Content and Resources based Mobile and Wireless Video Transcoding

Ashraf M. A. Ahmad

Abstract**—**Delivering streaming video over wireless is an important component of many interactive multimedia applications running on personal wireless handset devices. Such personal devices have to be inexpensive, compact, and lightweight. But wireless channels have a high channel bit error rate and limited bandwidth. Delay variation of packets due to network congestion and the high bit error rate greatly degrades the quality of video at the handheld device. Therefore, mobile access to multimedia contents requires video transcoding functionality at the edge of the mobile network for interworking with heterogeneous networks and services. Therefore, to guarantee quality of service (QoS) delivered to the mobile user, a robust and efficient transcoding scheme should be deployed in mobile multimedia transporting network. Hence, this paper examines the challenges and limitations that the video transcoding schemes in mobile multimedia transporting network face. Then handheld resources, network conditions and content based mobile and wireless video transcoding is proposed to provide high QoS applications. Exceptional performance is demonstrated in the experiment results. These experiments were designed to verify and prove the robustness of the proposed approach. Extensive experiments have been conducted, and the results of various video clips with different bit rate and frame rate have been provided.

Keywords**—**Content, Object detection, Transcoding, Texture, Temporal, Video.

I. Introduction and Related Work

CURRENT advances in mobile communications and portable client devices enable us to access multimedia content universally. However, when multimedia content becomes richer, i.e., including video and audio, it is difficult for wireless access to communicate due to the existence of many restrictions. The most important factor of all, wireless connections usually have a much lower bandwidth compared to wired ones. And communication conditions change dynamically due to the effect of fading in wireless communication. Another important factor is that portable client devices are equipped only with limited computing and display capabilities, which are not suitable for high quality video decoding and displaying.

Concerning the heterogeneity issue, the previous era has seen a variety of developments in the area of multimedia representation and communication. In particular, we are beginning to see delivery of all types of multimedia data for all types of users in all types of conditions. In a diverse and heterogeneous world, the delivery path for multimedia content to a multimedia terminal is not straightforward, especially in the mobile communication environment. Access networks are various in nature, sometimes limited, and differ in performance. The characteristics of end user devices vary increasingly (over the recent years), in terms of storage, processing capabilities, and display qualities; also, the natural environment changes: changes in position, elucidation or temperature; finally, users are different by nature, appreciating dissimilar preferences, special usage, disabilities, etc.

The advance of multimedia systems has had a major influence in the area of image and video coding. The problem of interactivity and integration of video data with computer, cellular, and television systems is relatively new and subject to a great deal of research world wide. As the number of networks, types of devices, and content representation formats increase, interoperability between different systems and different networks is becoming more important. Thus, devices such as gateways, multipoint control units, and servers must be developed to provide a seamless interaction between content creation and usage.

The transporting of multimedia over wireless channels to mobile users is becoming a research topic of rapidly growing interest [10, 12, 13, 16, 17]. With the emergence of small wireless handset devices such as PDAs, and video mobile and so forth, it is expected that interactive multimedia will be a major source of traffic to these handset devices. These devices could be carried by users inside buildings when they are connected by wireless LAN or in vehicles when they will be connected to the cellular network, such as GPRS [1, 8]. As well as, Wideband mobile communication systems such as IMT-2000 have emerged, and there should be a mechanism to cope with a variety of media such as video provided to mobile terminal.

Wireless transmissions use radio channel as the transmission media. Generally, radio links connect users to base stations. The base stations in turn, are connected to routers using wired links. The wireless segment "cells" provides mobility to a user while using the network. In contrast to wireline transmission links wherein the bandwidth can be easily increased and the channel quality can be guaranteed, whereas the bandwidth of a wireless channel is limited because of spectrum allocation and physical limitations. The transmission quality of radio is easily affected by environments such as buildings, moving objects, and

Manuscript received August 21, 2006.
Ashraf M. A. Ahmad is with the Department Computer Science, Prinsess Sumya University for technology, Amman, Jordan (e-mail: ashraf@csie.nctu.edu.tw).

atmosphere, shielded obstacles and so forth. Moreover, because of the mobile nature of the users, the access point of a mobile user changes continuously. All these factors in wireless networks give rise to issues such as effective bandwidth allocation, channel with high bit error rate, and user handover.

Moreover, because of its high traffic characteristics such as high bit rate, video will be the dominant traffic in multimedia streams and hence needs to be managed efficiently. Obviously for efficient utilization of network resources, video must be compressed to reduce its bandwidth requirement. Although there exist several compression techniques, MPEG [10, 13, 16] is one of the most widely used compression algorithms for networked video applications. A wireless handset device, for instance personal data assistant, can integrate voice, video, and data in one device. In contrast to solely text information, multimedia data can tolerate a certain level of error and fading. That is why a wireless network has a high bit error rate compared to a wireline network. And so it is possible that it will cost too much to transmit multimedia over wireless networks with acceptable quality.

For instance, The MPEG-2 compressed digital video content is being used in a number of products including the DVDs, camcorders, digital TV, and HDTV. In addition, tons of MPEG2 data have been stored already in different accessible multimedia servers. The ability to access this widely available MPEG-2 content on low-power end-user devices such as PDAs and mobile phones depends on effective techniques for transcoding the MPEG-2 content to a more appropriate, low bit-rate video.

Therefore mobile access to multimedia contents requires video transcoding functionality at the edge of the mobile network for interworking with heterogeneous networks and services, changing bit rates and so forth. This transcoding mechanism should tickle the aforementioned issues in transmitting video in mobile and wireless network.

This paper is organized as the following. First a general description of wireless and mobile technologies and architectures is presented. Then, Mobile Multimedia and issues of viewing a video stream in a mobile environment have been explained. In addition, the functions of transcoding techniques and transcoding architectures have been drawn and expressed. Then, new mobile and wireless video transcoding system architecture has been presented. Finally future works and conclusion has been anticipated.

II. Handheld and Wireless and Mobile Technologies

A. Handheld Features

The handheld device which is involved in mobile communication is supposed to have some specific features:

1) Inexpensive.
2) Compact.
3) Lightweight.
4) Power consumption.
5) Limited display capabilities.

B. Wireless Wide Area Networks

After the first-generation analog mobile systems, the second-generation (2G) mobile digital systems were introduced offering higher capacity and lower costs for network operators, while for the users, they offered short messages and low-rate data services added to speech services. An important evolution of the 2G systems, sometimes known as 2.5G, is the ability to use packet-switched solution in GPRS (General Packet Radio System). The main investment for the operators lies in the new packet-switched core network, while the extensions in the radio access network mainly is software upgrades.

C. Wireless and Mobile Network Features

In general, wireless and mobile network have some limitations.

These limitations include:

1) Limited bandwidth.
2) A high channel bit error rate
3) Wireless connections much lower bandwidth compared to wired ones
4) Communication conditions change dynamically due to the effect of fading

To address the technical solutions, several mechanism and technologies have been deployed such as Wideband CDMA, CDMA 2000, TD-SCDMA (Time Division-Synchronous CDMA), DECT (Digital Enhanced Cordless Telecommunications), UWC-136.

III. Issues of Viewing a Video Stream in a Mobile Enviroment

Although the constraints imposed by the heterogeneous nature of the communication network are quite different from those arising from the diversity of user terminals, both of them may deal with transcoding systems. In this work we address the problem of MPEG stream video transcoding, where the bandwidth of a coded video stream must be drastically reduced in order to cope with a highly constrained transmission channel, such as mobile. The variety in mobile devices also increases the difficulty of accessing content. For example, mobile devices are conveniently sized to fit in a pocket, but this size constrains their display area. Creating arbitrary trimmed versions of content could get around this constraint, but differences in display capabilities would easily make a device-specific authoring approach too costly to be practical and lower quality of service in some cases. Examples of device differences include screen sizes ranging from few hundred to thousands of pixels, and color depths ranging from two line black-and-white display to full-color display. Vide content can also be encoded in many different modes, such as MPEG-X series, and H26X series and so forth.

In these cases, transcoders can be used to transform multimedia content to an appropriate video format and bandwidth for wireless mobile streaming media systems. A conceptually simple and straightforward method to perform this transcoding is to decode the original video stream, downsample the decoded frames to a smaller size, and re-encode the downsampled frames at a lower bitrate. However, a typical CCIR601 MPEG-2 video requires almost all the cycles of a 300Mhz CPU to perform real-time decoding. Encoding is significantly more complex and usually cannot be accomplished in real time without the help of dedicated hardware or a high-end PC. These factors render the conceptually simple and straightforward transcoding method impractical. Furthermore, this simple approach can lead to significant loss in video quality. In addition, if transcoding is provided as a network service in the path between the content provider and content consumer, it is highly desirable for the transcoding unit to handle as many concurrent sessions as possible. This scalability is critical to enable wireless networks to handle user requests that may be very intense at high load times. Therefore, it is very important to develop fast algorithms to reduce the compute and memory loads for transcoding sessions.

In this section we identify issues involved in viewing a video stream in a mobile computing environment and propose methods of handling these issues. Video streams encoded by a high-bit-rate compression method (e.g., MPEG-1) require several megabits per second of bandwidth and are not suitable for a strictly band-limited environment. Therefore, video streams should be encoded by a low-bit-rate compression method (e.g., MPEG-4). MPEG-4 is appropriate for mobile access because it is robust to channel errors. From this point of view, some MPEG-4 codec large-scale integrations for mobile devices are under development, and MPEG-4 is apparently becoming the mainstream technique for mobile video usage. When communication conditions get worse and error rate increases in a wireless link, transmission jitter increases because error packets are retransmitted based on Radio Link Control protocol, located at the data link layer. If any packets are not recovered by retransmission, using a reliable transport protocol such as Transmission Control Protocol (TCP), error packets are retransmitted end to end. Consequently, this causes more transmission jitter and throughput reduction. Also, if an unreliable transport protocol such as User Datagram Protocol (UDP) is used, packets not recovered by retransmission and flooded by congestion on the communication link will be discarded, so the rate of packets that arrive at the client decreases. As a result, in both cases the application layer throughput reduces, and video data cannot be transferred stably. Therefore, the data rate of a video stream needs to be adapted in accordance with communication conditions. In addition, jitter can be absorbed by preserving some video data on the client buffer. As far as client device capabilities are concerned, display size, processing power, and memory size have to be taken into account. Since the display size of portable devices is small (on cellular phones both width and height are about 100 pixels at most), a video stream may be larger than the display size and is not easily viewed by mobile users. Hence, both the width and height of a video frame have to be fitted to the display size. In another method, the size may be reduced on the user client. However, it places an extra load on the user client and is undesirable for mobile where the processing power of the client device is low. If the client device does not have enough real-time video stream decoding power, a mobile user cannot fully view video streams. The amount of processing required for decoding is related to the number of video frames in one second (frame rate) and total number of pixels in one frame (frame size). Therefore, the frame rate and frame size need to be adjusted according to the processing power of the client device. As for memory size, the client device needs to have sufficient memory to preserve several decoded frames because, in encoding methods such as MPEG-4, the decoding process requires a few frames before and after the frame which is actually being decoded.

IV. Functions of Transcoding Techniques

To overcome the aforementioned limitations and obstacles in viewing video steams in wireless and mobile network, an effective transcoding mechanism is required. Building a good video transcoding for mobile devices poses many challenges. To overcome these challenges, a various kind of transcoding function is provided. The following paragraphs will describe these functions in details.

First function is bit rate adaptation. Bit rate adaptation has been the most significant function of video transcoding techniques. The idea of compressed video bit rate adaptation is initiated by the applications of transmitting pre-encoded video streams over heterogeneous networks. When connecting two transmission media, the channel capacities of the outgoing channel may be less than those of the incoming channel, so bit rate adaptation is necessary before sending the video bit stream over heterogeneous channels. In applications, such as video on demand, where video is off-line encoded for later transmission, the channel characteristics through which the resulting bit stream will be transmitted might be unknown. Through video transcoding, the bit rate of pre-encoded videos can be dynamically adapted to the obtainable bandwidth and variable communication circumstances. In most bit adaptation cases, a pre-encoded video with high bit rate and fine visual quality will be converted into low bit rate video with elegantly degraded visual quality.

Second function is frame size conversion. Video spatial resolution downscaling is significant since most current handheld devices are characterized by limited screen sizes. By inserting a downscaling filter in the transcoder, the resolution of the incoming video can be reduced. Because of downscaling the video into lower spatial resolution, motion vectors from the incoming video cannot be reused directly, but have to be resampled and downscaled. Based on the updated motion vectors, predictive residues are recalculated and compressed.

Third function is frame rate conversion. To transcode an arriving compressed video bitstream for a low bandwidth outgoing channel, such as a wireless network, a high transcoding percentage is often necessary. However, high transcoding ratios may result in intolerable video quality when the arriving bitstream is transcoded with the full frame rate as the arriving bitstream. Frame-rate conversion or frame-dropping is often used as an efficient scheme to assign more bits to the remaining frames, so that acceptable quality can be maintained for each frame. In addition, frame-rate conversion is also needed when an end system can play video only at lower frame rate due to the processing power limit. Frame rate conversion can be simply accomplished by random frame dropping. For instance, dropping every other frame in a sequential order leads to a half rate reduction in the transcoded sequence. When frames are dropped, motion vectors from the arriving video cannot be directly reused because they are pointed to the immediately previous frame. If the previous frame is dropped in the transcoder, the link between two frames is broken and the end decoder will not be able to reconstruct the picture by these motion vectors. Therefore, the transcoder is in charge for calculating new motion vectors that point to the previous un-dropped frames.

V. Transcoding Architectures

Generally speaking, transcoding can be defined as the manipulation or conversion of data into another more desirable format. Depending on the particular strategy that is adopted, the transcoder attempts to satisfy network conditions or user requirements in various ways. In the context of video transmission, compression standards are needed to reduce the amount of bandwidth that is required by the network. Since the delivery system must accommodate various transmission and load constraints, it is sometimes necessary to further convert the already compressed bitstream before transmission.

The simplest way to develop a video transcoder is by directly cascading a source video decoder with a destination video encoder, which is called the cascaded pixel domain transcoder [11]. Without using common information, this direct approach needs to fully decode input video and re-encode the decoded video by an encoder with different characteristics as described in fig.1. Obviously, this direct approach is usually intensive in computation. The architecture is flexible, because the compressed video is first decoded into raw pixels, hence a lot of operations can be performed on the decoded video. However, as we mentioned earlier, the direct implementation of the Cascaded Pixel Domain Transcoder is not desirable because it requires high complexity of implementation.

The alternative architecture for transcoding is an open-loop transcoding in which the incoming bitrate is downscaled by modifying the discrete cosine transform (DCT) coefficients. For example, the DCT coefficients can be truncated, requantized, or partially discarded in the optimal sense [9], [3] to achieve the desirable lower bitrate. In the open-loop transcoding, because the transcoding is carried out in the coded domain where complete decoding and re-encoding are not required, it is possible to construct a simple and fast transcoder. However, open-loop transcoding can produce “drift” degradations due to the mismatched reconstructed pictures in the front-encoder and the end-decoder, which often result in an unacceptable video quality. As a trade of between the aforementioned approaches limitation we [4] proposed adaptive motion vector refinement based approach for video transcoding mechanism. In this paper we incorporate this approach to construct our proposed system.

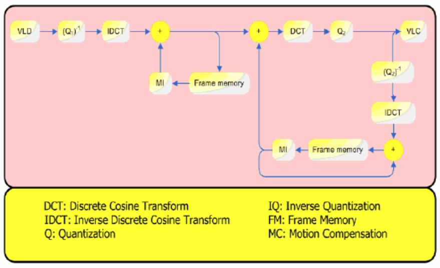

Fig. 1 Cascaded pixel domain transcoder

VI. Mobile and Wireless Video Transcoding System Architecture

In order to state the mechanism to deploy a transcoding scheme for high performance transcoding in mobile or wireless communication, we propose general system architecture for mobile and wireless video transcoding. In order to enable users to access video through wireless network and handheld devices, we propose using a transcoder as an intermediate node to dynamically convert the video according to network connections and user devices. In this approach, a content provider provides only one video stream, which is pre-encoded at high quality. The video transcoding performs transcoding dynamically for each user and provides converted video streams. The architecture of the video transcoding system is illustrated in Fig.2, as the proposed system contains four important parts which are user profile, video transcoder, control module and pre-encoded video content. In this system, the transcoder is integrated into the video streaming server. It can also be placed as an intermediate node along the transmission path.

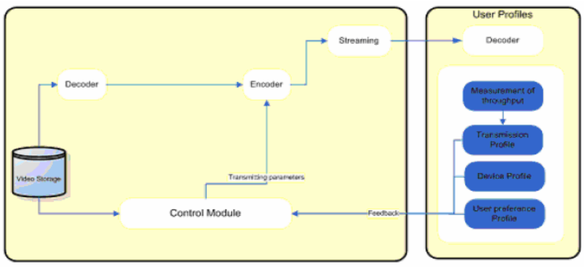

Fig. 2 Overview of Mobile and Wireless Resources, Network Conditions, and Content based Video Transcoding System Architecture

VII. Dynamic Resource Manager

The resource manager maintains several profiles. The handheld device profile includes hardware and software information on the devices, such as display size, processing power, storage capability and decoder information. The user preference profile includes user preference information such as user-preferred display size, user-preferred video presentation and user-preferred video player behavior. The communication profile will measure the download throughput and update the communication conditions. The information included in the device profile and user profile will be transmitted to the video stream server before the start of the video transmission session. The information included in the transmission profile will be sent to the stream server periodically to control the bit rate adaptation in the transcoder. All information in the user profile is used for parameters in the transcoding process. Table I is a clear example for user profile information.

TABLE I
User Profile Information

Information [a]
Frame rate
Pixel size
Monitor resolution
Graphic engine capability
CPU Usage
Memory usage
Visual objects summary
Visual combination report
Audio object summary
Scene description level

VIII. Video Transcoder

The video transcoder is the actual conversion engine of a video stream. It decodes a video stream, which is pre-encoded at high quality and stored in the video source, and then performs transcoding according to our proposed scheme. According to result we present in section 6, our proposed scheme has a very high performance in terms of visual quality. They are comparable to results which can be achieved by full-scale motion estimation based transcoding. When fast transcoding architectures are used, it is possible to execute transcoding in real time. Thus we can provide the handheld device user a smooth, online video presentation.

IX. Transcoding Control

The control module is responsible for creating a transcoding scheme according to the user profile and other information. The transcoding scheme will include some transcoding parameters. In order to decide appropriate transcoding parameters, decisions must be made by considering all of the factors adaptively. For example, when connection throughput is low, the bit rate of the video needs to be converted. At the same time, in order to ensure video quality, the frame rate of the video also needs to be reduced. In so doing, each frame will have enough bit budgets to maintain tolerable visual quality.

X. Conclusion

In this paper, several typical video transcoding architectures and major applications of video transcoding have been reviewed. We identify issues involved in access video streams through handheld devices and wireless networks. We state that the main functions of this transcoding include frame size downscaling, frame rate conversion, bit rate adaptation, color conversion, etc. To handle these issues, few video transcoding system architectures have been proposed in particular we introduce a system architecture recently suggested by us for intelligently transcoding pre-encoded video for different user devices and network connections in wireless network environment. Currently, a good trend towards the mobile TV has been brought to the academia and industries. Therefore, video transcoding will be very interesting issue to be addressed for mobile TV field.

References

[1] A. Eleftheriadis and D. Anastassiou, "Constrained and general dynamic rate shaping of compressed digital video," in Proc. IEEE Int. Conf. Image Processing, Washington, DC, Oct. 1995.
[2] A. Ghosh, et al., "Broadband wireless access with WiMax=8O2.16: current performance benchmarks and future potential", IEEE Commun.Mag., 2005, 43, (2), p. 129136
[3] A. Eleftheriadis and D. Anastassiou, "Constrained and general dynamic rate shaping of compressed digital video," in Proc. IEEE Int. Conf. Image Processing, Washington, DC, Oct. 1995
[4] A. M. Ahmad and S-Y. Lee, "A Novel Approach for Improving the Quality of Service for Wireless Video Transcoding" to be appeared In the International Journal of Mobile Multimedia
[5] A. Dornan, The Essential Guide to Wireless Communications Applications, From Cellular Systems to WAP and M-Commerce, Prentice Hall PTR, 2000
[6] A. Ephremides, T. Itoh, R. Pickholtz, M. Iskander, L. Katehi, R. Rao, W. Stark, J. Winters, "Wireless Technologies and Information Networks", International Technology Research Institute, World Technology (WTEC) Division, WTEC Panel Report, July 2000
[7] C.F. Chiasserini, M.A. Marsan, E. Baralis, P. Garza, Towards feasible topology formation algorithms for Bluetooth-based WPAN's", System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on 6-9 Jan 2003 Page(s):10 pp.
[8] G. Keesman et al., "Transcoding of MPEG bitstreams," Signal Process. Image Comm., vol. 8, pp. 481-500, 1996.
[9] H. Sun, A. Vetro, J. Bao, and T. Poon, "A new approach for memory-efficient ATV decoding," IEEE Trans. Consumer Electron., vol. 43, pp.517-525, Aug. 1997.
[10] J. Mitchell, W. Pennebaker, E. Fogg, Chad, & LeGall J. Didier. (1996). MPEG Video: Compression Standard. (1st ed.). New York, Chapman and Hall. [4,3]
[11] J. Youn, and M.-T. Sun, "Video Transcoding with H.263 Bit-Streams," Journal of Visual Communication and Image Representation, 11, 2000.
[12] P. Correia, S. M. Faria, P. A. Assuncao, Matching MPEG-1/2 Coded Video to Mobile Applications, 4th International Symposium on Wireless Personal Multimedia Communications, Vol. 2, pp. 699-704, Aalborg - Denmark, September 2001.
[13] R. Han, P. Bhagwat, R. LaMaire, T. Mummert, V. Perret and J. Rubas, Dynamic Adaptation in an Image Transcoding Proxy for Mobile Web Browsing, IEEE Personal Communications, pp.8-17, December 1998.
[14] Revision of part 15 of the Commission's rules regarding ultra-wideband transmission systems, First Report and Order, Federal Communications

Commission (FCC), Washington, DC, ET Docket 98-153, FCC 02-48, 2002
[15] R. Giuliano, and Mazzenga, F.: "On the coexistence of power-controlled ultrawide-band systems with UMTS, GPS, DCS 1800, and fixed wireless systems", IEEE Trans. Veh. Technol., 2005, 54, (1), pp. 62-81
[16] T. Warabino, S. Ota, D. Morikawa, and M. Ohashi, Video Transcoding Proxy for 3Gwireless Mobile Internet Access, IEEE Communications Magazine, pp. 66-71, October 2000.
[17] T. Shanableh and M. Ghanbari, Heterogeneous video transcoding to lower spatio-temporal resolutions and different encoding formats, IEEE Transactions on Multimedia, Vol. 2, No 2, pp. 101-110, June 2000.
[18] S.J. Vaughan-Nichols, : "Achieving wireless broadband with Wi-Max", IEEE Computer., 2004, 37, (6), pp. 10-13.
[19] M. Hamalainen, et al.: "On the UWB system coexistence with GSM900, UMTS=WCDMA, and GPS", IEEE J. Sel. Areas Commun., 2002, 20, (9), pp. 1712-1721.
[20] UMTS Forum Report No. 9: "The UMTS Third Generation Market - Structuring the Service Revenue Opportunities"' October 2000.
[21] UMTS Forum Report No. 10: "Shaping the mobile multimedia future"' October 2000.

Ashraf Ahmad obtained his PhD degree in Computer Science and Engineering from National Chiao Tung University (NCTU) in Taiwan. He obtained his B.Sc. degree from Princess Sumya University for Technology (PSUT) in Jordan. Mr. Ahmad is currently an assistant professor at the department of computer since in PSUT, Jordan. His interest area includes multimedia semantic features extraction, and analysis, multimedia retrieval, and multimedia communication. Prof. Ahmad has authored over 30 scientific publications including journal papers, conference papers and book chapters. In addition, Mr. Ahmad has several US and international patents in his field of expertise. He serves in program committee for several international conferences. He is also a reviewer and referee for several conferences and journals. His work has been published and presented at various international conferences. Mr. Ahmad has been listed in Who's Who in the World for the year 2006 and Who's Who in Asia for the year 2007. In addition, Prof. Ahmad has been elected as one of the 2000 Outstanding Intellectuals of the 21st Century for the year 2006 for his outstanding contribution in field of Video Processing and Communications. Ashraf Ahmad has been chosen as one of the recipients of Leading Scientist award in the year 2006.