RC 22096 (99055) 20 June, 2001 Computer Science

IBM Research Report

A Compression Scheme for Digital Cinema Application

Ligang Lu, Jack L. Kouloheris
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

COMPRESSION SCHEME FOR DIGITAL CINEMA APPLICATION

Ligang Lu and Jack L. Kouloheris

Multimedia Technologies Department
IBM T. J. Watson Research Center, P.O.Box 218
Yorktown Heights, NY 10298
lul@us.ibm.com

ABSTRACT

In this paper we present a compression scheme for the Digital Cinema application. Specifically, we developed a rate allocation algorithm for applying JPEG-2000 to coding Digital Cinema movies to achieve near lossless and smooth visual picture quality over the decoded movies at the required coding rate. First, a rate-distortion model is established based on the estimated source characteristics of the pictures, and a rate allocation algorithm is derived to determine a target rate for each picture based on the model so that 1) the overall distortion is minimized and 2) each picture has the same distortion at the required average bit rate. Then JPEG-2000 is employed to encode the high resolution motion picture according to the target rate allocation to achieve efficient compression performance and smooth picture quality over the picture sequence. The test results on Digital Cinema movie clips have shown that our scheme has achieved visually lossless coding performance and very smooth picture quality both visually and in PSNR.

Keywords: Rate Allocation, JPEG-2000, Digital Cinema, Motion Picture Coding, MPEG

1. INTRODUCTION

Compressed digital video is being used in more and more applications. Recently the multi-industry effort to develop a Digital Cinema Standard has received wide attention. This effort is to develop and standardize the techniques so that the films will be replaced by digital medium and movies will be produced, stored, and distributed in compressed digital format. Several major standardization groups and industry associations are involved in the standardization effort, such as SMPTE (Society of Motion Picture and Television Engineers), MPEG (Motion Picture Expert Group), MPA (Motion Picture Association), and NIST (National Institute of Standards and Technology). Currently the focus is on developing and determining the best enabling technologies.[1,2] The Digital Cinema application requires visually lossless video compression at a bit rate range from 35 to 90 Mbits/s for storage and playback.[3] The original movies for Digital Cinema application are 422 sources running at 24 frames/s with 1920x1080 resolution and a bit depth of 10-bits per component sample. This demands approximately a 30:1 compression ratio at the high end. The compression techniques currently under consideration are wavelet based coding, MPEG-2,[4] fractal coding, etc..[3] MPEG-2 is a motion picture coding standard widely used in digital TV broadcasting and DVD applications. However it was developed and optimized mainly for applications with bit rates under 35 Mbits/s. Also it does not support the 10-bit sample depth. The fractal encoding process is very computational intensive and thus very costly for real time implementation. Recently the joint photo expert group (JPEG) has finalized a new still image compression standard called JPEG-2000.[5] JPEG-2000 is largely based on a state-of-the art digital wavelet based coding technique called embedded block coding with optimal truncation (EBCOT). JPEG-2000 provides superior performance over existing still image compression standard on rate-distortion and subjective image quality. In addition, it includes many modern features, such as fine spatial and resolution scalabilities, strong error resilience, etc., which are very useful and important to many high-end and emerging applications. Therefore it is interesting and meaningful to investigate the potential of applying JPEG-2000 to the Digital Cinema application. However, JPEG-2000 is a still image compression technique and due to the content variations and camera operations, movies are statistically nonstationary video sources. Although, in compressing an image, JPEG-2000 has the ability to achieve a given target number of bits very accurately, the target number of bits must be first determined smartly. Compressing the movies at a fixed bit rate per picture will inevitably result in undesirable and often dramatic visual quality swing over the reconstructed picture sequence. Hence a rate allocation scheme is essential to use JPEG-2000 to compress movies. This rate allocation scheme should wisely determine the bits spent on each picture to achieve smooth visual quality over the picture sequence while maintaining the required total bit rate. In this paper, we present a scheme for using JPEG-2000 to compress the Digital Cinema movies. More specifically, we will illustrate our scheme in two steps. In the first step, a rate-distortion model will be established based on the estimated source characteristics of the picture, and a rate allocation algorithm is developed to determine the target bits for each picture based on the rate-distortion model and the required total bit rate. The objective is to achieve nearly lossless and constant picture quality over the movie sequence while meeting the bite rate requirement. In the second step, JPEG-2000 is used to encode each picture according to the target bit allocation from the first step to achieve efficient compression performance and very smooth picture quality over movie sequence.

In the next section, we first briefly describe the rate-distortion model and optimization operation in JPEG-2000. Then in Section 3, we will derive a rate allocation algorithm and a Digital Cinema compression scheme using JPEG-2000. In Section 4, we will show some test results and evaluate the performance of our scheme on coding the Digital Cinema movie clips. Finally we present conclusions in Section 5.

2. RATE-DISTORTION MODEL

In JPEG-2000, the probability density function (pdf) of the wavelet coefficients in each subband is assumed to be the generalized Gaussian density. The zero-mean pdf is given by

$$p(x) = \frac{\alpha}{2\sigma\Gamma(1/\alpha)}\sqrt{\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}}exp\{-\{\sqrt{\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}}(\frac{|x|}{\sigma})\}^{\alpha}\}, \tag{1}$$

where $\Gamma(x)$ is the *gamma function* defined as

$$\Gamma(x) = \int_0^{\infty} y^x e^{-y} dy \qquad x > 0. \tag{2}$$

The parameter α in Eqn. (1) determines the shape of the function, for instances, for the Gaussian density $\alpha = 2.0$ and for Laplacian density $\alpha = 1.0$. In JPEG-2000, the value of α for the i-th subband is estimated using the sample kurtosis,

$$K_i = \frac{\sum_j (x_{ij} - \bar{x}_i)^4}{\sigma_i^4}, \tag{3}$$

where x_{ij} is the j-th wavelet coefficient sample in the i-th subband, $\bar{x}_i$ is the sample mean and σ_i is the standard deviation of the i-th subband, respectively. α_i is obtained by solving

$$K_i = \frac{\Gamma(5/\alpha_i)\Gamma(1/\alpha_i)}{\Gamma(3/\alpha_i)}. \tag{4}$$

In JPEG-2000 α_i is rounded to one of the five allowable value in $(0.5, 0.75, 1.0, 1.5, 2.0)$ via a look-up table. To optimize the rate-distortion performance, the Lagrange technique and high rate assumption are used to derive the rate allocation and quantizer step for each subband. Let d_i be the distortion resulting from quantizing the i-th subband at the rate r_i, the JPEG-2000 rate allocation attempts to minimize the overall distortion in the image domain resulting from quantization in the wavelet domain subject to the constraint of the average rate R. The corresponding Lagrangain functional is

$$J = \lambda \sum_i \beta_i \gamma_i d_i + \sum_i \beta_i r_i, \tag{5}$$

where β_i is the fraction of all wavelet coefficients in the i-th subband and γ is the "energy weight" accounting for the range expansion/contraction due to inverse wavelet transform and any color transform performed, λ is Lagrangian Multiplier. The additive property of the distortion is provided by the assumption that the wavelet transform is orthogonal and the quantization errors for coefficient sample values are uncorrelated. In practice, the transform is usually only approximately orthogonal. Also the quantization errors using mean squared error metrics are not completely uncorrelated and only approximately additive. In most cases, these approximations are close enough.

The operational solution to Eqn. (5) can be obtained by finding the largest λ such that $\sum_i \beta_i r_i \leq R$ the desired rate. In JPEG-2000, once the subband pdf model parameter α_i is estimated, the rate r_i and the quantizer step Δ_i can be found in the look-up tables along with other parameters.

3. COMPRESSION SCHEME FOR DIGITAL CINEMA APPLCATION

In this section, we will derive a scheme for applying JPEG-2000 to compress the Digital Cinema movies. Our goals are 1) to minimize the overall compression distortion; 2) the decoded pictures have the same distortion or the movie has smooth quality.

3.1. Rate Allocation

From the last section, we know that once given a target rate allocation for coding a picture, JPEG-2000 will perform optimal rate allocation to all the subbands of wavelet coefficients to minimize the overall distortion in the image domain. So if the distortions are uncorrelated between the pictures, our first goal will be achieved in JPEG-2000's compression operation. To achieve the second goal, we need to devise a rate allocation scheme to determine the target rates for the pictures such that each decoded picture has the same distortion.

Suppose that there are N pictures in the movie and the pixel values of the k-th picture P_k is generated by the random variable $X_k, k = 1, 2, \ldots, N$. Let R_k be the target rate for coding the k-th picture P_k and D_k be the distortion resulting from coding P_k at the rate R_k. From the high rate theory, we have[6]

$$D_k = h_k \sigma_k^2 2^{-2R_k}, \tag{6}$$

where the distinctive distortion constant h_k is determined by the pdf $f_k(x)$ of the normalized random variable X_k/σ_k,

$$h_k = \frac{1}{12}\{\int_{-\infty}^{\infty} [f_k(x)]^{1/3} dx\}^3. \tag{7}$$

For example, if $f_k(x)$ is zero-mean Gaussian, $h_k = \sqrt{3}\pi/2$.

The optimal rate allocation is given by

$$R_k = \bar{R} + \frac{1}{2}\log_2 \frac{\sigma_k^2}{\rho^2} + \frac{1}{2}\log_2 \frac{h_k}{H} \tag{8}$$

where

$$\bar{R} = \frac{R}{N} \tag{9}$$

is the average rate per picture, and ρ^2 is the geometric mean of the variances $\sigma_k^2, k = 1, 2, \ldots, N$ of the pictures,

$$\rho^2 = (\prod_{k=1}^{N} \sigma_k^2)^{1/N}; \tag{10}$$

H is the geometric mean of $h_k, k = 1, 2, \ldots, N$.

$$H = (\prod_{k=1}^{N} h_k)^{1/N}. \tag{11}$$

The minimum overall distortion resulting from this rate allocation is

$$D = \sum_{k=1}^{N} D_k = NH\rho^2 2^{-2\bar{R}} \tag{12}$$

and each picture has the same average distortion

$$D_k = h_k \sigma^2 2^{-2R_k} = H\rho^2 2^{-2\bar{R}}. \tag{13}$$

The above rate allocation solution requires to know h_k for each picture, which may be estimated by fitting the picture pixel data to a pdf model. In the case that the pictures have the identical distribution after normalization, i.e., the normalized random variables $x_k = X_k/\sigma_k, k = 1, 2, \ldots, N$ have identical pdf's, then $h_k = h$ is a constant for all k. Eqn.(8) simplifies to

$$R_k = \bar{R} + \frac{1}{2}\log_2 \frac{\sigma_k^2}{\rho^2} \tag{14}$$

which is independent of the pdf of the random variables.

3.2. Compression Scheme

Based on the development in the previous sections, we propose a scheme to compress the Digital Cinema movies using the JPEG-2000 standard.

3.2.1. single pass scheme

Given the Digital Cinema movie picture sequence, $P_k, k = 1, 2, \ldots, N$, we first compute the sample variance $\sigma_k^2, k = 1, 2, \ldots, N$ and the geometric mean ρ^2 as in Eqn.(10). Then we determine the picture target rate allocation using Eqn. (14) and apply the JPEG-2000 to encode each picture at its target rate.

3.2.2. multi-pass scheme

Recall that we have invoked the high rate theory to derive the rate allocation in Eqn. (8) and the iid normalized pdf's assumption to derive the rate allocation in Eqn. (14). We have also retained the additive distortion property and orthogonal wavelet transform assumptions, etc. in JPEG-2000. In practice, these approximations may not all close enough. In the case that the resulting picture quality is not satisfactory or if the distortion varies significantly over the pictures, we can perform a multi-pass encoding scheme as in [7] on the whole movie or on the unsatisfactory portion to improve the picture quality smoothness using a revised rate allocation scheme. From Eqn. (6), we have

$$R_k = -\frac{1}{2}(\log_2 D_k - \log_2 h_k \sigma_k^2). \tag{15}$$

Taking the derivative with respect to D_k, we get

$$\Delta R_k = -\frac{1}{2\log 2}\frac{\Delta D_k}{D_k}. \tag{16}$$

Let $R_k^{'}$ be the rate to achieve the average distortion $\bar{D} = \sum_{i=1}^{N} D_k$ and substitute in Eqn. (16), we have

$$R_k^{'} = R_k + w\frac{D_k - \bar{D}}{\bar{D}}, \tag{17}$$

where w is a constant scaling factor. Sum Eqn. (17) over all the pictures,

$$\sum_{i=1}^{N} R_k^{'} = R, \tag{18}$$

which is the total bit rate.

4. RESULTS

Our scheme has been tested on Digital Cinema movie clips of *Star Wars Episode 1 – The Phantom Menace* from Lucasfilm Ltd. The movie clips are in 422 10-bits $YU'V'$ with 1920x1080 resolution. There are two formats and 8 clips in each format. The clips in Format 1 (FMT1) have only 1280x1024 valid pixels with the boundaries filled with black pixels to make up to the 1920x1080 resolution. While the clips in Format 2 (FMT2) have full 1920x1080 valid pixels. The clips have variety of contents and motion complexities. All the clips are coded at 50 Mbits/second using our scheme. The excellent PSNR results are shown in Table 4. More importantly, subjective quality evaluation performed by Lucasfilm Ltd. on the decoded clips indicates excellent quality, with visually transparent images even in split screen comparison with the source tape.[8] Fig. 1 and Fig. 2 show the frame PSNR the frame number of Clip 1 and Clip 2 for both formats as the examples of the picture quality smoothness.

To compare our scheme's performance with that of MPEG-2, we have coded 422 8-bit YUV 720x480 source *Cheer Leader* because MPEG-2 only supports 8-bit depth and MPEG-2 MP@ML is optimized in the bit rate range at and below 15 Mbits/s. We compared the performances of the two coding schemes at 10:1, 20:, 30:1 and 40:1 compression ratios. Figure 3 to Figure 6 show the plots of PSNRS vs. the frame number. The plots have shown that our scheme performs far better than MPEG-2 intra coding only mode and only 0.2dB to 0.7dB less than MPEG-2 intra and inter coding mode. More importantly the picture quality of our scheme is much smoother than that of MPEG-2. In conclusion, our scheme has achieved visually lossless and very smooth picture quality on Digital Cinema movie clips and has comparable performance to MPEG-2. However, our scheme does not need motion estimation/compensation so it can be used for low cost implementation. In addition our scheme does not use the motion compensated prediction technique thus can greatly facilitate the editing operations in movie production and post-processing.

Video Clip	FMT1 PSNR(dB)	FMT2 PSNR(dB)
Fogy Forest	53.01	49.26
Marching	47.90	45.06
Tanks	51.41	50.25
Robot Army	49.32	46.14
Monster	52.75	49.50
Fighter	51.49	48.79
Spaceship	55.47	54.98
Fighting	50.20	49.59

Table 1: Results on Digital Cinema Movie Clips of *Star Wars Episode 1-The Phantom Menace* at 50 MBits/sec

5. CONCLUSIONS

We have proposed a compression scheme for the Digital Cinema application. In our scheme We derived a rate allocation algorithm for applying JPEG-2000 still image coding standard to compress Digital Cinema movies. At the Digital Cinema required bit rate and source formats, our scheme has achieved 1) near lossless coding performance and 2) very smooth picture quality over the decoded movie clips. The testing results have also shown that our scheme is far better than the MPEG-2 intra coding mode and has very close coding performance to that of the MPEG-2 intra and inter coding mode. The subjective quality of the decoded clips are also excellent and impression. However, this scheme does not use the motion compensated prediction coding technique therefore it significantly facilitates the editing operation in movie production.

6. ACKNOWLEDGMENTS

The authors wish to thank Lucasfilm Ltd. for providing the movie source from *Star Wars Episode 1-The Phantom Menace* and for the permission to identify the source in the paper.

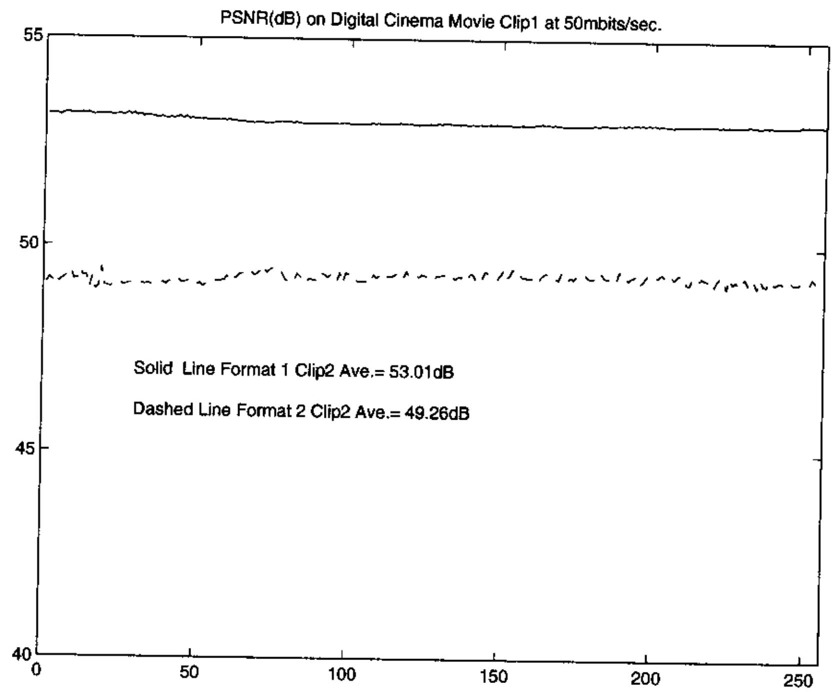

Figure 1: PSNR(dB) on Clip1 of *Star Wars Episode 1–The Phantom Menace* at 50Mbits/sec.

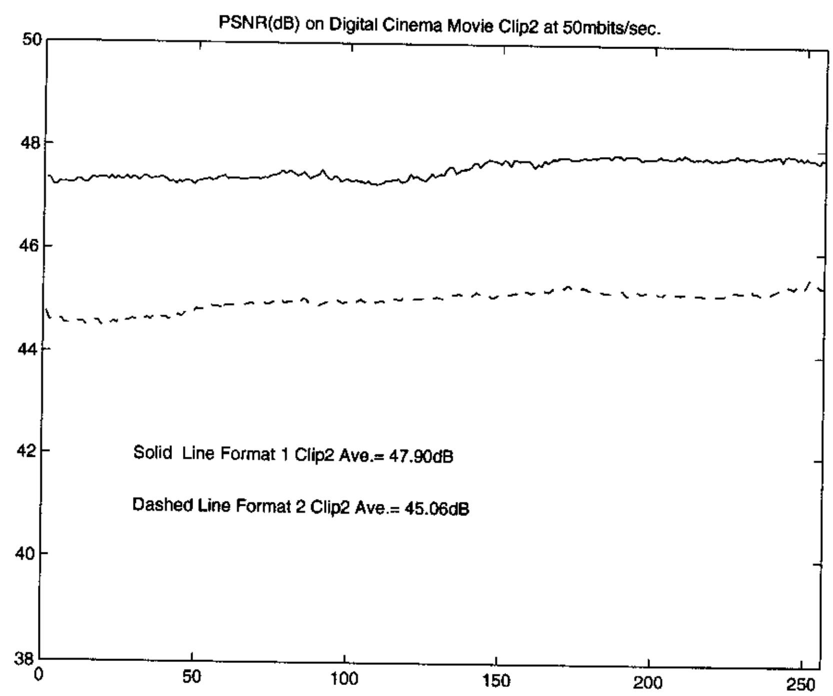

Figure 2: PSNR(dB) on Clip2 of *Star Wars Episode 1–The Phantom Menace* at 50Mbits/sec.

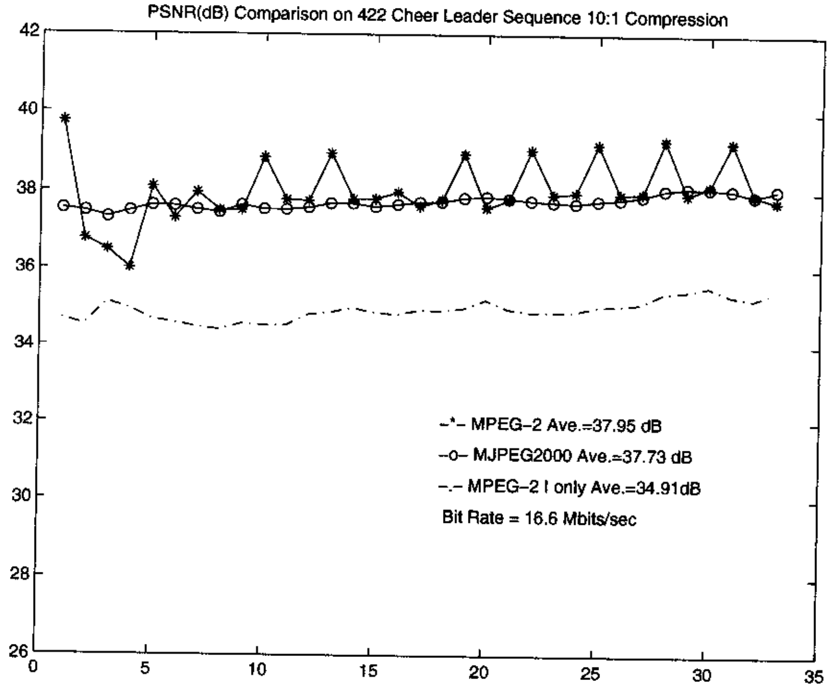

Figure 3: Comparison Results on 10:1 Compression

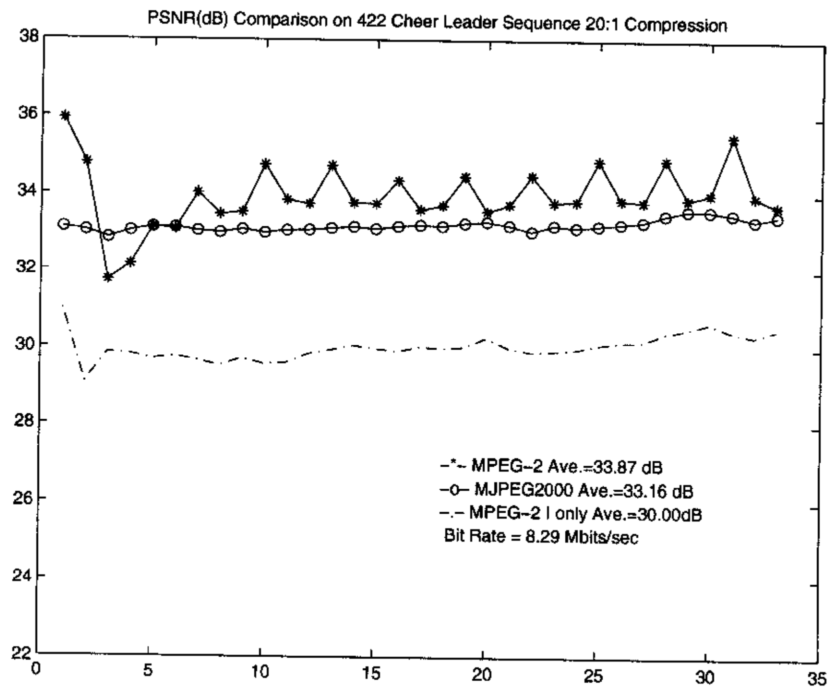

Figure 4: Comparison Results on 20:1 Compression

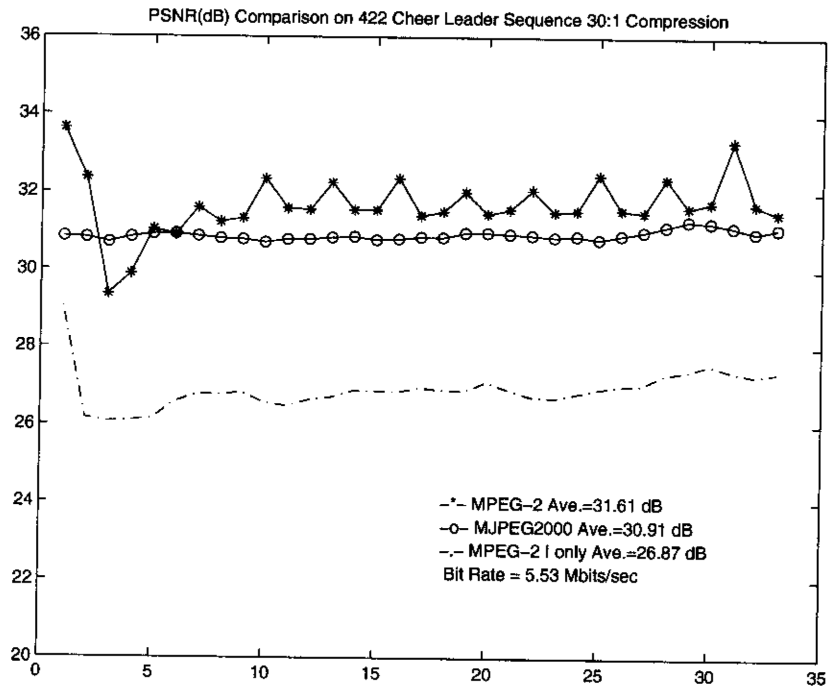

Figure 5: Comparison Results on 30:1 Compression

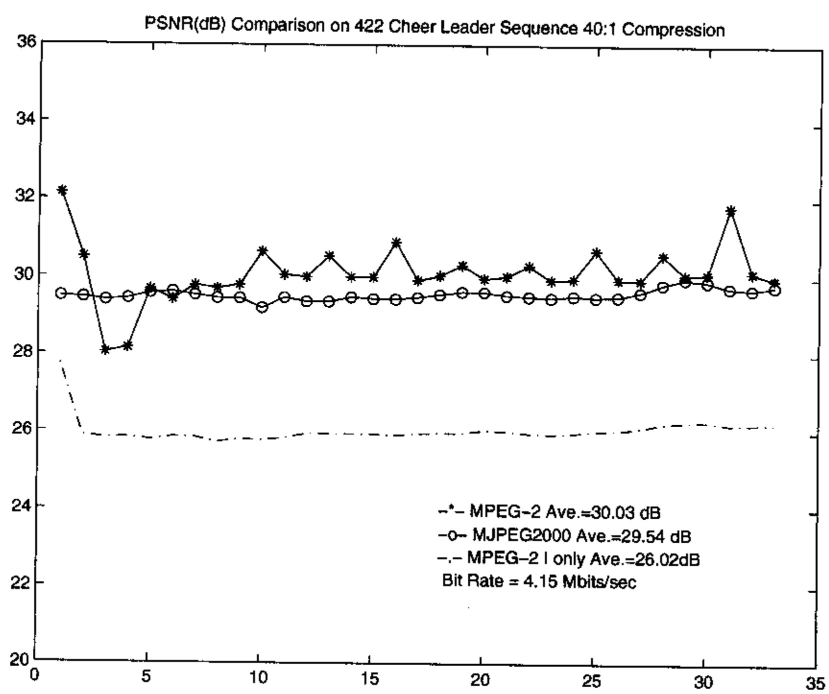

Figure 6: Comparison Results on 40:1 Compression

REFERENCES

1. ISO/IEC JTC 1/SC 29/WG 11 N3758, "Digital Cinema Requirements," Oct. 2000.
2. ISO/IEC JTC 1/SC 29/WG 11 N3663, "Ad Hoc Group on Digital Cinema," Oct. 2000.
3. P. Lelyveld, "Overview of Digital Cinema," *Proceedings of Digital Cinema Conference*, Jan. 2001, Gaithersburg, MD.
4. "Information Technology–Generic Coding of Moving Pictures and Associated Audio Information: Video," First Edition, *ISO/IEC 13818-2*, May 1996.
5. ISO/IEC JTC 1/SC 29/WG 1 N1878, "JPEG2000 Verification Model 8.5," Sept. 2000.
6. A. Gersho and R. Gray, *Vector Quantizations and Signal Compression*, Kluwer Academic, 1992
7. L. Lu, J. Kouloheris, C. Gonzales, et al, "Segment Reencoding of Buffer Constrained Variable Bit Rate Video Streams" *Proceeding of IEEE International Conference on Multimedia&Expo 2000*, New York, NY, July 2000.
8. Dave Schnuelle, Private Communications, Lucasfilm Ltd. Nov. 6, 2001.