Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Better Compression Rate For Images With The Deployment of ATSVD

A. Sharifinejad

The University of New South Wales, Sydney 2052, Australia
ashar@ieee.org

Abstract. The reduction of output bitrate of video source (of I-frames) and consequently the improvement of multiplexer's gain are the main target of this paper. In reducing the bitrate, the SVD transform coding as an attractive alternative to the DCT coding was adopted. The SVD transformation just like DCT is a lossy image compression, but it can achieve a higher rate. The rank of SVD matrix transform was limited to maximum 21% of original value. Then our new algorithm so called ATSVD was introduced in further adaptively reducing the bitrate according to the details of image. Consequently, the output bitrate of video source was reduced to a fraction of its original value.

1 Introduction

In this paper, a congestion recovery algorithm was proposed to improve the delivery of real-time multimedia, more specifically video streams in a more effective manner while sustaining the video quality subsequent to packet and cell losses. The high-bandwidth requirement, real-time delivery constraints, enhancement of congestion control policies in high-speed networks and maintenance of high-quality videos motivated the design for further improvement in multimedia delivery techniques. Except for a small number of traffic management and congestion control policies that mainly attempt to deal with the sources of congestion, majority of proposed techniques aim rather to distribute the network resources so as to have a fairer operating system for all applications. Take the video rate-change based transcoders and video coarse vector quantization techniques as examples of the source dealt congestion control system whereby many proposed techniques [1], [2], [3], so far, are experienced problems such as excessive-delay response and access latency, inappropriate for real-time multimedia applications, progressive and unmanaged degradation and deterioration of video quality, the need for a large storage capacity, incompatible for high-speed network and the need for a high computing power. The challenge in designing an effective traffic management algorithm for the real-time multimedia over high-speed networks is to consider some compromises at source as well as network level, compromises such as fast processing and delivery speed verses high-standard multimedia quality or jitter-free bit stream, etc. Of course, video streams with smooth, predictable, error resilient, scalable and most importantly low bit-rate are by far the most desirable feature of any source generated multimedia streams. On the other hand, flexibility in bandwidth allocation, guaranteed quality of service, maximization of connection per requested call, none or short queuing delay, multiple protocol exchangeability and support, etc. can be encountered as the most attractive advancements in network delivery system.

In response to network requirements, dynamic bandwidth allocation algorithms that can adaptively respond to traffic variations have been proposed [4], [5], and [6] in order to overcome the above difficulties. Tanthawichian et al. [7] proposed a dynamic bandwidth allocation based on a heuristic approach made up of two functions, a time ε-quartile function to characterize the source behaviour and a function to bound the amount of bandwidth served by the multiplexer. Others such as Benjapolakul et al. [8] used a neural network method to propose an aggregate bandwidth allocation scheme for heterogeneous sources in the ATM network. The problem with these proposals is due to occasional under- and overestimation of the actual required bandwidth that yields to inefficient use of bandwidth, network congestion and/or loss of data. Liu at el [9] and others [10], [11] utilized the autocorrelation between bitrate of multimedia streams to create a traffic model. Because of the diversity of traffic models proposed for source generated multimedia streams [12], [13], [14], etc., hence, we cannot reliably implement an implicit traffic policy for high-speed networks including ATM network services based on these models.

In dealing with the sources of video traffic, many components of the video compression structure can be addressed to increase the compression ratio individually and independently, so as to reduce the transmission bitrate for the same faithful reconstructable video, and consequently and particularly for ATM network increase the gain of ATM statistical multiplexer by increasing the number of call admission per channel. In general, the throughput of the system will dramatically increase. The hierarchical structure of video encoders whose generating the traffics enabled the researchers to tackle effectively each module in isolation, modules such as redundant bit removal, variable length coding, variable quantization and etc. Although the order of processing modules is fixed, the details of each module can be manipulated independently of others as long as the output of module conforms to the input structure of the subsequent modules. Various types of video application employing different encoding procedure result in different bitrate characteristic. However, the most predominant encoder of all, MPEG, is open to any modification if only the structure of output bit stream remains recognizable by the standard. The improvement of the transform coding can play major role not only in achieving lower output bitrate but also the quality of decoded video, speed of encoding and the magnitude of SNR.

A typical image or video coding system consists of a number of components such as analyzer, quantizer, and entropy coder, with a number of possibilities for each component. Analyzer can operate on entire or portion of the signal, or on error predicted signal by one of many available mechanisms such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Fast Fourier Transform (FFT), Discrete Fourier Transform (DFT), Sub-bands or Singular Value Decomposition Transform (SVD). These transform methods in the context of video coding operate on individual blocks of frames. Among all mentioned transform methods, most commonly adopted technique, DCT, has gained a lot of popularity because of its relatively high speed. The DCT has relatively only a good energy compaction property, nevertheless, it is widely used for image coding. Also, the DCT transform can be computed efficiently, the basis functions are fixed and only the transform coefficients are quantized for the storage or transmission.

In contrast, the use of singular value decomposition (SVD) in image compression is motivated by its relatively excellent energy compaction property. Moreover, the SVD transformation has the characteristic of optimal energy compaction in the least square sense, which makes it most useful for the bandwidth compression coding ignoring its computational complexity. The disadvantage of SVD transformation is due to its need for the recalculation for each subimage. However, SVD, unlike some widely used spectral algorithms in signal processing such as DFT, DCT and DWT, has a boosted performance for the spikes and abrupt jumps of the input signals, and can be applied to heterogeneous and M-dimensional vectors. The SVD is often considered as the most potential candidate for transformation but is seldom used in practical applications. The singular vectors in SVD are image dependent, and must be, therefore, coded along the associated singular values. To exploit the optimal energy compaction properties of the SVD, most of the effort in designing a more efficient SVD coder was put into the effective coding of singular values and singular vectors as well as the reduction of computational cost.

2 SVD and Video Compression

Images are nothing but two-dimensional matrices from the mathematical point view. Moreover, SVD takes an image in a rectangular m×n matrix A format and calculates three matrices of U, S, and V where U and V are unitary matrices with dimensions of m×m and n×n, and S is a diagonal m×n matrix (the same dimensions as matrix A), respectively.

$$\vec{\mathbf{U}} = \left[\overrightarrow{u_1} \quad \overrightarrow{u_2} \quad \cdots \quad \overrightarrow{u_k} \quad \cdots \quad \overrightarrow{u_m} \right]$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2k} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} & \cdots & a_{mn} \end{bmatrix}$$

$$\vec{\mathbf{V}}^{\mathbf{T}} = \left[\overrightarrow{v_1} \quad \overrightarrow{v_2} \quad \cdots \quad \overrightarrow{v_k} \quad \cdots \quad \overrightarrow{v_n} \right]^T$$

Calculating the SVD consists of finding the eigenvalues and eigenvectors of A^TA. The eigenvectors of A^TA make up the columns of V (i.e. v_i) and the eigenvectors of AA^T make up the columns of U (i.e. u_i). The eigenvalues of A^TA or AA^T are the squares of the singular values for A. The singular values are the diagonal entries of the S matrix and are arranged in descending order. The singular values are always real numbers and if the matrix A is a real matrix, then U and V are also real. In order to calculate the eigenvalues of matrix A, it is required to solve the characteristic polynomials of:

$$\det(A^T A - \lambda_i . I) = 0 \qquad (i = 0, \dots, n) \tag{1}$$

where I is an n×n identity matrix, $\lambda_1 \dots \lambda_n$ are the eigenvalues and 0 is zero vector of n×1 dimension. Then, the columns of eigenvectors V can be computed by solving the following characteristics equations:

$$(A^T A - \lambda_i . I) . v_i = 0 \qquad (i = 0, \dots, r) \tag{2}$$

The columns of the eigenvector U can be found by:

$$\vec{\mathbf{U}} = \begin{bmatrix} \frac{1}{\delta_1}\mathbf{A}\overrightarrow{v_1} & \frac{1}{\delta_2}\mathbf{A}\overrightarrow{v_2} & \cdots & \frac{1}{\delta_k}\mathbf{A}\overrightarrow{v_k} & \cdots & \frac{1}{\delta_n}\mathbf{A}\overrightarrow{v_n} \end{bmatrix}$$

$$\text{or} \quad \mathbf{u_i} = \frac{1}{\sqrt{\boldsymbol{\lambda}_\mathbf{i}}}\mathbf{A}\mathbf{v_i} = \frac{1}{\boldsymbol{\delta}_\mathbf{i}}\mathbf{A}\mathbf{v_i} \qquad (i = 0, \dots, r)$$

The singular value δ_i is equal to the square root of the descending ordered eigenvalues, λ_i, of $A^T A$ and, therefore, organized in descending ordered of magnitude $\delta_1 > \delta_2 > \dots > \delta_r > \delta_{r+1} = \dots = \delta_n = 0$. Thus, the result of the transform in terms of singular value matrix is given by:

$$S = U^T A V$$

$$\mathbf{S_r} = \begin{bmatrix} \delta_1 & 0 & & & \cdots & & 0 \\ 0 & \delta_2 & 0 & & & & \\ & 0 & \ddots & & & & \\ & & & \delta_r & & & \vdots \\ \vdots & & & & 0 & & \\ & & & & & \ddots & \\ 0 & & & \cdots & & & 0 \end{bmatrix}$$

and the SVD representation of the block A is:

$$A = USV^T$$

U and V are orthogonal matrices and S is a diagonal matrix with the singular values along the main diagonal with rank r. The image block A can also be written as the column vectors of eigenvalues by:

$$\mathbf{A} = \sum_{\mathbf{i=1}}^{\mathbf{r}} \mathbf{u_i \delta_i v_i^T} \tag{3}$$

3 Truncated SVD

Singular value decomposition matrix **S** with rank r has very heavy computational burden imposing a major drawback in practical applications such as image and video signal processing for real time multimedia communications, therefore, to reduce this burden one possibility is to approximate this transformation by quantization and/or truncation of the singular values by parameter k < r.

$$S_r = \begin{bmatrix} \delta_1 & 0 & & & \cdots & 0 \\ 0 & \delta_2 & 0 & & & \\ & 0 & \ddots & & & \\ & & & \delta_k & & \vdots \\ \vdots & & & & \ddots & \\ & & & & \delta_r & \\ 0 & & & \cdots & & 0 \end{bmatrix} \approx S_k = \begin{bmatrix} \delta_1 & 0 & & & \cdots & 0 \\ 0 & \delta_2 & 0 & & & \\ & 0 & \ddots & & & \\ & & & \delta_k & & \vdots \\ \vdots & & & & 0 & \\ & & & & & \ddots \\ 0 & & & \cdots & & 0 \end{bmatrix}$$

So, the approximated version of equation (3) can be represented by:

$$\hat{\mathbf{A}} = \sum_{\mathbf{i=1}}^{\mathbf{k}} \mathbf{u_i \delta_i v_i^T} \tag{4}$$

Hence, the approximation error matrix $\mathbf{E_K}$ is dependent on the performance accuracy of the quantization and/or truncation by parameter k, which can be described by:

$$\mathbf{E_k} = \mathbf{A} - \hat{\mathbf{A}} \tag{5}$$

And the second norm of approximation error is calculated by:

$$\mathbf{E}_{\mathbf{k}}^{2} = \left\| \mathbf{A} - \hat{\mathbf{A}} \right\|_{2} = \left\| \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{r}} \mathbf{u}_{\mathbf{i}} \boldsymbol{\delta}_{\mathbf{i}} \mathbf{v}_{\mathbf{i}}^{\mathbf{T}} - \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{k}} \mathbf{u}_{\mathbf{i}} \boldsymbol{\delta}_{\mathbf{i}} \mathbf{v}_{\mathbf{i}}^{\mathbf{T}} \right\|_{2} = \left\| \sum_{\mathbf{i}=\mathbf{k}+\mathbf{1}}^{\mathbf{r}} \mathbf{u}_{\mathbf{i}} \boldsymbol{\delta}_{\mathbf{i}} \mathbf{v}_{\mathbf{i}}^{\mathbf{T}} \right\|_{2}$$

That is:

$$\mathbf{E}_{\mathbf{k}}^{2} = \sum_{\mathbf{i}=\mathbf{k}+\mathbf{1}}^{\mathbf{r}} \boldsymbol{\delta}_{\mathbf{i}}^{2} = \sum_{\mathbf{i}=\mathbf{k}+\mathbf{1}}^{\mathbf{r}} \boldsymbol{\lambda}_{\mathbf{i}} \tag{6}$$

The range of 19 to 21 percent SVD matrix rank was determined to enable a graceful non-deteriorated reconstructable video, while no significant visual improvement was observed beyond 21 percent. To adopt the best image quality for the Truncated Singular Value Decomposition (TSVD) and benefit from the maximum compression, that is, the lowest achievable bit rate, meanwhile keeping the mean square error relatively minimum, the TSVD algorithm was set to 21 percent of original rank.

4 ATSVD Video Encoder

The performance of SVD in terms of transformation speed ultimately has an impact on the encoder’s processing time and the output bitrate of encoded sequence of images. This performance is very much dependent on the size of the image and the chosen operation rank for the singular value decomposition matrix. The larger the image size, the longer processing time the SVD transformation coding takes. So, two phases were realized in developing our final proposed algorithm (i.e. ATSVD). Firstly, by applying the TSVD on full image size, we need to verify its capability on bitrate reduction and maintenance of video quality. The development of Adaptive Truncated Singular Value Decomposition (ATSVD) emerged by noticing the problem with the use of SVD in terms long processing time associated with its compression rate. As mentioned earlier, truncation of SVD matrix to 21 percent of its original rank has an acceptable outcome in terms of video quality and achieving a lower output bitrate but not for the processing time. The significant of this investigation is that the acquired percentage value for truncated matrix rank is valid for all image sequences regardless of their sizes. Secondly, applying ATSVD on smaller blocks of image reduces the processing time and the bitrate even further while maintaining all other beneficial aspects of TSVD intact. This is because smaller blocks may generally contain less luminance variation, that is, the standard deviation of this variation is very narrow; therefore, the rank of SVD in transforming the smaller sized blocks can shrink even further. This is not a trade-off situation and in fact is necessary for obtaining a good rate of compression. Therefore, in order to adopt into the potential variation range of truncation values for each block, the first two blocks were set to the upper limit of 21 percent of SVD original rank.

To determine the optimum truncation rank for the rest of the blocks containing different features, the SVD truncation values were adaptively obtained based on applying a second order autoregressive function AR(2) to the resulting error functions (e.g. mean square error, mean absolute error, mean error, etc.). These error measurements are obtained from the encoded and original block mismatches for the two previously transformed neighboring blocks. In ATSVD algorithm, the coefficients of AR(2) terms were empirically determined and set to 0.62 and 0.38 for the closest and second closest neighboring blocks to the current block, respectively. The chosen coefficient values were fine tuned on the basis of the correlation between the further neighboring blocks are less influential than the nearer neighboring blocks. Longer order of autoregressive function can be employed in this procedure for more possible precision in achieving better image quality at the cost of processing time increase. But as one of our primary objective to win over any increase in processing delay, the second order autoregressive function was adopted in here. AR process can be described by:

$$z(n) = \sum_{i=1}^{p} \phi_i z(n-i) + a(n)$$

A second order AR is chosen in ATSVD algorithm (i.e. p = 2). Φ_1 and ϕ_2 are the empirically determined coefficients as 0.62 and 0.38, respectively. Z(n-i) is the error values obtained as a result of mismatch between the original block value and the block transformed by the TSVD. a(n) is an independent shock value which was set to zero for this algorithm.

5 Experimental Results and Discussion

Our proposed ATSVD algorithm was applied on our five test sequences and the results were compared with SVD and DCT. The results for different block sizes of 8 and 16 are tabulated in I and II, respectively. In addition to the five test video sequences (of Independent I-frames only) the algorithm was tested on a number of other video sequences with different frame sizes to verify whether or not the results are consistence for any image dimension. Only one video with 320×240 frame size is shown in Figure 1 due to space constraint. It is worth to note that for some cases the rank of 19 percent produces the original frame with no noticeable visual degradation subjectively for those videos (which is even less than the rank of 21).

Examining the table entries reveals that ATSVD has a better performance for the smaller block sizes in terms of low processing time (PT) and low bitrate. The loss of around 1 or 2 dB on PSNR in comparison to gain in shorter processing time and higher compression rate in regard to SVD and DCT is negligible. The loss in PSNR value and gain in computation complexity are due to adaptive rank truncation of the singular value decomposition matrix. In achieving to produce a lower bitrate for video source coders by the deployment of ATSVD, more number of video streams can now be multiplexed per channel. In fact, the result signifies that for the same allocated bandwidth, 2 or 3 additional video streams can now be accommodated with the utilization of ATSVD. Hence, fewer channels for multimedia streams are required, that is, the number of call connections per channel is increased. This is noticeable for high-speed network such as ATM where the gain of the ATM statistical multiplexer is increased. The performance of statistical multiplexing gain can also be evaluated based on the buffer size and queuing delay where a larger buffer size implies to a larger service delay for the individual quantities of traffic (i.e. cells for ATM network), which in turn implies to a smaller statistical multiplexing gain. Now, for the fixed multiplexing buffer size and the same number of traffic sources, the statistical multiplexing gain would be greater for the sources with a lower output bitrate than the otherwise, and that is what has been achieved in here.

8×8	Average per frame	Lecture	Icehouse	Ronin	Indiana Jones	Speed
DCT	Bitrate (KB)	5.614	5.774	5.655	5.875	5.8592
	PSNR	28.47	30.63	30.78	32.18	31.82
	PT (%)	100	100	100	100	100
	Total Size (KB)	449.12	461.95	452.40	470.06	468.74
SVD	Bitrate/f (KB)	3.563	3.845	3.761	3.893	3.852
	PSNR	29.83	31.87	31.60	32.74	31.96
	PT (%)	201	204	203	205	204
	Total Size (KB)	285.09	307.56	300.84	311.47	308.19
ATSVD	Bitrate/f (KB)	1.436	1.491	1.459	1.512	1.508
	PSNR	28.02	29.54	29.22	30.84	30.18
	PT (%)	95	100	101	101	102
	Total Size (KB)	114.90	119.24	118.71	120.97	120.64

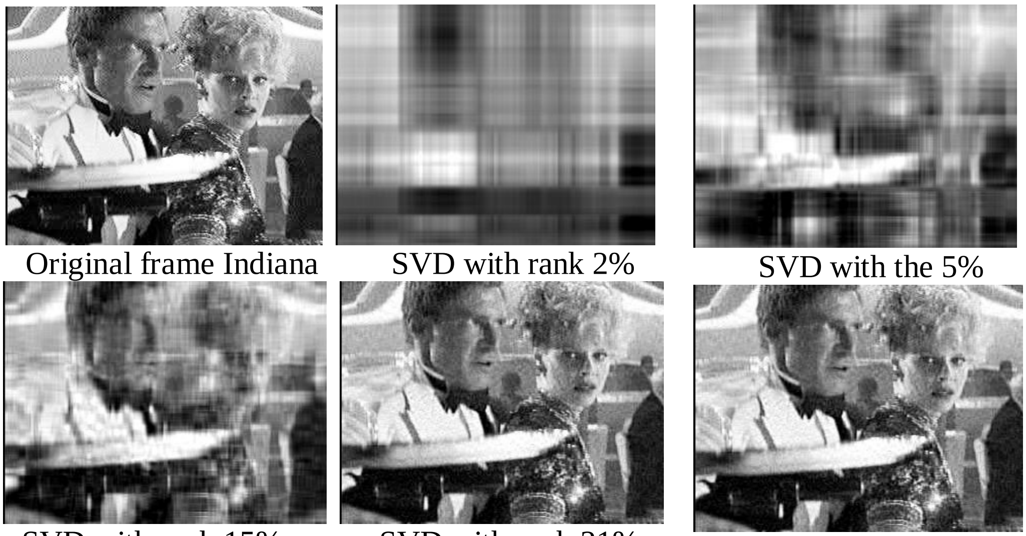
Table I

Obtained results for 5 different test series encoded with given algorithm (block 8×8)

16×16	Average per frame	Lecture	Icehouse	Ronin	Indiana Jones	Speed
DCT	Bitrate (KB)	5.428	5.517	5.371	5.655	5.559
	PSNR	28.88	31.10	30.97	32.89	32.15
	PT (%)	100	100	100	100	100
	Total Size (KB)	434.24	441.36	429.68	452.40	444.72
SVD	Bitrate/f (KB)	3.714	4.079	4.019	4.176	4.007
	PSNR	29.01	31.53	31.98	33.15	32.87
	PT (%)	441	440	447	446	449
	Total Size (KB)	297.11	326.33	321.50	334.08	320.52
ATSVD	Bitrate/f (KB)	1.732	1.784	1.804	1.926	1.878
	PSNR	28.67	29.81	29.63	30.95	30.40
	PT (%)	100	107	104	109	105
	Total Size (KB)	138.57	142.70	144.33	154.07	150.21

Table II

Obtained results for 5 different test series encoded with given algorithm (block 16×16)



Original frame Indiana | SVD with rank 2% | SVD with the 5%

SVD with rank 15% | SVD with rank 21% | ATSVD with adaptive rank AR(2)

Figure 1
Quality comparison of "Indiana Jones" scene with different algorithm

6 Conclusion

The reduction of output bitrate of video source (I-frames) and consequently the increase of call connections per communication channel were the sole aim of this experiment presented in this paper. Moreover, in reducing the bitrate, the SVD transform coding as an attractive alternative to the DCT coding was investigated. However, the required processing time and high computational complexity have a crucial impact on the selection of transform coding method for any video encoder. To utilize the high compression capability of the SVD transformation while diminishing the processing delay in consequence of the SVD deployment, a new algorithm called ATSVD was introduced. In our proposed algorithm, the SVD was truncated adaptively based on the second order autoregressive function with the upper limit was set to 21 percent of the original SVD rank of the image dimension. The range of 19 to 21 percent SVD matrix rank was determined to enable a graceful non-deteriorated reconstructable video, while no significant visual improvement was observed beyond 21 percent. The significant of this investigation is that the acquired percentage value for truncated matrix rank is valid for all image sequences regardless of their sizes.

As a result, further errors were introduced due to truncation of the SVD matrix rank besides the losses already imposed by the deployment of the SVD lossy compression process, however, objectively the difference in the PSNR measurement values between this method and other methods is negligible. Also, these errors are not subjectively noticeable in the decoded video. More importantly, the bitrate of video source is reduced by the incorporation of ATSVD into video coders, and consequently, the processing time for real-time multimedia applications over the high-speed network has been improved.

References

1. O. H. Werner, "Generic quantizer for transcoding of hybrid video", *Proceeding of picture coding symposium '97*, Berlin, Sep. 1997, 10-12.
2. P. Tudor and O. Werner, "Real-time transcoding of MpEG-2 video bit streams", *Proceedings of the international broadcast convention (IBC)*, Amsterdam, 1997, p.p. 206 – 301.
3. G. Keesman, "Transcoding of MPEG bitstreams", *Signal processing image communication*, Vol. 8, 1996, p.p. 481 – 500.
4. F.M. Porikli and Z. Sahinoglu, "Dynamic bandwidth allocation with optimal number of renegotiations in ATM networks", *Proceedings tenth international conference on computer communications and networks*, 2001, p.p. 290 – 295.
5. K. Jung-Taek and I. Koh, "A development of an intelligent algorithm for bandwidth allocation in ATM networks using Petri Nets", *2001 IEEE international conference on Systems, Man, and Cybernetics*, Vol. 2, 2001, p.p. 1131 – 1136.
6. M.C. Yuang and P.L. Tien, "Multiple access control with intelligent bandwidth allocation for wireless ATM networks", *IEEE Journal on selected areas in communications*, Vol. 18, Issue 9, Sept. 2001, p.p. 1658 – 1669.
7. P. Tanthawichian, A. Fujii and Y. Nemoto, "Bandwidth allocation in ATM networks: Heuristic approach", *IEEE Proceedings 7th international conference on computer communications and networks*, 1998, p.p. 20 – 25.
8. W. Benjapolakul and T. Rangsihiranrat, "Aggregate Bandwidth allocation of heterogeneous sources in ATM networks with guaranteed quality of service using a well-trained neutral network", *IEEE Asia-Pacific conference on circuits and systems*, 2000, p.p. 384 – 351.
9. H. Liu, N. Ansari and U. Q. shi, "Markov-modulated self-similar processes MPEG coded video traffic modeler and Synthesizer", *Globecom Telecommunication conference*, Globecom'99, p.p. 1184 – 1188.
10. H.-W. Ferng and J.-F. Chang, "Characterization of the output of an ATM output buffer receiving self similar traffic", *Global telecommunications conference, IEEE Globecom'01*, Vol. 4, 2001, p.p. 2650 – 2653.
11. S. Chandramathi and S. Shanmugavel, "A novel fuzzy approach to estimate cell loss probability for self similar traffic in ATM networks", Proceedings sixth *IEEE symposium on computers and communications*, 2001, p.p. 260 – 265.
12. D. P. Heyman and T. V. Lakshman, "Source models for VBR broadcast-video traffic", *IEEE/ACM Transactions on Networking*, Vol. 4, no.1, June 1996, p.p. 301- 317.
13. D.P. Heyman, "The GBAR source model for VBR videoconferences", *IEEE/ACM transactions on networking*, Vol. 5, No. 1, Aug. 1997, p.p. 554-560.
14. H. Liu, N. Ansari and U. Q. shi, "A simple model for MPE video traffic", *Proceedings of IEEE international conference of multimedia and expo*, New York city, July 2001.