International Journal of Computer Applications (0975 – 8887)
Volume 65– No.8, March 2013

Efficient Design of Multimedia Transcoding for Clientside Application over Mobile Interface

Shakti Awaghad
Research Scholar, Dept of Electronics & Communication
G.H. Raisoni College of Engg, Nagpur, India

Sanjay Pokle, PhD.
Professor and Head: Dept. of Electronics & Communication, R.C.O.E.M, Nagpur, India

ABSTRACT
The proposed system introduces a novel framework that can optimize the bandwidth and maintain better visual qualities by performing transcoding on the multimedia contents mainly targeted over the mobile devices. The system designed in Matlab considers both image and video as input file and performs multimedia transcoding considering BMP, GIFF, and JPEG formats supportability on mobile devices. The framework considers designing linear rate regulation policy for enhancing the estimation accuracy and minimizing computational complexities. Results shows that there is a drastic decrease in download time when transcoding is performed maintaining better visual quality estimated by PSNR.

Keywords
Transcoding, Multimedia File sharing, rate and distortion, transform coding, video coding and transmission

1. INTRODUCTION
Recent advances in computing and communication technology have stimulated the research interest in digital techniques for recording and transmitting visual information. The exponential growth in the amount of visual data to be stored, transferred, and processed has created a huge need for data compression [1]. Compression of visual data, such as images and videos, can significantly improve the utilization efficiency of the limited communication channel bandwidth or storage capacity. The demand for image and video compression has triggered the development of several compression standards, such as JPEG [2], JPEG-2000 [3], MPEG-2 [4], H.263 [5], and MPEG-4 [6]. In both the compression standards and the algorithms reported in the literature, transform coding has become the dominant approach for image and video compression. The transform, either discrete wavelet transform (DWT) or discrete cosine transform (DCT), is applied to the input picture. Here, a picture can be either a still image or motion-compensated video frame. After quantization, the quantized coefficients are converted into symbols according to some data representation scheme [7][8]. In transform coding of images and videos, the two most important factors are the coding bit rate and picture quality. The coding bit rate determines the channel bandwidth required to transfer the coded visual data. One direct and widely used measure for the picture quality is the mean-square error (MSE) between the coded image/video and the original one. The reconstruction error introduced by compression, often referred to as distortion. In typical transform coding, both rate and distortion are controlled by the quantization parameter of the quantizer. The major issue here is how to determine the value of to achieve the target coding bit rate, or target picture quality. To this end, we need to analyze and estimate the R-D behavior of the image/video encoder; this behavior is characterized by its rate-quantization (R-Q) and distortion-quantization (D-Q) functions, R(q) and D(q), respectively [9], [10]. It is well known that the R-D behavior of an image/video encoder is determined both by the characteristics of the input source data and by the capability of the coding algorithm to explore these characteristics. Analysis and estimation of the R-D functions have important applications in visual coding and communication. First, with the estimated R-D functions we can adjust the quantization setting of the encoder and control the output bit rate or picture quality according to the channel condition, the storage capacity, or the user's requirement [11]–[12]. Second, based on the estimated R-D functions, optimum bit allocation as well as other R-D optimization procedures can be performed to improve the efficiency of the coding algorithm and, consequently, to improve the image quality or video presentation quality. We develop a generic source modeling framework for transform coding of images and videos by the following two major steps. In the first step, we introduce the concepts of characteristic rate curves and rate curve decomposition to characterize the input source date and to model the coding algorithm, respectively. In the second step, we propose a linear regulation scheme to improve the accuracy and robustness of the R-D estimation. The proposed system discusses about a model that performs transcoding on multimedia contents especially targeted on mobile devices. In section 2 we give an overview of related work which identifies all the major research work being done in this area. Section 3 highlights about the transcoding system describing some of the prominent attributes system. Section 4 discusses the significance of the transcoding of multimedia contents. Proposed system is discussed in Section 5 followed by implementation and result discussion on Section 6 and finally in section 7 we make some concluding remarks.

2. RELATED WORK
There has been a substantial amount of the work being done of transcoding witnessed in the prior literature. Yang e.t. al. [13] have presented a rate control scheme for MPEG-2 to H.264 transcoder with the consideration of peak signal-to-noise ratio (PSNR) fluctuation control and re-quantized quality loss reduction. Yasmin e.t. al [14] have presented a scheme for transcoding document images for presentation on handheld devices like PDA's, e-books etc. The paper has presented a scheme for transcoding Devnagari script based Indian language documents. Zhang, Perkis, and Georganas [15] have reviewed new features of H.264/AVC. Brandt and Wolf [16] have proposed a new video transcoding Scheme that posses a multidimensional transcoding approach for MPEG-4 encoded video, which smartly combines existing transcoding techniques to enable fine grain adaptation for different video devices. Vetro [17] presents two types of techniques suitable for rate reduction transcoding for wireless

video streaming applications. The first type of transcoder is based intra refresh architecture and is able to offer robust transcoding from MPEG-2 format to reduced spatial resolution MPEG-4 format. The second type of transcoder is based R-D function of the pre-encoded video and is suitable for a variety of transcoding tasks because of its simplified architecture, its high performance rate control, smart frame dropping, and channel adaptive streaming strategy. Chandra e.t. al [18] has presented a work to understand the nature of typical Internet images and their transcoding characteristics. The authors showed that most GIF images accessed on the Internet are small; about 80% of the GIF images are smaller than 6 KBs. JPEG images are larger than GIF images; about 40% of the JPEG images are larger than 6 KBs. It was also shown that for JPEG images, the JPEG compression metric and a transcoding that reduces the spatial geometry are productive transcoding operations (saves at least 50% of the file size for 50% of the images). It can be visualized from the literature that majority of the transcoding techniques are highly applicable on video as multimedia contents, however, our proposed work will equally put emphasis on both video as well as image as multimedia contents. From the literature, it can be seen that a straightforward realization of a transcoder to cascade a decoder and an encoder: the decoder decodes the compressed input video and the encoder re-encodes the decoded video into the target format. It is computationally very expensive. Therefore, reducing the complexity of the straight-forward decoder-encoder implementation is a major driving force behind many research activities on transcoding.

3. TRANSCODING

Transcoding is the direct digital-to-digital data conversion of one encoding to another, such as for movie data files or audio files. The key drawback of transcoding in lossy formats is decreased quality. Compression artifacts are cumulative, so transcoding causes a progressive loss of quality with each successive generation, known as digital generation loss. For this reason, transcoding is generally discouraged unless unavoidable. Figure 1 shows the schema of Transcoder.

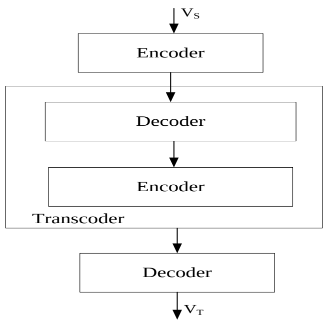

Figure 1 Schema of Transcoder

For image editing users are advised to capture or save images in a raw or uncompressed format, and then edit a copy of that master version, only converting to lossy formats if smaller file sized images are needed for final distribution. As with audio, transcoding from lossy format to another format of any type will result in a loss of quality.

For video editing, images are normally compressed directly during the recording process due to the huge file sizes that would be created if they were not, and because the huge storage demands being too cumbersome for the user otherwise. However, the amount of compression used at the recording stage can be highly variable, and is dependant on a number of factors, including the quality of images being recorded (e.g. image resolution), and type of equipment available to the user, which is often related to budget constraints – as highest quality digital video equipment, and storage space, may be very expensive. Effectively this means that any transcoding will involve some cumulative image loss, hence the best practicable solution is that the original format of recording is deemed the master copy, and subsequent finished versions (that are often in another format and/or at a much smaller file size) are edited from copies of that master

4. SIGNIFICANCE

Quality of experience is used to measure the real efficiency of the transcoding algorithms and video coding algorithms in general. User experience is dependent on both the video and audio quality of the streamed content. Depending on the content type (news shows, sport games, movies, etc.) people react differently to the proportion of video/audio quality. In particular lower video quality is acceptable in a news show if audio is of good quality. Opposite to that lower audio quality with good video would be acceptable for some sport games [3]. Delay time is also important factor that needs to be carefully planned not to degrade the quality of experience. After requesting a stream from a server and defining its capabilities, end user has to wait some time for the transcoding process to begin and for the transmission of the content to his device. With some perceptual tricks like displaying the channel logo, waiting time can be shorten, but just perceptually [4]. Transcoding algorithms should be able to provide a lower bandwidth video when degradation occurs on the available network bandwidth. This will help from video being stopped and it will be played at lower quality. Unfortunately this approach won't help if the bit-error or jitter is high. Researches of quality of experience for the standard devices like desktop computers, video consoles and television has been done but further researches is needed to see the difference concerning mobile devices.

The first and most important challenge in the context of a video conferencing is to provide transcoding on the fly with real-time speed and without any interruption of video flow [5], [6]. There are three basic requirements in transcoding [7]: 1) the information in the original bitstream should be exploited as much as possible; 2) the resulting video quality of the new bitstream should be as high as possible, or as close as possible to the bitstream created by coding the original source video at the reduced rate; 3) in real-time applications, the transcoding delay and memory requirement should be minimized to meet real-time constraints. A video transcoder can provide several functions, including adjustment of bit rate and format conversion. We illustrate these functionalities and their classification in Fig. 2.

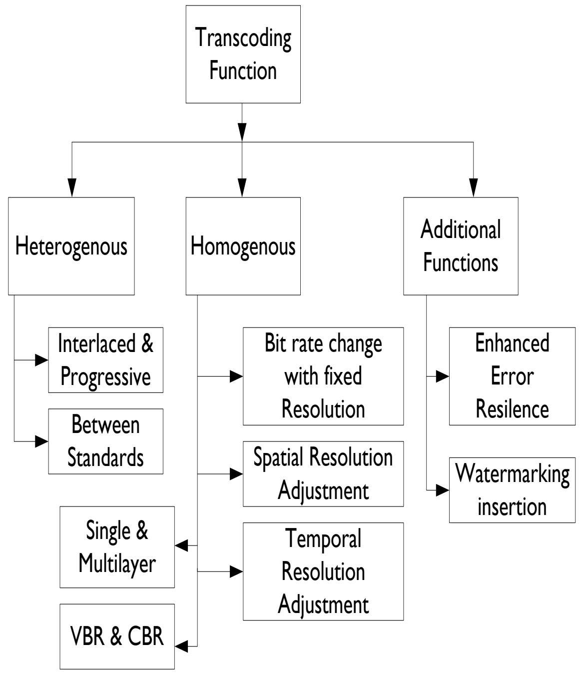

Figure 2 Different Transcoding techniques

5. PROPOSED SYSTEM

The proposed work is based on the fact that for low bandwidth mobile devices, transcoding can reduce download time considerably. The proposed system defines a technique for transcoding image and video to optimal file size that are considerably lower compared to the original file size. The proposed system introduces a novel framework that takes the input of multimedia file (image and video) and performs transcoding considering various attributes e.g. decoding time, encoding time, down-sampling, and DCT time. The proposed system also introduces a methodology termed as Transcoding Core Evaluation (TCE) that effectively enhances the performance of proposed multimedia transcoding. It has been observed that zeros play a key role in transform coding, especially at low bit rates. All typical coding algorithms treat zeros in a special way and address most of the effort to efficient coding for them. In transform coding of images and videos, the two most important factors are the coding bit rate and picture quality. After the DCT coefficients are quantized with a quantization parameter, the evaluation process estimates the percentage of zeros among the quantized coefficients. Hence the proposed system will study the rate and distortion score for the multimedia contents while transcoding.

6. Implementation and Result

The proposed system is designed on Windows 32-bit OS with 1.84 GHz processor with broadband connectivity of 100 Mpbs. The framework is designed using Matlab. The proposed framework is experimented over three types of image formats e.g. JPEG, BMP, and GIF with different size and dimensional factor of the image. The video clip of 3 min and 13 second of mpg-format is selected with size of 32 MB, frame width of 352, frame height of 288, and 25 frames per second. In case of video, it is converted to frames per second and each frame is subjected to cumulative proposed transcoding policy. Each images or frames are then subjected to JPEG compression followed by Transcoding Core Evaluation. Performing till this phase results in yielding following information in the framework:

- Original size
- Client bandwidth
- Server bandwidth
- Download Time Transcoding Time
- Quality value
- Parameter of saving in download time
- Image format
- Transcoded Size
- PSNR

The details of the image and video inputs are as follows:

Table 1 Image Input Considered

Image	Size(KB)	Dimension	Format
Image-1 	21.8	194x259	GIF
Image-2 	20.9	255x197	GIF
Image-3 	21.7	275x183	GIF
Image-4 	7.68	225x225	JPEG
Image-5 	8.56	224x225	JPEG
Image-6	10.1	204x204	JPEG

			
Image-7 	147	259x194	BMP
Image-8 	148	292x173	BMP
Image-9 	148	278x182	BMP

Table 2 Video Input Considered

Size	32 MB
Length	00:03:13
Frame width	352
Frame height	288
Frame rate	25 frames per second
Data rate	1150 kbps
Total bitrate	1374 kbps

The proposed system considers the quality of the compressed image or frame that can vary within the range of 0 to 100. Hence, the proposed system is computationally simple and reliable to estimate the necessary value of quality attribute for accomplishing the targeted size of compressed image.

Figure 3 highlights the subsampling time in seconds with respect to the image size in terms of bytes. The above results represents the time needed to downsample the input image (or frame) and obtain the coefficients of DCT. However, it can be seen that the time required for performing downsampling is pretty small and therefore this time is superimposed by the time needed for estimation of DCT. Hence, the results shows that this is linearly dependent on the number of the image (or frame) pixels. The results depicted in Fig 4 highlights basically image quality with respect to rate in bytes/pixels. The result is expressed in curve as estimate rate R and actual rate Rm. An error is introduced by neglecting the fact that the quantization step size is limited to 255. However, the error is small enough for the purposes

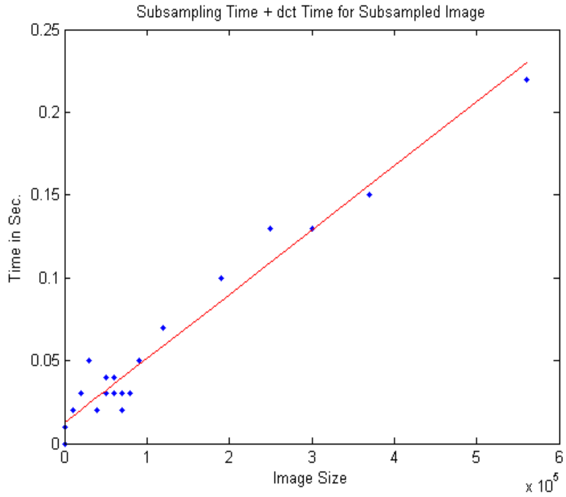

Figure 3 Subsampling time in seconds Vs Image size in bytes

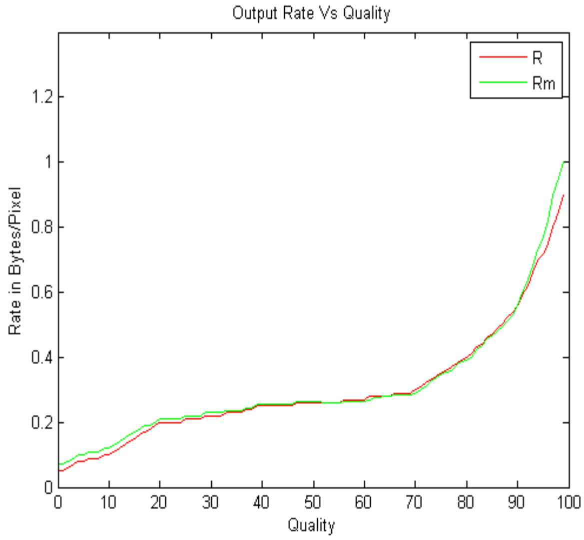

Figure 4 Actual rate and estimated rate Vs. quality for an image

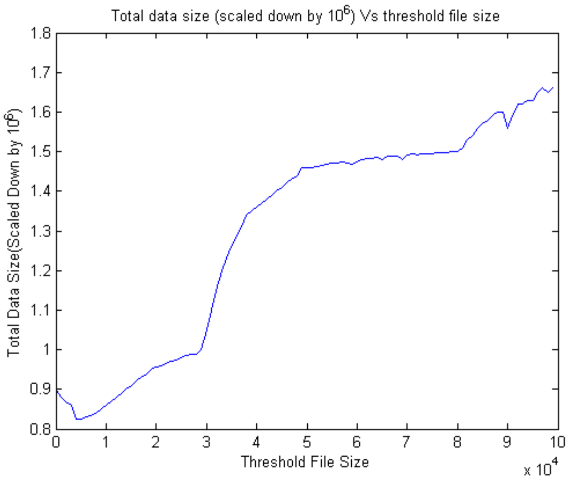

Figure 5 Total size of images Vs. Threshold for conversion to JPEG

The compression for GIF images can be maximized by either eliminating color tables or by downsampling. The proposed system doesn't use color table elimination as it results into loss of image details. As majority of the GIF images available on web are already of small size, therefore, it is preferable not to downsample by factor of more than 2. Other approach for GIF image compression is conversion to JPEG. For the data used, total size of GIF images was 1.1138 Mega bytes. If all are converted to JPEG, the data size reduces to 0.8881 Mega bytes. Also for most of the GIF images available on web, image details are not lost by converting to JPEG. But images are blurred depending upon quality factor used in JPEG. This little blurring is acceptable compared to other options –down sampling or color table truncation. Thus for GIF images, decision needs to be taken whether to convert to JPEG or to keep in GIF format. Ideally all images having graphics should be kept in GIF format and all photographs should be converted to JPEG, as photographs are better compressed by JPEG. But from the data obtained from web, it is observed that, as file sizes of GIF images increase, images become more complex. It is found that most of the large GIF images are either photographs or have embedded photographs in graphics. Hence it was found that large files should be converted to JPEG and smaller files should be kept in GIF format. To decide the threshold file size, threshold file size was varied from 0 to 10000 in steps of 100 bytes. All files above this threshold were converted to JPEG and files below threshold were kept in GIF format. Total data size was then measured. This data size versus threshold file size is plotted in figure 5.

7. COMPARATIVE EVALUATION

For the purpose of comparative evaluation, we chose to select the work of Stephane Coulombe [19], Karun B. Shimoga [20], and Johan Garcia [21], which discusses about the process of image transcoding by their respective techniques.

Stephane Coulombe [19] has proposed a novel low-complexity JPEG transcoding system which delivers near-optimal quality. The system is capable of predicting the best combination of quality factor and scaling parameters for a wide range of device constraints and viewing conditions.

Karun B. Shimoga [20] present an algorithm for adapting video image size to multiple heterogeneous clients while maximizing information content in the transcoded image. The work has considered a 512x512 gray scale image was considered for transcoding for five types of client displays:. workstation (256x256), desktop (192x192), TV-browser (128x128), hand-held (96x96), and personal digital assistant (PDA) (64x64). However, according to author, the results exhibited that region-of-interest based cropping, with maximized visual attention value, is the most natural method for cropping an image for transcoding for display on heterogeneous clients.

Johan Garcia [21] has presented progressive parsing, which is an efficient method for performing the transcoding of progressive JPEG images by truncating the data-stream. The suggested transcoding method has a lower delay and shows image quality advantages over other suggested approaches.

The transcoder produces images that adhere to the JPEG standard, but the transcoding is simpler and faster and can provide a better rate/distortion performance than other suggested transcoding approaches. All the above 3 works presents an optimal results of image transcoding that motivates us to adopt and compare with our technique too.

The results of comparative evaluation are performed by conducting a series of experiments and following the visual observation on average download time and rate in bytes/pixels being made in course of experiment as shown below:

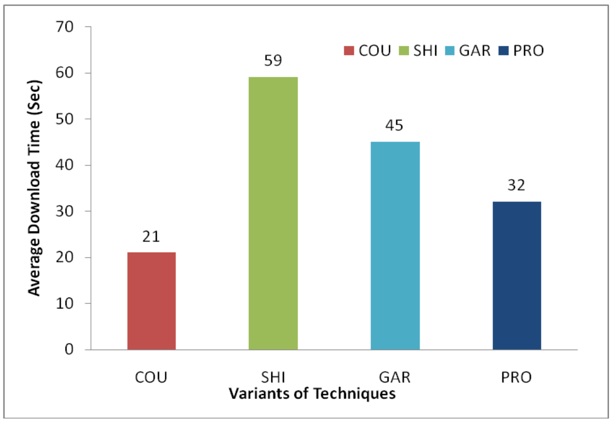

Figure 6 Average Download Time of different approaches

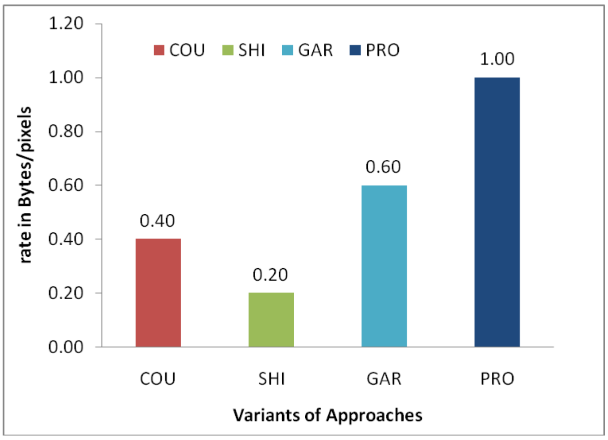

Figure 7 Rate in bytes/pixels of different approaches.

Figure 6 highlights the average download time in second considering all the 9 images in Table 1 for Stephane Coulomb [COU], Karun B. Shimoga [SHI], Johan Garcia [GAR], and the proposed system [PRO]. It can be seen that download time of COU is 21 sec due to the usage of scaling parameters. Moreover the work of [COU] has used better complexity analysis as compared to [SHI] and [GAR]. However, [COU] also exhibits failing rate in transcoding due to consideration of parameters that are limited to scaling factors of 10% or more, which is not in our case PRO. Hence PRO exhibits better download time in every aspects especially when it comes to mobile communication channel. Figure 7 highlights the rate in bytes/pixels of different approaches to mark the quality. In this case also, the proposed system out performs the other variants of techniques due to adaption of optimized DCT time for performing sub-sampling of images. COU address the optimal quality of images but posses failure rates when it comes to transcoding multiple images in wireless channel. SHI uses region of interest with maximized visual attention value but again lacks the consideration of wireless channel. However, GAR uses JPEG encoding that splits each color component into 8x8 sample point block followed by DCT to obtain better quality, but lacks optimization.

8. CONCLUSION

The proposed system introduces a novel framework where multimedia transcoding is applied on contents mainly targeting at mobile devices. The result accomplished proved that downloading time using multimedia transcoding is comparatively less than the process without transcoding. The result also ensure saving of download time when the bandwidth is maximized. Hence, the proposed system is highly adaptable to mobile devices with java environment which can efficiently be used for multimedia content sharing. One of the major limitation of the proposed work is that the design is done considering without security over transcoded multimedia files in wireless environment. Wireless environment is much prone for various attack like Denial of Service, routing attack, and Wormhole attack, the existing framework should be accompanied with efficient security work which will be addressed in our next chain of work

9. REFERENCE

[1] Mahoney, M. 2012. Data Compression Explained, Dell Inc, 2012

[2] http://www.jpeg.org/committee.html. Accessed on 5th Jan, 2013

[3] http://www.jpeg.org/jpeg2000/. Accessed on 5th Jan, 2013

[4] ITU-T and ISO/IEC JTC 1, "Generic coding of moving pictures and associated audio information – Part 2: Video," ITU-T Recommendation H.262 – ISO/IEC 13818-2 (MPEG-2), Nov. 1994.

[5] http://www.itu.int/rec/T-REC-H.263/en. Accessed on 5th Jan, 2013

[6] ISO/IEC JTC1, "Coding of audio-visual objects – Part 2: Visual," ISO/IEC 14496-2 (MPEG-4 visual version 1), April 1999; Amendment 1 (version 2), February, 2000; Amendment 4 (streaming profile), January, 2001.

[7] Wallace, G. K. 1991. The JPEG still picture compression standard. Communication of ACM, vol. 34, pp. 30–44..

[8] LeGall, D. 1991. MPEG: A video compression standard for multimedia application. Communication of ACM, vol. 34, pp. 46–58.

[9] Gish, H. and Pierce, J. N. 1968. Asymptotically efficient quantizing. IEEE. Trans. Inform. Theory, vol. IT-14, pp. 676–683..

[10] Berger, T. 1984. Rate Distortion Theory. Englewood Cliffs, NJ: Prentice- Hall.

[11] Corbera, J.R. and Lei, S. 1999. Rate control in DCT video coding for low-delay communications. IEEE Trans. Circuits Syst.Video Technol., vol. 9, pp. 172–185.

[12] Chiang, T. and Zhang, Y. -Q. 1997. A new rate control scheme using quadratic rate distortion model. IEEE Trans. Circuits Syst.Video Technol., vol. 7, pp. 246–250.

[13] Yang, J., Dai, Q., Xu, W., Ding, R. 2005. A Rate Control Algorithm for MPEG-2 to H.264 Real-time Transcoding. Visual Communications and Image Processing.

[14] Yasmin, T., Chaudhury, S., Jain, R. 2004. Transcoding of Document Images for Mobile Devices. Proceedings of the Fourth Indian Conference on Computer Vision, Graphics & Image Processing, Kolkata.

[15] Zhang, J., Perkis, A., Georganas, N.D. 2004. H.264/AVC and Transcoding for Multimedia Adaptation, In Proceedings of the 6th COST 276 workshop Thessaloniki, Greece

[16] Brandt, J. and Wolf, L. 2007. Multidimensional Transcoding for Adaptive Video Streaming, NOSSDAV.

[17] Vetro, A., Cai, J., Chen, C.W. 2002. Rate-reduction transcoding design for wireless video streaming, DOI:10.1109/ICIP (In proceeding of: Image Processing).

[18] Chandra, S., Gehani, A., Ellis, C.S., Vahdat, A. 2001 Transcoding Characteristics of Web Images, In ACM/SPIE Multimedia Computing and Networking.

[19] Coulombe, S., Pigeon, S. 2010. Low-Complexity Transcoding of JPEG Images with Near-Optimal Quality Using a Predictive Quality Factor and Scaling Parameters, IEEE

[20] Shimoga, K.B. 2002. Region-of-Interest Based Video Image Transcoding for Heterogeneous Client Displays, Proc. Packet *Video* Workshop

[21] Garcia, J., Brunstrom, A. 2000. Progressive Parsing Transcoding of JPEG Images, 7th Int. Workshop on Mobile Multimedia Communications (MoMuC2000), Tokyo, Japan