# A Survey of Optimization of Rate-Distortion Techniques for H.264 Scalable Video Coding

**G.S. Sandeep, B.S. Sunil Kumar, G.S. Mamatha and D.J. Deepak**

**Abstract** Rate-Distortion optimization is a process of improving a video quality during video compression. This will mainly concentrate on amount called rate and is a measure of distortion against data required to encode the video. The main usage of video encoders is to improve quality in encoding image, video, audio, and others to decrease the file size and to improve the quality. The classical method of making encoding decisions for the video encoder is to choose the result that yields highest quality output. Scalable video coding (SVC) is a process in which encoding and decoding can be done on a bit stream. The resulting video has lower temporal or spatial resolution or a reduced fidelity, while retaining reconstruction quality that is close to the original. SVC reduces bit rate necessarily required to represent a given level of perceptual quality and also improves coding efficiency. The main advantage of this Rate-Distortion is applied to measure the quality obtained from the video signals. Some of the rate-distortion models are prediction, approximation, Lagrangian multiplier, empirical, and parametric. Among these Lagrangian multiplier provides good rate-distortion optimization. The MSE- or PSNR-based quality metrics are usually used to assess the visual quality in the cases that the spatial resolution and frame rate are fixed. Our main aim is to demonstrate gain in PSNR with optimized value of bit rates. Finally, we conclude by presenting the efficient techniques to achieve improved quality of video.

**Keywords** Bit rate · CGS · FGS · MGS · MSE · Peak Signal-to-Noise Ratio Rate Distortion · Scalability · Scalable Video Coding

---

G.S. Sandeep (✉) · B.S. Sunil Kumar · D.J. Deepak
G M Institute of Technology, Davangere, India
e-mail: sandeepgs@gmit.ac.in

B.S. Sunil Kumar
e-mail: sunilkumarbs@gmit.ac.in

D.J. Deepak
e-mail: deepakdj@gmit.ac.in

G.S. Mamatha
R.V. College of Engineering, Bengaluru, India
e-mail: mamatha.niranjan@gmail.com



M.S. Reddy et al. (eds.), *International Proceedings on Advances in Soft Computing, Intelligent Systems and Applications*, Advances in Intelligent Systems and Computing 628, https://doi.org/10.1007/978-981-10-5272-9_22

## 1 Introduction

Rate-Distortion is a quality metric for measuring both deviation of source and cost of each bit stream in a possible set of decision outcomes. These bits are measured mathematically by multiplying the cost of each bit with Lagrangian value and are a set of values representing relationship between bit cost and quality level. To minimize true Rate-Distortion cost in hybrid coding, a proper designing of framework is required that jointly designs motion compensation, quantization, and entropy in H.264. Some of the rate-distortion algorithms are available and are named as: firstly, graph-based algorithm for soft decision quantization requires motion compensation and step sizes. Secondly, iterative algorithm uses residual coding in H.264 with only motion compensation. Some of the innovative features of H.264/SVC combine temporal, spatial, and quality scalabilities into a single multilayer stream. The provision of scalability in terms of picture size and reconstruction quality in these standards comes with a considerable growth in decoder complexity and a significant reduction in coding efficiency. The spatial and temporal scalabilities have bit stream to represent content of the source with reduced picture size and frame rate, respectively. The usage of SVC in spatial scalability with arbitrary resolution ratios gives a value and is a ratio of picture size for complete bit stream to the included sub-stream.

In the year 2007, the scalable extension of the H.264 codec was invented named as SVC. In the year 2010, a joint collaboration by ITU-T VCEG and ISO MPEG derives a new model for SVC. To increase flexibility of video encoder at application layer for packet level, we need to use bit rate adaptation. Compression efficiency has been improved by applying inter-layer estimation for the video frames.

In recent years, the video traffic in Internet forms a significant increase of information. In 2011, this share of internet video was 51% of all consumer Internet traffic and is expected to rise to 55% in 2016. These high numbers do not include video exchanged through peer-to-peer sharing. Finally, the sharing of video on the overall Internet traffic is expected to reach 86% in 2016. The important aspect related to video quality is a measure of bit rate of the video carried over the network. When this measure is too low, means the visual quality of the decoded video may be degraded. On the other hand, when this is too high, it may lead to video freezes or longer waiting times or loss of data.

The aim of this paper is to survey the different papers to identify the different scalability measure and that uses prediction parameters and list out the peak signal-to-noise ratio values that are related with the rate-distortion parameters. In this work, we are focus on the measures of Rate-Distortion and statistical values are obtained from different authors and finally concluded that by using Lagrangian rate-distortion optimization, we get good Rate-Distortion results and is measured with PSNR values. From the analysis, we can identify that Rate-Distortion mainly depends on CGS and MGS scalability parameters, and it will produce an acceptable value of PSNR with lesser bit rates.

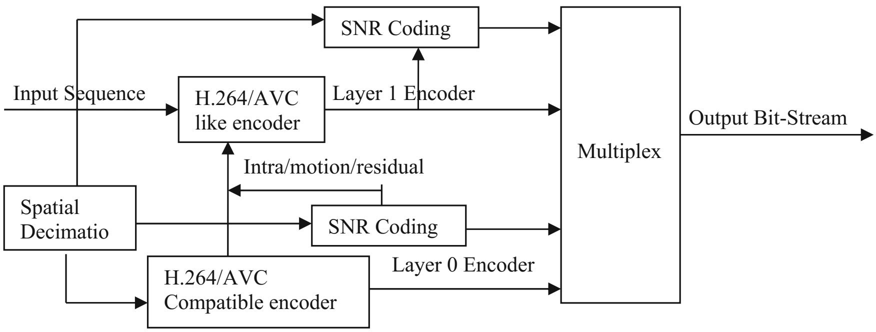


**Fig. 1** Simplified SVC encoding structure with two spatial layers

## 2 Overview of SVC

Figure 1 shows the architecture of the SVC encoder. The architecture proposes a more flexible design to improve the functional efficiency of encoder. Basically, the prediction of blocks using motion-compensated and intra prediction within the same layer are introduced. However, dependency between layers is carried out by inter prediction section, which again utilize the motion vector, similar structures in a frame, so as to improve the ratio of compression in the encoding process.

The base layer is designed with coarse grain scalability to code it as first. The enhancement layer uses fine grain scalability to improve the resolution which may increase the complexity of hardware design.

There are three types of scalability in SVC: (a) Temporal scalability, (b) Spatial scalability, and (c) SNR scalability.

### *2.1 Temporal Scalability*

By using frame rates and bit-stream subsets, we can be able to represent video. Encoded video streams have three distinct types of frames: I (intra), P (predictive) or B (bi-predictive). I frames only explore the spatial coding within the picture, whereas both P and B frames do have interrelation with different pictures. While in P frames inter-picture predictive coding is performed based on one preceding reference and B frames consist of a combination of inter-picture and bi-predictive coding.

### 2.2 Spatial Scalability

It represents layered structure videos with distinct resolutions; i.e., each enhancement layer is responsible for improving the resolution of lower layers.

The three prediction techniques supported by this module are

- *Inter-Layer Motion Prediction*: The motion vectors and their attached information must be rescaled.
- *Inter-Layer Intra Texture Prediction*: Supports texture prediction for internal blocks within the same reference layer (intra). The intra block predicted in the reference layer can be used for other blocks in superior layers.
- *Inter-Layer Residual Prediction*: In SVC, the inter-layer residual prediction method can be used after the motion compensation process to explore redundancies in the spatial residual domain.

### 2.3 SNR Scalability

We legitimize transporting the complementary data in different layers in order to produce videos with distinct quality levels. In SVC, SNR scalability is implemented in the frequency domain.

The H.264/SVC standard supports three distinct SNR scalability modes;

(a) *Coarse Grain Scalability*: Each layer has an independent prediction procedure, and consecutive layers have the same resolution.
(b) *Medium Grain Scalability*: It increases efficiency by using a more flexible prediction module, whereas both base and enhancement layers can be referenced.
(c) *Fine Grain Scalability*: It employs an advanced bit-plane technique where different layers are responsible for transporting distinct subset of bits corresponding to data information.

## 3 Rate Distortion Done by Different Authors

By calculating peak signal-to-noise ratio (PSNR) values of the frames of a video sequence as objective video quality measure. Here we consider a frame with pixels (8bit per pixel), and now Calculating PSNR from mean squared error (MSE).

$$\mathrm{MSE} = \frac{1}{Nx \cdot Ny} \sum_{x=0}^{Nx-1} \sum_{y=0}^{Ny-1} [F(x,y) - R(x,y)]^2 \tag{1}$$

$$\text{PSNR} = 10 \cdot \log_{10} \frac{255^2}{\text{MSE}} \tag{2}$$

In Sun et al. [1], authors have analyzed sub-bit-plane technique of scalable video coding for fine granular SNR and also checks that MSE-based distortion rate (D-R) function should be linear within a FGS layer and finally achieve good coding efficiency with reduced complexity by using model-based quality layer assignment algorithm.

In Mansour et al. [2], authors have analyzed rate-distortion prediction model for medium grain scalable, and it allows for video encoder to predict size and distortion of video frame. Here, cross-layer optimization capabilities are used to achieve best picture quality.

In Thomas et al. [3], authors describe secure scaling for bit streams at intermediate nodes, and it mainly supports medium grain scalability and coarse grain Scalability and finally incorporates FGS packets and produces more robust pictures.

In Sun et al. [4], authors analyzed the rate-distortion function of SVC FG coding and provide a solution using approximation model of the rate-distortion function. This solution consists of integer transform coefficients and extracts properties of Rate-Distortion using Gaussian model and apply approximation to obtain good quality picture.

In Li et al. [5], authors describe Lagrangian multiplier for rate-distortion optimization. It directly supports multilayer scenario, but is not efficient because correlation between the layers is not considered. To overcome this, finally a new selection algorithm for rate-distortion optimization was proposed.

In [6], authors have made an assumption for fixed display scenario at constant rate of the frame, and it consists of one CGS and another MGS for creating quality layers for bit stream. Author proposed hierarchical coding which yields dyadic decomposition of temporal layers.

In [7], authors proposed a scheme named rate-distortion optimization with ROI slices to improve coding efficiency. It discards background slices, and it uses base layer information to generate prediction signal of EL. And lastly, author derives Lagrange multiplier for performance improvement.

In [8], authors have proposed new rate control scheme for principle component analysis of scalable video coding. An improvement has made and create a new model named TMN8 for rate control on all the frames like P, B, I. This scheme is more accurate in case of bit rate and gains much peak signal-to-noise ratio.

In [9], authors also proposed updated rate control scheme for spatial scalability for scalable video coding. The adaptive approach is used to decide a quantization parameter and by applying transform coefficients on spatial scalability. A new model called Cauchy Density is used to derive natural logarithm domain Rate-Distortion model this matches spatial scalability characteristics this in turn will provide an excellent performance over the earlier rate control schemes.

In [10], authors have analyzed empirical model for Rate-Distortion, and it checks for channel rates and capacity. By utilizing optical rate prediction architecture with cross-layer design side information, a new empirical model has been created to get good video distortion method. Finally, author uses low-density parity-check approach to estimate the performance of video Rate-Distortion.

In [11], authors have analyzed that medium grain scalability is more superior than coarse grain scalability in case of rate-distortion complexity, and it also improves error resilience. By applying inter-layer prediction; it improves rate-distortion complexity of the scalable video coding. On usage of low-delay group of picture-structure, mainly in handheld devices, reduces decoding complexity also the structural delay.

In [12], authors have described adaptive pre-filtering system for SVC computational complexity-rate-distortion analysis; as a result the visual presentation quality at decoder side is improved with using limited resources. SVC preprocessing will provide efficient coding for desired region of interest.

In [13], authors have analyzed large-scale Rate-Distortion and rate variability distortion characteristics of CGS and MGS approach and found that CGS achieves low bit rate compared to single-layer encoding. From the analysis, the traffic variability of CGS is lower compared to single-layer streams. Similarly, MGS can achieve slightly higher Rate-Distortion efficiency than single-layer encoding. By using hierarchical B frame structure of MGS layer, a high Rate-Distortion performance is achieved.

In [14], authors have developed a model named as parametric Rate-Distortion for medium grain scalability and that uses spatial and temporal complexity of video sequences with variable bandwidth that will improves the quality of video sequences.

In [15], authors have analyzed optimal bit-stream extraction scheme for SVC, and it produces scalable layer representation for multicasting network with varying bandwidth. This scheme also extracts the optimal path for multicast nodes and also helps in reducing optimal path extraction for less capable devices through path truncation. Finally, it gives good PSNR and MSE.

In [16], authors have analyzed rate-distortion function of SVC FGS pictures by using Gaussian model, and it proves distortion-rate curve will be concave. By using sub-bit-plane technology, distortion rate is inferred to be linear under MSE. To find drift error, author uses effective distortion model in SVC. A new virtual group of picture concept and new priority setting algorithm is designed to achieve good optimal Rate-Distortion performance.

In [17], authors have proposed rate-distortion model in scalable wavelet video coding to find efficiency. Generally, Lagrangian multiplier provides optimized solution for mode decision and rate-constrained motion estimation. This operates on multiple bit rates and open loop structure. Finally, proposing motion prediction gain metric to measure efficiency and always improved.

In [18], authors have analyzed rate-distortion optimization scheme to utilize transform coefficients in spatial SVC for solving l1-regularized least squares problems and produce larger PSNR in reducing bit rate. Finally, this will improve the coding efficiency.

## 4 Results and Analysis

An exhaustive survey is made over 12 papers published in international conferences or in journals of Scopus index. The very purpose of this survey is to identify the best technique to produce highest bit rate of transmission; at the same time it should ensure an appreciable quality (PSNR) of video. Table 1 enlists the different techniques and their performance parameters. The observation indicates all the techniques proposed over the timeline are not efficient. This can be observed in Fig. 2. However, few papers showed an aggressive performance in bit rate results retaining a video quality near to the original. Such techniques are identified and are listed in Table 2. The graph of the results in these techniques clearly shows an improvisation

**Table 1** Comparison of PSNR versus bit rate for different rate-distortion techniques

| Ref. paper no. | Rate-distortion techniques | PSNR (dB) | Bit rate (kbps) |
|---|---|---|---|
| 10 | Empirical model | 29.87 | 379 |
| 11 | Inter-layer prediction | 30.45 | 320 |
| 15 | Bit-stream extraction scheme | 31.05 | 135 |
| 14 | Parametric Rate-Distortion | 31.60 | 246 |
| 18 | Inter-layer residual prediction | 31.80 | 190 |
| 09 | Rate control scheme | 32.38 | 132 |
| 13 | Large-scale Rate-Distortion and rate variability distortion | 35.00 | 200 |
| 05 | Lagrangian multiplier | 35.10 | 300 |
| 01 | Sub-bit-plane technique | 35.35 | 192 |
| 17 | Motion prediction gain metric | 36.50 | 253 |
| 16 | Bit-stream extraction scheme | 37.53 | 384 |
| 12 | Adaptive pre-filtering system | 40.40 | 183 |

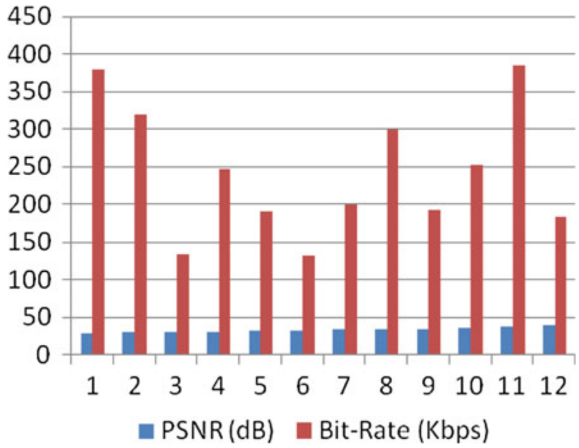


**Fig. 2** Comparison of PSNR versus bit rate for different rate-distortion techniques (color figure online)

**Table 2** Values of optimized PSNR versus bit rate in different techniques

| PSNR (dB) | Bit rate (kbps) |
|---|---|
| 29.87 | 379 |
| 30.45 | 320 |
| 31.60 | 246 |
| 35.00 | 200 |
| 35.35 | 192 |
| 40.40 | 183 |

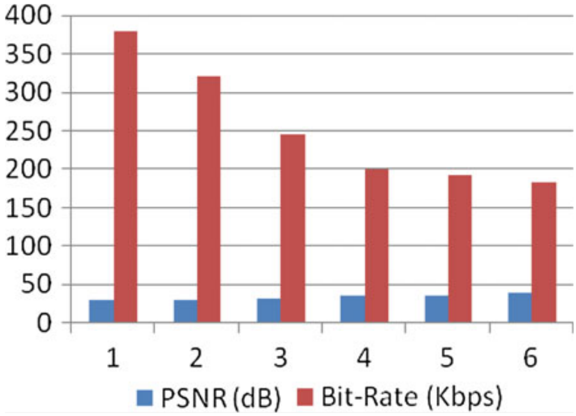


**Fig. 3** Values of optimized PSNR versus bit rate in different techniques (color figure online)

in the bit rate with slight reduction in the PSNR value. In Fig. 3 the techniques with highest bit rate have reduced PSNR and vice versa. But the values of extreme, i.e., PSNR and bit rate are not the region of interest to us. The techniques which produce a better bit rate for the moderate PSNR values give us a better Rate-Distortion with highest bit rate.

## 5 Conclusion and Future Work

The main goal of measuring distortion is to use the different models to achieve good quality video sequences and at the same time it should yield a better bit rate of transmission. On this view, the results show that the inter-layer prediction and parametric Rate-Distortion with Lagrangian multiplier perform the best. These methods produce the results with higher complexity. This can be an objective to reduce complexity in the future work and able to derive a successful technique in obtaining lower Rate-Distortion with higher bit rates.

## References

1. Sun, Jun, Wen Gao, and Debin Zhao. 2009. On rate-distortion modeling extraction of H.264/SVC fine-granular SNR scalable video. *IEEE Transactions on Circuits and Systems for Video Technology* 19 (3).
2. Mansour, Hassan, Vikram Krishnamurthy, and Panos Nasiopoulos. 2008. Rate and Distortion Modeling of Medium Grain Scalable Video Coding. *Proc ICIP08*.
3. Thomas, Nithin, David Bull, and David Redmill. 2008. *A Novel H.264 Scalable Video Coding Encryption Scheme for Secure Bit-Rate Transcoding*. UK: University of Bristol.
4. Sun, Jun, Wen Gao, and Debin Zhao. 2007. Statistical Analysis and Modeling of Rate-Distortion Function in Scalable Video Coding Fine-Granular SNR Scalable Videos. In *Proc ICME 2007*, 1135–1138.
5. Li, Xiang, Peter Amon, Andreas Hutter, and André Kaup. Lagrange Multiplier Selection for Rate-Distortion Optimization in Scalable Video Coding. *Picture Coding Symposium, (PCS)*.
6. Slanina, Martin, Michal Ries, and Janne Vehkaperä. 2013. Rate Distortion Performance of H.264/SVC in Full HD with Constant Frame Rate and High Granularity. In: *Proc ICDT 2013, the Eighth International Conference on Digital Telecommunications*, 7–13.
7. Wang, Hongtao, Dong Zhang, and Houqiang Li. 2015. *Rate-Distortion Optimized Coding Method for Region of Interest in Scalable Video Coding*. *Advances in Multimedia*, vol. 2015. Hindawi Publishing Corporation, 1–11 p.
8. Xu, Long, Wen Gao, Xiangyang Ji, Debin Zhao, and Siwei Ma. 2007. *Rate Control for Spatial Scalable Coding in Scalable Video Coding*. Chinese Academy of Sciences 2007.
9. Zhu, Tao, and Xiong-wei Zhang. 2011. *Rate Control Scheme for Spatial Scalability of H.264/SVC*. PLA University of Science & Technology.
10. Cho, Yongju, Hayder Radha, Jeongju Yoo and Jinwoo Hong. 2008. A Rate-Distortion Empirical Model for Rate Adaptive Wireless Scalable Video. In *CSIS 2008*.
11. Gan, Tong, Bart Masschelein, Carolina Blanch, Antoine Dejonghe, and Kristof Denolf. 2009. Rate-Distortion Complexity Performance Analysis of the Scalable Video Coding decoder. In *Proc: ICME 2009*, 213–216.
12. Grois, Dan, and Ofer Hadar. 2011. Efficient Adaptive Bit-Rate Control for Scalable Video Coding by Using Computational Complexity-Rate-Distortion Analysis. In *Proc MM* 11–07.
13. Gupta, Mohan, Akshay Pulipaka, Patrick Seeling, Lina J. Karam, and Martin Reisslein. 2012. H.264 Coarse Grain Scalable (CGS) and Medium Grain Scalable (MGS) Encoded Video: A Trace Based Traffic and Quality Evaluation. *IEEE Transactions on Broadcasting* 58 (3): 428–439.
14. Haseeb, Abdul, Maria G. Martini, Sergio Cicalò and Velio Tralli. 2012. Rate and Distortion Modeling for Real-Time MGS Coding and Adaptation. *Wireless Advanced (WiAd)* 85–89.
15. Peng, Wen-Hsiao, John K. Zao, Hsueh-Ting Huang, Tse-Wei Wang, and Lun-Chia Kuo. 2008. Multidimensional Scalable Video Coding Bit-Stream Adaptation and Extraction for Rate-Distortion Optimized Heterogeneous Multicasting and Playback. In *Proc ICIP 2008*, 2476–2479.
16. Sun, Jun, Wen Gao, Debin Zhao, and Weiping Li. 2009. On Rate-Distortion Modeling and Extraction of H.264/SVC Fine-Granular Scalable Video. *IEEE Transactions on Circuits and Systems for Video Technology* 19 (3): 323–336.
17. Tsai, Chia-Yang, and Hsueh-Ming Hang. 2009. *Rate-Distortion Model for Motion Prediction Efficiency in Scalable Wavelet Video Coding*. National Chiao University, Hsinchu.
18. Winken, Martin, Heiko Schwarz, and Thomas Wiegand. 2008. Joint Rate-Distortion Optimization of Transform Coefficients for Spatial Scalable Video Coding Using SVC. In *Proc ICIP 2008*, 1220–1223.