COMPARISON OF OPEN AND FREE VIDEO COMPRESSION SYSTEMS
A Performance Evaluation

Till Halbach
Norwegian Computing Center, Norway

Keywords: Still-image and video compression, Performance comparison, Dirac, Theora, H.264, Motion JPEG2000.

Abstract: This article gives a technical overview of two open and free video compression systems, Dirac and Theora I, and evaluates the rate distortion performance and visual quality of these systems regarding lossy and lossless compression, as well as intra-frame and inter-frame coding. The evaluation shows that there is a substantial performance gap of Theora and Dirac when compared to H.264- and Motion JPEG2000-compliant reference systems. However, an algorithm subset of Dirac, Dirac Pro, achieves a performance comparable to that of Motion JPEG2000, and which can be less than one dB below the PSNR performance of H.264 with TV-size and HD video material. It is further shown that the reference implementations of the codecs of concern still have potential for efficiency improvements.

1 INTRODUCTION

The video compression community has recently been witness to an interesting development concerning royalty-free specifications. Both the British Broadcasting Corporation (BBC) and the Xiph.org Foundation have released video codecs which are open-technology and open-source and free to use for the public. They compete therefore with patented codecs such as MPEG-*x* and H.26*x* which are, though specified in international standards, typically subject to a license fee, and not entirely open. While the lack of royalty claims makes these specifications interesting candidates for use by industry and individuals likewise, the technologies provided should offer a performance close to or better than today's state-of-the-art standards to ensure widespread acceptance.

So far, a detailed performance assessment of both Dirac and Theora relative to existing standards has been lacking in the literature; a gap this article wants to fill. All involved technologies are presented below with technical details. Dirac and Theora are then compared to each other and to reference standards to evaluate their performance. Conclusions are drawn at the end of the article.

2 DIRAC

The BBC have been working on a family of general-purpose video codecs with the umbrella name Dirac for some time now. As a result, two major products have been released to the public at the time of writing.

Dirac Pro, which is version 1.0 of the Dirac family, is an intra-frame video codec which targets lossless or near lossless image processing with low latency, suitable e.g. for studio and professional applications (BBC, 2008a). Dirac Pro has been submitted to the Society of Motion Picture and Television Engineers and is expected to become international standard VC-2 (W3C, 2008), with Microsoft's WM-9 being VC-1 (SMPTE, 2006). Version 2.1 of Dirac is a super-set of Dirac Pro, including motion estimation/ motion compensation (ME/MC) (BBC, 2008b). It is developed mainly with broadcasting and streaming in mind, targeting application areas like TV, digital cinema, and Internet. According to the Dirac Pro specification, there is the intention to extend VC-2 to include the full Dirac specification.

In particular for the end user it is of interest that the BBC claim the standards suite to be patent-/royalty-free and open-technology (Borer, 2005). The reference codec developed is licensed under the Mozilla Public License version 1.1.

Halbach T. (2009).
COMPARISON OF OPEN AND FREE VIDEO COMPRESSION SYSTEMS - A Performance Evaluation.
In *Proceedings of the First International Conference on Computer Imaging Theory and Applications*, pages 74-80
DOI: 10.5220/0001809700740080

2.1 Technology Overview

Like international video compression standards, Dirac defines the byte stream and decoding procedure only. However, for illustration purposes, the discussions below are made with regard to standard-compliant encoders.

Dirac Pro (BBC, 2008a) accepts RGB and YUV video, with chrominance subsampling factors of 4:4:4, 4:2:2, and 4:2:0. There are no restrictions on the bit resolution, and neither on frame size or frame rate. Predefined formats defined in the reference implementation include picture resolutions from below QCIF size to 4K Digital Cinema. The nature of the original video can be both interlaced and progressive. Dirac Pro does not involve any MC. A compliant encoder comprises three main compression components. A frame of a video sequence is first transformed by a wavelet filter bank to exploit spatial redundancies, before the transform coefficients are quantized and entropy encoded to exploit statistical correlation. The filter bank can be implemented in a very efficient way by means of lifting. The quantizers used are basically scalar dead-zone quantizers and are left out for lossless encoding. Entropy encoding can either mean variable-length coding (VLC) or arithmetic coding (AC). Concerning data migration, later transcoding without loss is possible by compressing video sequences in Dirac Pro's lossless mode.

Another feature of Dirac Pro is the use of independent chunks of data inside the byte stream, which allows for fast forward and backward functionality. All data are stored in a single byte stream, such that the transport in container formats like MPEG transport stream or OGG is possible. Due to the use of a filter bank hierarchy in the decomposition stage, incremental frame decoding is easily achieved, allowing for scalability in quality and resolution. Temporal scalability is supported implicitly as no temporal prediction is involved.

Though Dirac Pro defines the from other standards well known concept of profiles (sets of algorithms) and levels (constraints on decoder resources like frame and data buffer sizes), none have been defined yet. However, a coding option for high quality and another one for low latency exists. High quality is achieved by processing the entire frame at a time instance and implies thus a delay of at least one frame. Low latency is achieved by processing subregions of the frame at a time, and by replacing AC by VLC. This comes typically at the expense of less efficient compression. By the use of a wavelet filter bank and predicting subband coefficients, Dirac Pro can be related to the well know SPIHT video compression algorithm (Said and Pearlman, 1996) and the international standard Motion JPEG2000 (ISO/IEC, 2001).

As mentioned before, Dirac 2.1.0 (BBC, 2008b) deploys MC to exploit temporal redundancies and can therefore be counted as a member of the family of block-based hybrid video codecs. Dirac is identical to Dirac Pro, except that here not the original image is passed to the filter bank, but a motion compensated one. As a consequence, the functionality of scalability is lost. The motion prediction process is quite flexible, allowing any combination of various unidirectional and bidirectional modes, including weighting. An additional option for global motion estimation exists. The block sizes utilized in the prediction process can be of varying size and overlap each other, also known as overlapped block motion compensation. The accuracy can be either half-pel, quarter-pel, or 1/8-pel.

3 THEORA

Theora I is a video codec which is developed by members of the Xiph.org Foundation. It is more and more frequently found on the Web and used by large sites complying with the Wikimedia Commons like Wikipedia. However, the main application areas remain undefined. Xiph.org claims that Theora contains solely royalty-free and open technology. Its reference implementation is provided under a BSD-style license.

3.1 Technology Overview

Also this codec can be classified as belonging to the family of block-based hybrid video codecs. As with international standards, the specification details the decoder operation and data field order in the byte stream. As a concatenation of data packets, this stream can easily be encapsulated in any suitable transport container format. The standard allows progressive video material with 8 bpp accuracy and arbitrary dimensions, ranging from below QCIF to significantly more than 4K. The sequence may have a YUV color space with 4:2:0, 4:2:2, and 4:4:4 chroma subsampling.

A compliant decoder comprises five main components. The byte stream data are first entropy decoded employing Huffman codes. The decoded transform coefficients are then passed to the inverse quantization process and, subsequently, to an inverse 8×8 DCT, producing a frame difference signal, or frame delta.

The frame delta pixels are then added to the predicted frame to form the reconstructed frame which is processed by a deblocking filter before picture display and storage in the frame buffer. The filter has hence an in-loop position, and it is applied to block edges.

The block-based MC process utilizes forward prediction (P-frames) based on a single reference frame, as well as simple block copying, but no bi-predictive (B-) frames. The accuracy is either full pel or half pel. Otherwise it is of interest that that most algorithms make use of new data scanning orders in order not to infringe any patents. Arbitrarily accessing I-frames in the code stream is possible under the constraint that the byte stream header must have been transmitted and processed for decoder initialization previously. Other than that, it is worth mentioning that Theora does not provide the possibility for scalable picture decoding, and it lacks further the option of later lossless data migration from data stored in Theora format to another format, as it does not provide any means for lossless coding.

4 PERFORMANCE COMPARISON

In this section, Dirac and Theora are compared to each other performance-wise, and to Motion JPEG-2000 and H.264 as reference codecs and international state-of-the-art video compression standards (ISO/IEC, 2001; ITU-T, 2003). To the author's knowledge, only a single independent evaluation of Dirac has been undertaken previously (Onthriar et al., 2006), where the authors found that H.264 outperforms the at that time emerging Dirac both in *PSNR* and *SSIM* value, especially at low bot rates. Theora remains untested in that respect.

All encoders are operated in rate distortion optimization (RDO) mode or near RDO mode (if provided) and identical or near identical parameter sets required for fair codec comparisons. The involved motion compensation implies 250 P- or B-frames between I-frames and a single pass.

To assess the quality of Dirac 2.1.0, the version 0.9.1 of the reference implementation is used, which complies with the standard. The encoder is quality controlled by setting a quality factor similar to that of the JPEG standard. Near RDO is accomplished by requiring a full ME search with an area of 32×32 pixels. Concerning Theora I, version 0.19 of the reference implementation with the name libtheora is used. The software is quality controlled like Dirac with a quality parameter and is operated in RDO mode. For compressing video material compliant with H.264, the implementation named x264 is used. The encoder is rate controlled but operated without any particular RDO due to a lack of such an option. Two B-frames are inserted between two P-frames, for which the ME process is based on a single reference frame. The search area is of the size 32 square pixels and has a hexagonal form.

All video sequences used in the experiment comprise 8 bpp image material and a YUV color space with 4:2:0 subsampling. AKIYO is a 300-frame head-and-shoulder sequence in QCIF resolution with little motion, MOBILE has SD/PAL resolution and consists of 220 frames with a moderate amount of motion, while CREW comprises HD 720p material with flashing effects.

4.1 Inter-frame Comparison

The comparison of inter-frame coding schemes treats lossy and lossless coding separately.

4.1.1 Lossy Coding

The original video sequence is encoded and decoded involving temporal prediction, and the average luminance *PSNR* (Y-*PSNR*, in dB) is recorded, as well as the achieved bit rate in Kbps. The rate distortion (RD) curves are drawn over six different codec operation points, while only five points are recorded with Theora due to its limited allowed range for the quality factor. The bit rate is measured for raw video data except Theora, with which the data is encapsulated inside the OGG container format, implying the use of (a negligible amount of) overhead data. The size of the video data is naturally somewhat lower.

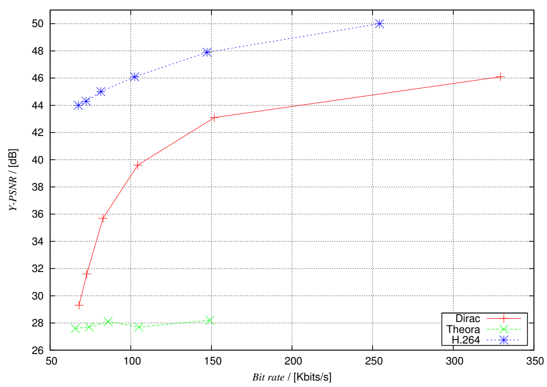

Figure 1: RD comparison for AKIYO (QCIF).

It can easily be seen in Fig. 1 through Fig. 3 that there are significant performance gaps among all involved codecs. The reference system H.264 outperforms both Dirac and Theora by a substantial amount, in terms of *PSNR*, at all rates. The difference between Dirac and H.264 can be more than 10 dB at low rates and up to 5 dB at moderate to high rates. Dirac's target application range is obviously high bit rate. Besides the aforementioned, Fig. 3 also bears proof of the fact that Dirac provides only a few coarse quantizers, which limits its RD curve at low RD products. Dirac also copes better with SD- and HD-sized video material than with smaller formats. This can be derived from the fact that the *PSNR* difference between Dirac and H.264 is significantly less (roughly 3 dB) in Fig. 2 and Fig. 3 than in Fig. 1.

Theora performs worse than Dirac. In Fig. 1 and Fig. 3, the reference implementation does not gain an improved performance with a bit rate increase, which is in contrast to what would have been expected intuitively. This is, however, likely an implementation issue, as the codec performs as expected with MOBILE. In either case, it can be concluded that neither Theora nor Dirac are suitable codecs for video sequences with less-than-SD resolution, considering *PSNR* values below 30 dB at low bit rates. There are manifold reasons for H.264's superiority, and its hierarchical ME mechanism with variable block sizes and a very accurate interpolation, the adaptive in-loop deblocking filter, as well as the highly efficient arithmetic encoder are among them.

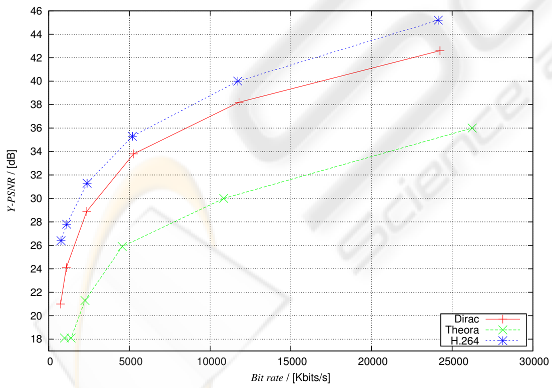

Figure 2: RD comparison for MOBILE (SD).

The gap in *PSNR* between Theora and H.264 can be as large as 10 dB, when the malfunctioning of the implementation is ignored. Fig. 2 shows that the gap decreases with roughly 2 dB towards low bit rates. Theora is obviously designed for low to moderate bit rates. The container overhead is measured to be around 1% and is hence without influence on the final result.

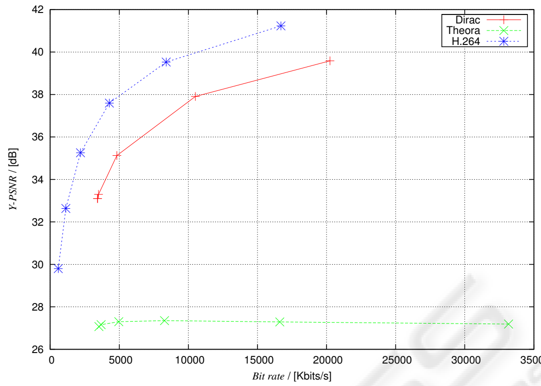

Figure 3: RD comparison for CREW (HD).

The efficiency of the involved implementations is measured in Tab. 1. The respective software is invoked 20 times on a single general-purpose CPU with 2.2 GHz clock frequency, and the CPU time needed for encoding (in seconds) is averaged over all runs. This comparison has to be interpreted with care, as implementations like the one for Dirac are for a proof of concept only and not optimized for speed. However, the comparison serves as a pointer for a rough assessment of the encoders' complexity.

Table 1: Implementation efficiency, in fps.

	Dirac	Theora	H.264
Akiyo	21	**379**	323
Mobile	2	9	**13**
Crew	0.03	0.3	**8**

The Dirac implementation appears to be the most inefficient software of the three measured implementations. Dirac's frame processing frequency is lowest in all cases, while the Theora and H.264 implementations are alternately best. However, it is of advantage that the processing frequency of Dirac's software is almost constant at varying bit rates, or in other words, its standard deviation σ_{Dirac} is much smaller than 1. The standard deviation σ_{Theora} equals roughly 31, 1, and 0.07 with AKIYO, MOBILE, and CREW, respectively, and $\sigma_{\mathrm{H.264}}$ is roughly equal to 49, 5, and 2, respectively. Summarizing, the frame processing frequency of Theora's implementation varies most.

It is concluded that Theora and H.264, which both deploy a transform instead of a filter bank like Dirac, can be operated much faster than the latter mentioned codec, despite the lifting structure as described in the Dirac specification. Dirac is close to real-time performance with QCIF-size video, while

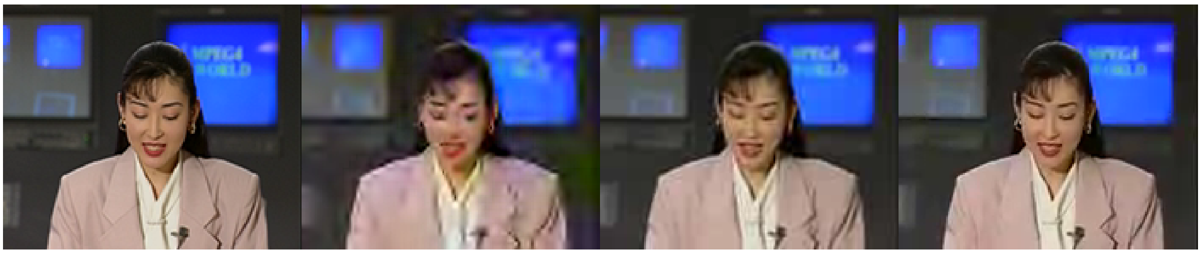

Figure 4: Original (left) and decoded frame 100 of AKIYO. Middle left: Dirac; Middle right: Theora; Right: H.264. The images may be slightly scaled due to space considerations.

Theora and H.264 are way above this requirement. Neither of the mentioned codecs is capable of encoding in real time with SD-size video or larger (i.e., without additional implementation optimization and without extra hardware). Concerning Dirac, it should be mentioned that a high-performance implementation named Schrödinger is being developed, and that specialized hardware exists for coding of HD signals in real time (Borer, 2007).

A visual comparison of all involved codecs at a very low bit rate reveals significant artifacts with both Dirac and Theora. Fig. 4 shows a single decoded frame of the AKIYO video. All videos were encoded as described previously. The bit rates of Theora and H.264 are 17.8 Kbps and 17.0 Kbps, respectively, while the for the reference implementation lowest possible bit rate with Dirac is 67.2 Kbps. It can easily be seen that, despite the significant higher bit rate, Dirac is the codec which performs worst, i.e. which introduces much more compression artifacts than the two competing codecs. The artifacts are mainly ringing along edges and blurred areas, which can be attributed to the use of a filter bank instead of a transform as with Theora and H.264. With Theora, the decoder outputs a slightly blurred picture when compared to H.264, and also minor blocking can be observed (in the women's face). Dirac and Theora yield 26.5 dB and 30.5 dB with the frame of concern, respectively, whereas H.264 achieves a *PSNR* of 34.3 dB. Here, the decoded picture is slightly blurred compared to the original, but no blocking artifacts are found, despite the approach of block-based MC and a block-based transform. The explanation lies in the good performance of H.264's in-loop deblocking filter, which obviously outperforms the filter employed in Theora. Note finally that, while the visual comparison (exemplified by a single frame) suggests that Theora is superior to Dirac at very low bit rates, the RD points in Fig. 1 are averaged over the entire sequence.

4.1.2 Lossless Coding

As already mentioned, Theora does not offer a lossless option. Concerning Dirac, it was found that the reference implementation was not stable enough to conduct these experiments, hence this topic remains for future research.

4.2 Intra-frame Comparison

This section investigates the efficiency of Dirac Pro and Theora when disregarding motion compensation, i.e. in pure intra-frame coding mode. With Motion JPEG2000 or short MJ2, an additional international state-of-the-art standard is taken as reference system. Lossy and lossless coding are again treated separately.

Version 1.3 of the implementation named OpenJPEG has been used to encode/decode to and from MJ2 format. Other implementations with differing efficiency exist (Pearson and Gill, 2005). The encoding process generates a single quality layer, and four and five decomposition levels for AKIYO and MOBILE/MOBILE, respectively. The lossy scenario involves an irreversible 9/7 wavelet filter bank, while the lossless scenario demands reversible 5/3 integer wavelets.

4.2.1 Lossy Coding

The experiments in this section are as described in Sec. 4.1.1, except that MC is turned off, i.e. the encoded bit stream consists solely of I-frames. Fig. 5 through Fig. 7 lack RD curves for Theora as its reference implementation is not capable of running in intra-frame mode.

Again, H.264 outperforms the other codecs significantly at high bit rates, while the difference in *PSNR* can be less than 1 dB at low rates, depending on the video material. It is further observed that Dirac Pro has roughly the same RD behavior as Motion JPEG-2000. In fact, Dirac's intra-frame coding appears to be optimized for compressing TV and HD material at low bit rates, given a *PSNR* close to that of H.264 with

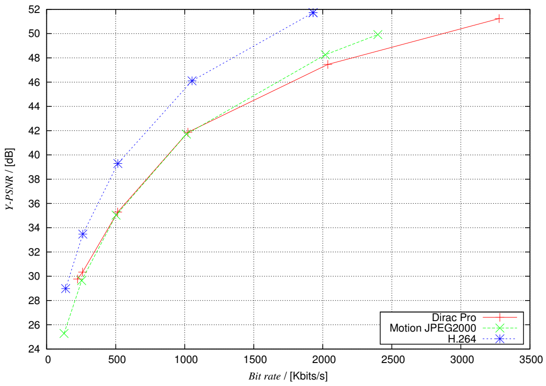

Figure 5: RD comparison in intra mode for AKIYO (QCIF).

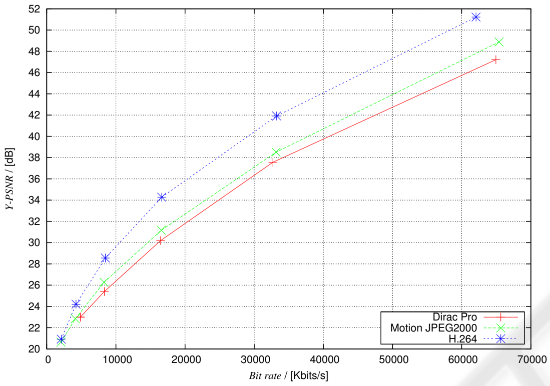

Figure 6: RD comparison in intra mode for MOBILE (SD).

MOBILE/CREW. This is in contrast to the inter-frame RD behavior as seen in Fig. 1 but can be explained by a different set of quantizers and differing statistical signal properties.

The latency of the reference implementations is measured as described in Sec. 4.1.1, and listed in Tab. 2. It is found that the H.264 software is best with a substantial margin in terms of frame processing rate,

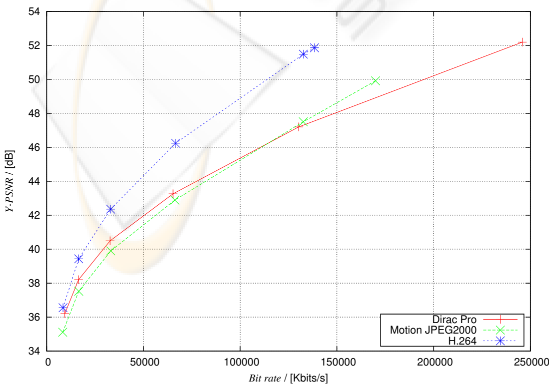

Figure 7: RD comparison in intra mode for CREW (HD).

Table 2: Implementation efficiency, in fps.

	Dirac Pro	Motion JPEG2000	H.264
Akiyo	128	62	**266**
Mobile	8	3	**15**
Crew	5	2	**11**

and the implementation of Motion JPEG2000 is the most inefficient of the three implementations of concern. However, the latter one performs in the most uniform manner with a σ_{MJ2} less than or equal to 2. The deviations $\sigma_{\mathrm{Dirac\ Pro}}$ and $\sigma_{\mathrm{H.264}}$ are 47 and 79 with QCIF-size material, 4 and 7 with PAL material, and 2 and 3 with HD material, respectively. I.e., the performance efficiency of H.264's implementation varies most.

4.2.2 Lossless Coding

Dirac Pro in lossless intra-coding mode is compared to both reference codecs. Tab. 3 shows the achieved bit rate of all coding schemes. Dirac Pro is outperformed by both reference codecs. H.264 appears to draw advantage of its highly efficient intra-frame prediction mechanisms which are superior to those employed in Dirac Pro. Motion JPEG2000, which in turn is based on JPEG2000, can rely on superior subband prediction and a highly efficient arithmetic coding engine. As already mentioned, the filter banks of Dirac Pro and MJ2 are set to be identical, so the arithmetic encoder of Dirac Pro has potential for improvement when compared to Motion JPEG2000. A side result of the experiments is that Motion JPEG2000 appears to be the best choice for professional studio applications.

Table 3: Comparison of lossless intra-frame coding schemes. The rate is given in Mbps.

	Dirac Pro	H.264	Motion JPEG2000
Akiyo	4.8	3.6	**3.1**
Mobile	102.3	86.8	**76.5**
Crew	335.1	276.2	**224.8**

5 CONCLUSIONS

Two open and free codecs were presented in this article. Their technical details were discussed, and a performance evaluation was given. It is stressed that the assessment is by no means exhaustive, but rather must be interpreted as a starting point for further research.

The experiments showed that the reference system H.264 outperforms the two codecs significantly, both in the objective and the subjective/visual evaluation. The low performance (in terms of *PSNR*) of Dirac and Theora surprises, since both codecs have been designed to compete with current cutting-edge technology. While Dirac is suitable for compression of SD/HD-sized video, its reference implementation is the slowest of the three systems of concern. The implementation of Theora comes closest to the implementation of H.264 in terms of efficiency. However, this system achieves the lowest *PSNR* of all involved codecs. Dirac Pro can compete with Motion JPEG-2000 over the entire rate spectrum and comes with less than 1 dB difference in *PSNR* close to H.264 with TV and HD material.

It is concluded that the algorithms of the two open and free codecs need improvement to be able to compete with state-of-the-art technology. With today's performance both standards can, however, be compared to former video compression systems like MPEG-2 and H.263+, as the performance gap between those standards and e.g. H.264 is well known (Wiegand and Sullivan, 2007). The battle for the most used video compression format on the Web will hence continue, and the candidates still have to position themselves in a better way, i.e. with improved RD behavior and other functionality. In particular, Theora should offer options for lossless coding and intra-frame coding, and Dirac should provide for compression at very low bit rates. At the same time, the encoding/decoding libraries have to be optimized for speed to increase the possibility for a wide acceptance of the offered technologies.

It is finally stressed that the comparisons presented in this article are not only comparisons of specifications, but also comparisons of encoder implementations. Though all standard-compliant, different encoders typically yield different performances. The results presented here can therefore not be generalized to be valid for other implementations or even other versions than those involved.

REFERENCES

BBC (2008a). Dirac Pro specification, version 1.0.0.

BBC (2008b). Dirac specification, version 2.1.0.

Borer, T. (2005). Open technology video compression for production and post production. *EBU Technical Review*.

Borer, T. (2007). Open technology video compression for production and post production. In *International Broadcasting Convention*, Amsterdam (Netherlands).

ISO/IEC (2001). JPEG2000 Image Coding System, Part 3.

ITU-T (2003). *Advanced video coding for generic audiovisual services*. Recommendation H.264.

Onthriar, K., Loo, K. K., and Xue, Z. (2006). Performance comparison of emerging dirac video codec with h.264/av. In *ICDT '06: Proceedings of the International Conference on Digital Telecommunications*, page 6, Washington (DC, USA). IEEE Computer Society.

Pearson, G. and Gill, M. (2005). An evaluation of motion JPEG 2000 for video archiving. In *Proc. IS&T Archiving*, pages 237–243, Washington (DC, USA).

Said, A. and Pearlman, W. A. (1996). A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. Circuits, Syst. for Video Technol.*, 6:243–250.

SMPTE (2006). VC-1 compressed video bitstream format and decoding process.

W3C (2008). Video on the web, workshop report. Accessed: 2008-02-28.

Wiegand, T. and Sullivan, G. J. (2007). The H.264/AVC video coding standard. *IEEE Signal Processing Magazine*, pages 148–153.