HDTV compression for storage and transmission over Internet

JAIME LLORET[1], MIGUEL GARCIA[2], ANTONIO J. FERRER[3] AND MARIA MORANT[4]
Communications Department
Polytechnic University of Valencia
Camino Vera s/n, 46022, Valencia
SPAIN
1

Abstract: - Because of the lacks in analogic television signal, researches proposed High Definition Television (HDTV), a new approach to transmit the television signal with higher definition. HDTV has quite better image quality than traditional TV. The recent transition to broadband access networks allows multimedia and TV services for Internet users, so it is needed an analysis and an evaluation of HDTV compression codecs for its storage and transmission over Internet. In this paper, we will show how to create a video with HDTV properties. Then, we will choose an adequate codec for HDTV compression to measure the compression gain, the compressed file size and the time needed to compress a second of video.

Key-Words: - HDTV, video compression, transmission over Internet, video measurements.

1 Introduction

The transition to High Definition Television has been different in Europe than in Japan and U.S.A. The first one who began to research on HDTV was the public Japan Television (NHK, Nippon Hoso Kyokai). They wanted to obtain better quality images and with vivid colours in a greater size format and higher resolution. This system was called MUSE (MUltiple sub-nyquist Sampling Encoding) and it is not compatible with their traditional TV (NTSC system [1]). In Europe, researchers began HD-MAC project [2]. European researchers developed a standard compatible with their traditional TV (PAL system). Finally, North-American researchers tried to improve and adapt Japanese proposal to their specifications.

There are four key concepts that indicate the main HDTV issues:

• Aspect ratio: It is defined as the relation between the height and the width of the screen. This parameter is used for other factors such as the transmission bandwidth. HD-MAC aspect ratio is 16:9 and MUSE aspect ratio is 5:3.

• Number of lines: It is the number of horizontal lines in each image. HD-MAC system has 1250 lines and MUSE system has 1125 lines.

• Definition: Its main goal is to enhance horizontally and vertically the image definition. Although, vertical resolution is related to the number of lines, it depends on other factors such as the persistence of the phosphorus in the screen, the frequency or the exploration system (continuous or interlaced). European system aspect ratio is 16:9, so the number of details in horizontal is: 1152 (effective lines) x 16:9 = 2048. Japan system has an aspect ration of 5:3, so the number of details in horizontal is: 1035 (effective lines) x 5:3 = 1725.

• Screen size: It is wanted a sweeping environment in HDTV, so it is necessary a vertical visual angle of 20 degrees. The screen size depends on the position of the people that will be watching TV (1,30x0,75 meters typically).

• Frequency: It is the number of images shown per second. HD-MAC uses 50 frames per second in continuous mode or 100 frames per second in interlaced mode. This values guarantee PAL and SECAM compatibility. MUSE system uses 60 frames in interlaced mode, but it has not compatibility with NTSC system because it uses 29,97 Hz.

Frequency parameter is very important because it is related with others parameters such as:

• Blinking,
• Movement resolution.
• Static resolution of the images in motion.
• Noise reduction.
• Bandwidth.

There are five main differences between HDTV and traditional TV:

1.- Bandwidth: Analogical traditional TV needs 6 MHz to be transmitted, but HDTV needs higher bandwidth values (over 30 MHz when it is not compressed).

2.- Number of lines: NTSC system uses 525 lines and PAL and SECAM systems use 625 lines. HDTV uses more than 1000 lines (HD-MAC 1250 lines and MUSE 1125 lines).

3.- Exploration system: Analogical traditional TV is based on an interlaced exploration, where the first image has odd lines (from line 1 to line 525) and the second image has even lines (from line 2 to line 524). HDTV uses progressive exploration format, where the image is explored from line 1 to the last one.

4.- Transmission format: Analogical traditional TV uses analogical signal for transmission, but HDTV uses digital signals.

5.- Aspect ratio: HDTV uses for each 16 units in wide 9 units in high, but analogical traditional TV uses 4:3.

The paper is organized as follows. Section 2 describes the problem of HDTV transmission over Internet. MPEG-4 standard is explained in section 3. We explain how we have created a sequence of video with HDTV properties in section 4. Section 5 shows how HDTV compression can be done. The results of our measurements are shown in section 6. Finally, section 7 gives the conclusions and future works.

2 Problem Formulation

A non-compressed HDTV video file has a very big size (10 seconds of video with one frame of video and one frame of audio needs 1,2 GBytes of storage memory). DVD does not have enough capacity to store HDTV videos, but FMD (Fluorescent Multilayer Disc) technology has higher capacity. Its read speed is 1 Gbyte per second and it is able to store 140 GBytes in a single optical disc [3].

The main problem in video transmission over IP is the bandwidth needed to transmit it and the quality of service needed between end systems. It is necessary the use of video high-compression algorithms for low bandwidth consumption in its transmission and to use few Bytes in storage memory [4].

There are many codification techniques. Some of them are based on image spatial compression (such as Motion-JPEG), others, such as H.261 and H.263, are based on temporal video sequence compression (they analyze the motion variation between an image and the next one). Their goal is to achieve good quality image meanwhile it has high compression.

A compressed video allows reducing costs in its transmission and promotes its distribution. One of the best compression systems that offers good quality with high compression is MPEG-4 standard.

3 MPEG4 standard

MPEG-4 [5] is a video compression ISO/IEC Standard. The application fields where it can be applied are Digital Television, interactive graphical applications and interactive multimedia. MPEG-4 provides high audiovisual data compression to store or stream video and, at the same time, it provides audio and video quality. MPEG-4 compression is based on visual-objects coding [6]. Visual-objects could be audio, images, text and graphics in 2 and 3 dimensions. It defines an audiovisual scene as a representation of codified video-objects (VOCs) and audio-objects (AOCs) that are treated independently and they have relationship over the time and the space, so the compression and the decompression are different. These objects are organized hierarchically, where the basic objects are static images, video objects and audio objects. Objects have some properties enclosed such as special coordinates, the scale, locality, zoom, rotation and so on. MPEG4 uses a system based on layers, where the first plane of the scene is separated from its environment. These layers, called planes, are the following ones:

• Video-objects plane (VOP): A VOP is every simple form of video-object in any time slot. Every VOP could be codified independently (I-VOP), or related with other VOP’s (P o BVOP) using the motion compensation. A VOP is split in macrobocks. In YUV 4:2:0 format, each macroblock has 4 luminance blocks and 2 crominance blocks. A macroblock has all information about the form, the texture and the motion of the coded object. It uses the same type of compression than MPEG-2. It is based on Discrete Cosine’s Transform with I-frames (keyframe), P-frames (predictive) and B-frames (bidirectional), but it introduces more efficient algorithms, offering better performance than MPEG-1 and 2 at low bitrates.

• Groups of VOP’s (GOV): It is a group of video-objects planes such as GOP’s in MPEG-1 and MPEG-2. They are optional.

These characteristics allow a user reconstitute the original video after decoding all layers of objects.

MPEG-4 can vary the bitrate from 5 kbps to 10 Mb/s. Video resolution can be from sub-QCIF (128 x 64) to HDTV (1440 x 1152). It allows pause, rewind and fast-forward with synchronized objects. Audio quality varies from the audio telephony (4 kHz) to stereo CD quality (20kHz). It allows interlaced video.

4 HDTV video creation

It is very difficult to find an uncompressed video with HDTV features, so we had to create our HDTV video but without using a High Definition video-camera. We chose HD-MAC, whose characteristics are: 1250 lines (1280x720 pixels), aspect ratio 16:9 and 50 frames per second. We have not found any video in Internet with these characteristics.

In order to create a HDTV video, we have to take many pictures and, later, we have to join them in order to make a motion sensation. So, we decided to take them from a dusk, when the sun was going down, fixing the photograph camera face to the sun (it is very difficult to get motion sensation without jumps while moving the camera).

Figure 1. Lanscape to take pictures

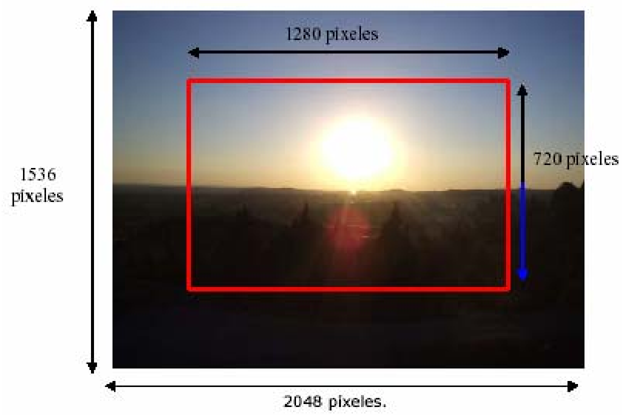

Figure 2. Pamscan getting the image from the picture.

Because our photographic camera was not able to take 50 pictures per second, we took pictures from a place where things were running slow. Therefore, we obtained the motion sensation without jumps. We had also avoided animals, cars, too much wind (because of clouds fast movement), and so on. Then, we added these pictures to a video editor in a ratio of 50 images per second. We wanted a minute of video, so we needed 3000 pictures (60 seconds x 50 pictures per second = 3000 pictures) for our purposes. We used a 3.1 Megapixel photographic camera that was able to make a picture every 3 seconds.

Once we were in the landscape, we began to take a picture every 3 seconds, so in 15 minutes we had 900 seconds and we took 300 pictures. Now, if we use 5 pictures for every second of video, to create 60 seconds of HDTV video we need 300 pictures. It will give us the sensation of a normal sun movement of 50 frames per second without jumps. The landscape used to take pictures could be seen in figure 1.

When we had all pictures, we had to touch up some of them because some times appeared a bird, a plane, cars, etc. in the scene. If we do not touch up those photos, we would have a sensation of marks in the video so it decreases the video quality.

The desktop program used to create the video was Adobe Premiere © [7]. We configured the project output properties with HDTV characteristics (50 frames per second, 1280x720 pixels and 16:9 aspect ratio) and uncompressed avi-format video using millions of colors in depth. We put 5 pictures per second (we changed the picture every 10 pictures). Last problem was the size of the pictures. They were 2048x1536, so they have higher resolution than the one permitted by HDTV. This time we used the Panscan technique to avoid the loose of quality (it gets the resolution needed from the centre of the image). Figure 2 shows how Panscan gets the image from the picture.

Finally, we obtained an uncompressed HDTV video. It lasted 60 seconds and its size was 3,86 GB. Now, we can see that a video with HDTV quality is too big to be stored or to be transmitted over Internet. Let's see which types of codecs are able to compress a HDTV video.

5 HDTV video compression

We have found 3 codecs that are able to admit HDTV as a compression profile:

• XviD mpeg-4 codec 1.1.0 [8]. It was created from DivX 4 and it is based on the same libraries as OpenDivX. It is open code and there are many developers improving it selflessly. XviD profiles allow Home DivX DVD player compatibility. "SA" (Simple Advance) profile limits the resolution to 352x288 and the data rate to 384 Kbits/s. "ARTS" (Advanced Real Time Streaming) profiles have the same resolution than "SA" profile, but their data rate are 4 Mbits/s. Each profile has several levels where resolution and bitrate can be varied, but it can be configured freely without using any profile. Its configuration window can be seen in figure 3 a). The following are the parameters that can be configured:

- The number of I, P y B frames at once. We can also choose frame zones to apply different compression.
- Quarter pixel motion compensation. It helps to fine the movement more precise, to fine the details and improve the compression. It is useful for low-resolution videos, but it is not compatible with Home DVD-DivX players.
- Global motion compensation. It is a technique that reduces temporal redundancy in an image sequence.
- Luminance hide or adaptable quantification. It reduces image quality using low bit rates in bright and dark zones, because quality looses are less appreciable.
- Framework quantization. It allows higher compression but it can diminish the image quality and the compression speed.
- Kind of quantizer. We can choose H.263 (for low bitrates) and MPEG (for high bitrates).
- Interlaced. It runs each frame separately.

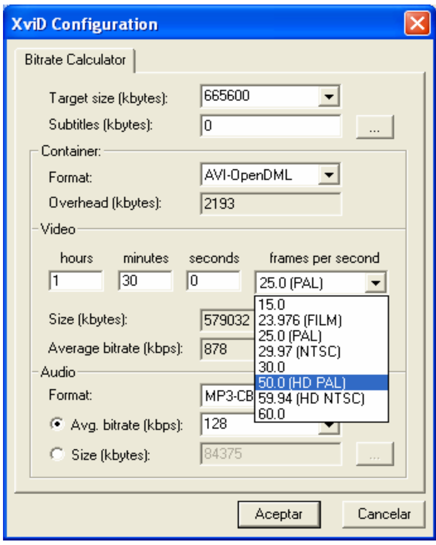

Figure 3. a) XviD 1.1.0 codec

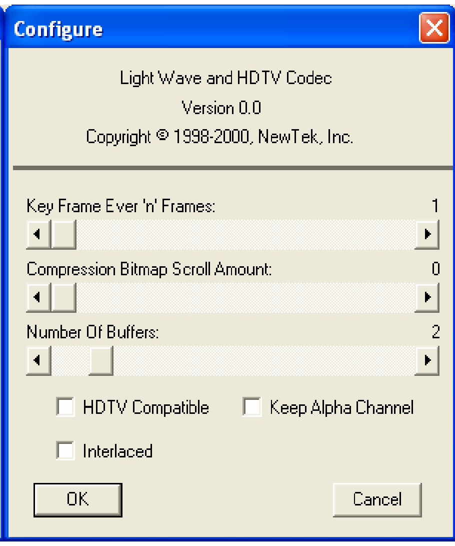

Figure 3. b) LightWave HDTV 0.0 codec

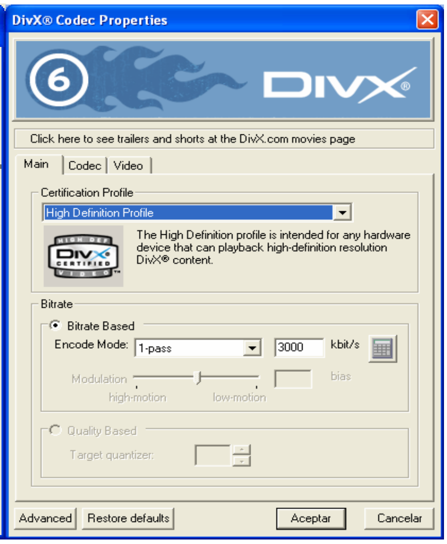

Figure 3. c) DivX 6.1 codec

• NewTek LightWave HDTV codec 0.0 [9]. NewTek Inc. created it. We can configure how many frames will be for every key frame. We can choose the HDTV compatibility and if we want it interlaced. Figure 3.b) shows its configuration window.

• Divx® 6.1 codec [10]. DivX appeared in 1999. Its main characteristics are the following ones:

- It has 3 codification speed modes (there is a relationship between the performance and the quality): Fastest, High performance, Balanced.
- Psycovisual enhancement based on DCT.
- Pre-processing: Spatial and temporal filters.

This codec allows choosing the number of I, P and B frames, the maximum interval between key frames, the global motion compensation, the type of quantization (H.263 and MPEG-2) and the interlaced. Figure 3.c) shows its configuration window.

The desktop application used to compress the HDTV video using the codecs aforementioned was “VirtualDubMod 1.5.4.1” [11]. After a comparison, we saw that XviD was the codec with best performance and with fewer problems during the compression. On one hand, sometimes NewTek LightWave HDTV codec hanged up the application and, on the other hand, when we used DivX codec, and we configured the application to maintain the same resolution in the output video, it displayed a window telling that the image format was not the correct one. We tested all these codecs in other desktop applications, but sometimes the frequency of the video obtained were always 30 frames per second or fewer and other times the number of pixels were not enough to accomplish high definition features.

While VirtualDubMod is compressing, we can see the codec status (as it is shown in figure 4) and how each frame is processed. We can also see the number of I type, P type and B type images and their size in the output video. I type images are not compressed, so its size is greater than P type and B type images. Last ones use motion compensation algorithms (unidirectional for P type images and bidirectional for B type images). We can also see that images graphically using VirtualDubMod Status. Figure 5 shows I type images (red lines) that are the largest ones (biggest size), P type images (medium size blue lines) and B type images (smallest size blue lines). These figures show us the GOP (Group of Pictures) codification structure. It is IPBBPBBPBBPBBPBB (there are few I type images).

We can configure the application to use 1-pass or 2-pass. The difference between 1-pass and 2-pass is that 1-pass compress the video directly, but the 2-pass first analyzes scenes complexity and, then, stores it in a log file that will be used in the second pass to have higher performance.

Once the 1-pass compression process is finished, we have obtained a minute of a compressed HDTV video. The size of the uncompressed video file was 3,86 GB and now it has 8,8 MB with HDTV characteristics, so it is 0,22% of its original size. The file obtained in 2-pass have higher size than the one obtained in 1-pass. The file in the second pass has a size of 57,4 MB, so it is 1,45% of its original size. It is 6,5 times more, but it is not significant compared to the size of the original HDTV video. The video quality of the 2-pass video file is higher than the 1-pass video file, but it is not appreciated at first glance.

6 Measurements

In order to analyze HDTV video compression for storage and transmission over Internet, we have measured the size of the files obtained for several compression bitrates, the time needed to compress the video and the compression gain for several bitrates.

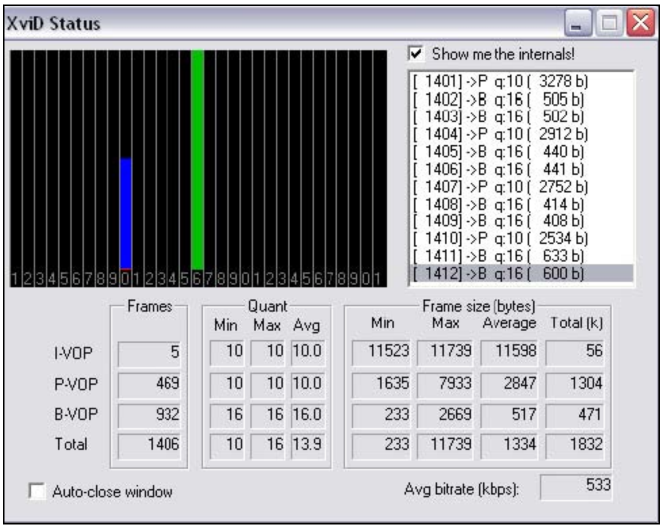

Figure 4. Xvid Status

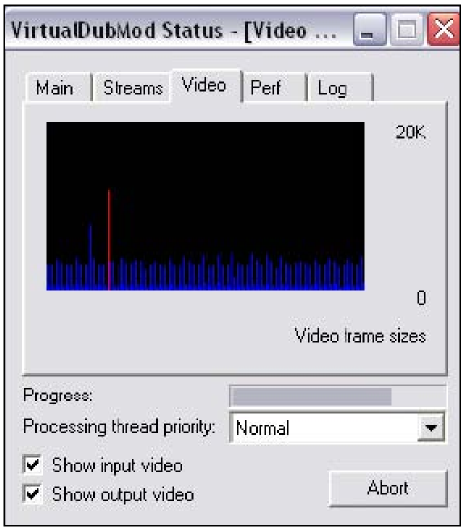

Figure 5. Video frame sizes

The bit rates used to take measurements have been 800, 1200, 1600, 2000 and 5000 Kbps.

We have compressed all files using 1-pass (because it could be used for transmission) and 2-pass (because it has better quality and it could be used for storage purposes).

Figure 6 shows the size of the compressed file as a function of several bitrates. The size of the file increases as the bitrate increases, but we have detected an anomaly at about 2000 Kbps in the 2-pass (see figure 7). We think the codec has better performance at that bitrate because its size has not increased linearly as for other bitrates (it should be 18 MB instead of 14,3 MB).

Figure 8 shows the time needed to compress the HDTV video for 1-pass. It is between 4 minutes and 42 seconds and 5 minutes and 1 second. The more is the bitrate, the time longer is. But it its less than when we use 2-pass. Figure 9 shows that the time needed to compress the HDTV video is between 4 minutes and 29 seconds and 4 minutes and 51 seconds. 2-pass range has lower values than the one for 1-pass.

Figure 10 shows the gain for 1-pass. There is a very high gain (692:1 when we compress using a bitrate of 800 Kbps using 1-pass). The main reason for that high gain is that the back scene of the video was always the same, there is only one object moving: the sun, so there was few I type images and a lot of images P and B. When we increased the bitrate, the graphic decreases exponentially. The higher is the bitrate, higher size has the file obtained. We obtain higher compression gain for 1-pass than for 2-pass. Figure 11 shows the compression gain for 2-pass.

7 Conclusion

The quality of the HDTV format is quite higher than standard television quality, so in a medium-term, TV will begin to transmit in this format. In few years people will consider this video format for storage or transmission.

We have explained how to create a video with HDTV characteristics without using a HDTV video camera nor any HDTV device.

We have described several available codecs to compress video with HDTV characteristics keeping its characteristics after the compression.

An uncompressed HDTV video file with a size of 3,86 GB has a size of 5,71 MB when it is compressed using 800 kbps bitrate for 1-pass and 52 MB when it is compressed using 5000 kbps bitrate for 2-pass.

We have detected an anomaly for the XviD codec when it is used for HDTV compression. At 2000 Kbps, the performance of the codec is higher because it compresses more the file than the value expected.

We have shown the time needed to compress a second of video for 1-pass and for 2-pass for several bitrates. 2-pass values are lower than 1-pass.

We have obtained high gains values (between 692:1 and 76:1) for several bitrates and for 1-pass and 2-pass. For a bitrate of 800 kbps and 1-pass we obtain a gain of 692, and for 5000 kbps and 2-pass we obtain a gain of 76.

Future works will develop a new HDTV codec accomplishing all HDTV characteristics, because we have found few codecs to achieve our goal. All measurements have been done using only a sequence of video without any audio multiplexed with the video. We will try to create a video sequence with audio multiplexed in HDTV format.

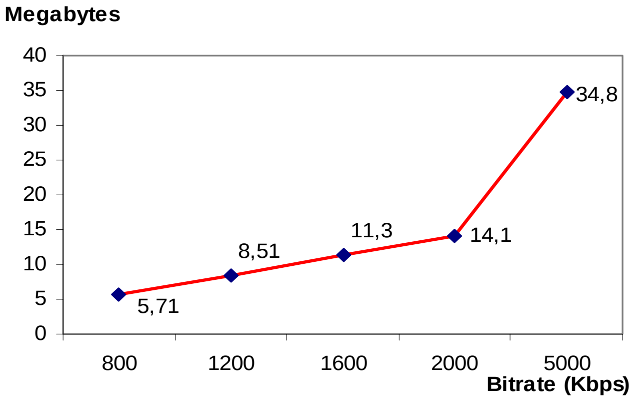

Figure 6. Size of the files for 1-pass

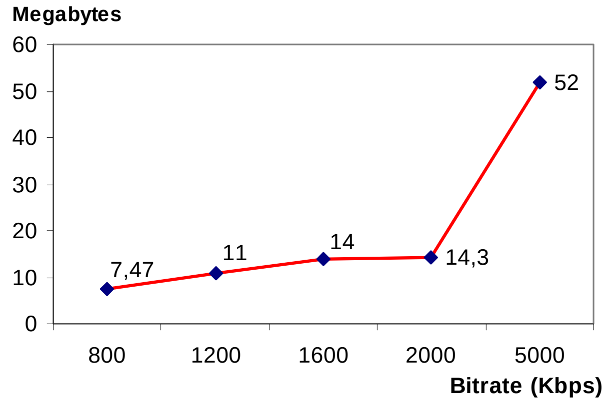

Figure 7. Size of the files for 2-pass

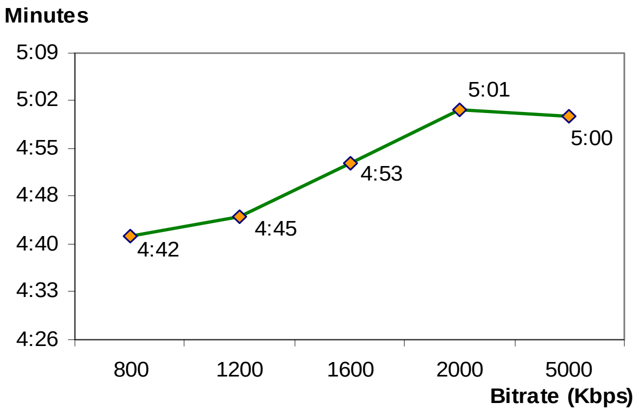

Figure 8. Time to compress a file for 1-pass

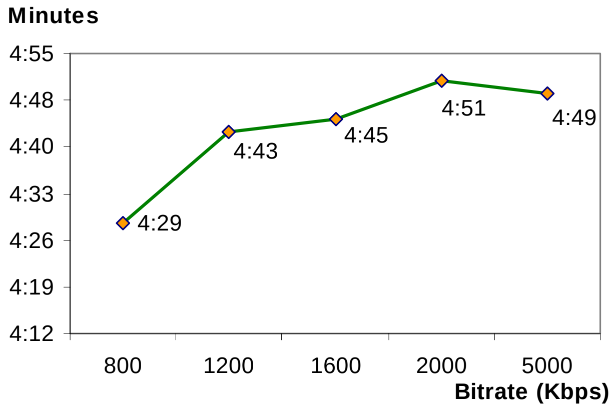

Figure 9. Time to compress a file for 2-pass

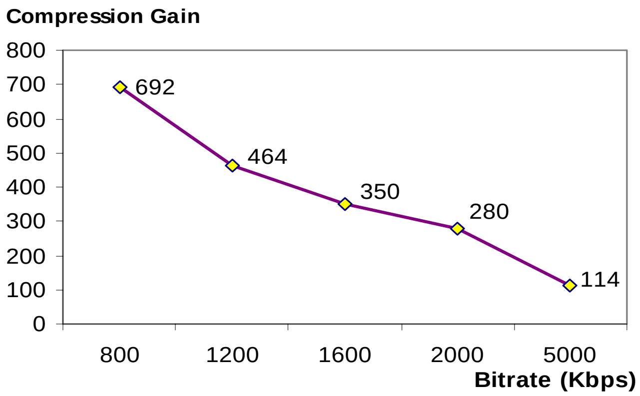

Figure 10. Compression Gain for 1-pass

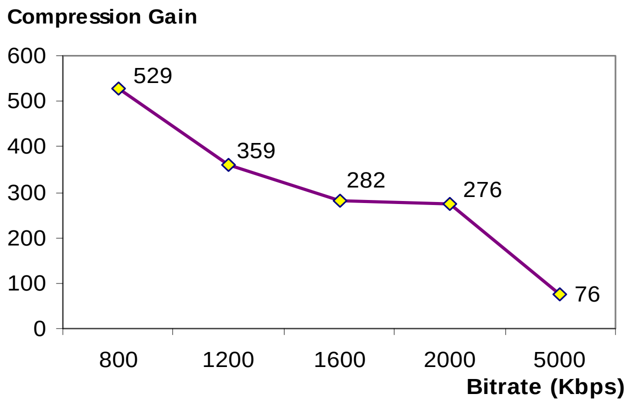

Figure 11. Compression Gain for 2-pass

References:

[1] Masafumi Saito. Digital Satellite Broadcasting and HDTV Services in Japan. ITU Seminar ISDB-S. Kiev, Ukraine, 13 - 15 November 2000.

[2] Council Directive 86/529/EEC of 3 November 1986.

[3] S Uhrig, F Betriebssysteme. Future Trends-Visionen des Mobile Computing im 21. Jahrhundert. At http://www.informatik.tu-darmstadt.de/BS/Lehre/Sem99_00/Texte/T12.pdf

[4] M. Bramberger, J. Brunner, B. Rinner y H. Schwabach, "Real-time video analysis on an embedded smart camera for traffic surveillance". Real-Time and Embedded Technology and Applications Symposium, 2004.

[5] R. Koenen, "Overview of the MPEG-4 standard", ISO/IEC JTC1/SC29/WG11 N1730, Luglio 1997.

[6] Coding of Audio-Visual Objects, Part-2 Visual, Amendment 4: Streaming Video Profile, ISO/IEC 14 496-2/FPDAM4, July 2000.

[7] Adobe Premiere website, http://www.adobe.com/products/premiere/

[8] XviD website, http://www.xvid.org/

[9] NewTek website, http://www.newtek.com/

[10] DivX website, http://www.divx.com/

[11] VirtualDubMod website, http://virtualdubmod.sourceforge.net/