve as I outlined in the previous section.

Otherwise, Beamr seems best suited for larger media and broadcast companies who create their own custom encoding workflow and also distribute the volumes of videos where the bandwidth savings can be fully realized. The next obvious targets would be high volume encoding companies like OVPs and cloud encoders, who can also easily integrate Beamr into their workflows.