Excerpt from book "Multimedia Communications: Directions & Innovations", chapter 15, ed. Jerry D. Gibson, Academic Press, 2000.
For private use only.

Transcoding of the Internet's Multimedia Content For Universal Access

Richard Han and John R. Smith
{rhan, jrsmith}@watson.ibm.com

1. INTRODUCTION

Transcoding is a technique that transforms multimedia, e.g. text, images, audio and/or video, from the original format in which the multimedia was encoded into a second alternative format. Within the emerging world of wireless connectivity to the Web, transcoding of Internet multimedia enables Internet content providers, e.g. Web sites, and Internet service providers (ISP) to transform Web content so that a wider collection of Internet-enabled client devices, such as cell phones and personal digital assistants (PDA), can have access to that content. As a second equally important benefit, transcoding of Internet multimedia enables Web sites and ISP's to increase the perceived speed of access to that content by additional compression of text, image, audio, and/or video over slow Internet access links.

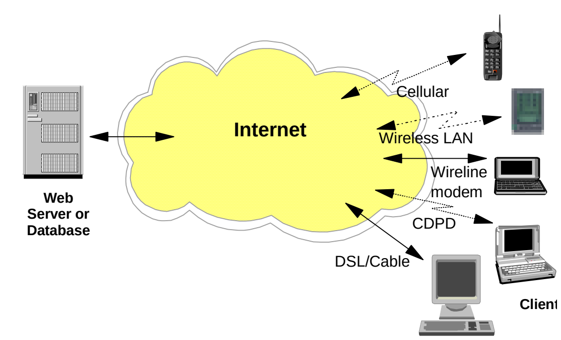

Figure 1. The emerging world of the wireless Web consists of a wide variety of fixed and mobile clients connected to the Internet over a wide range of link speeds, all trying to access multimedia content stored on a remote Web server or database.

As shown in Figure 1, the Internet is experiencing a rapid diversification in the types of network links used to connect to the Internet, e.g. 56K modems, cable modems, digital subscriber loops (DSL), cellular data links such as the European/Asian GSM standard, wireless data services such as Mobitex, wireless LAN's, etc. Transcoding enables re-compression of multimedia content, i.e. reduction in the content's byte size, so that the delay for downloading multimedia content can be reduced. Compression-based transcoding over links with the most limited bandwidth can dramatically improve the interactivity of Web access. An additional advantage of reducing the byte size of multimedia content is reducing the cost over tariffed access links that charge per kilobyte of data transferred.

At the same time, the Internet is experiencing a rapid proliferation in the kinds of client devices used to access the Internet, e.g. desktop workstations, laptops, Web-enabled screen phones, Internet-enabled cellular phones, integrated television/Web set-top boxes, handheld Palm PDA's with Web access, etc. Transcoding enables the transformation of multimedia content so that each of these increasingly diverse client devices can have access to the Internet's rich multimedia content, albeit in transcoded form.

1.1 Adapting To Bandwidth Heterogeneity

A common motivation for applying transcoding is to reduce the delay experienced when downloading multimedia-rich Web pages over Internet access links of limited bandwidth, e.g. modem links and/or wireless access links [Liljeberg95, Fox96a, Bickmore97, Fleming97, Han98, Smith98a, Bharadvaj98]. The transcoding function is typically placed within an HTTP proxy that resides between the content provider's Web server and the client Web browser. The transcoding proxy reduces the size in bytes of the multimedia content via lossy compression techniques (e.g. images are more aggressively compressed, HTML text is summarized). The proxy then sends the re-compressed multimedia content over the modem/wireless access link down to the client. The reduction in byte size over the access link, typically a bottleneck, enables an often significant reduction in the perceived response time. Transcoding has also been applied to continuous media to reduce a video's bit rate and thereby enable clients connected over bandwidth-constrained links to receive the modified video stream [Amir95, Yeadon96, Tudor97]. The terms distillation, summarization, dynamic rate shaping, and filtering are frequently used as synonyms for transcoding when reducing latency (via compression over low-bandwidth links) is the primary motivation for transcoding.

While transcoding reduces response time over very slow links, compression-based transcoding can actually increase the response time over high bandwidth links [Han98]. The operation of transcoding adds latency, especially for images, because decompression and compression are collectively compute-intensive. Over high bandwidth access links, the decrease in download time achieved by reducing the byte-size may be less than the increase in delay caused by transcoding (decompression and re-compression), thereby resulting in inflated response times. Figure 2 summarizes this behavior and suggests that transcoding should only be performed when the network poses a sufficient bottleneck. The precise conditions are derived in later sections.

In addition to compression, a transcoding proxy can also reduce response time by choosing an output transcoded format that minimizes client-side decoding delay. .... See textbook for further information .... Comparisons of the tradeoffs in response time obtained by moving decoding complexity away from the PDA's browser towards a transcoding proxy have been analyzed [Fox98b, Han99]. An early study found that partitioning most of a text-only Web browser's complexity away from a PDA into a proxy enabled response times on the order of several seconds (in conjunction with caching and prefetching) [Bartlett94]. General architectures have been proposed that partition application functionality [Watson94] or distribute code objects [Joseph95] between the network (a server or transcoding proxy) and a mobile client, especially as a means of dynamically adapting to changing network and client conditions [Nakajima97, Noble97]. The commercial Palm VII PDA conducts wireless access to the Internet via a partitioned application structure; streamlined "query" applications on the PDA request information from selected Web servers, whose content is then transcoded by an intermediate "Web clipping" proxy to a PDA-specific format [PalmVII99].

.... See textbook for further information ....

Compression-based transcoding includes both compression within a media type as well as summarization across media types. .... See textbook for further information ....

A convenient means of summarizing the multiple ways to transcode, i.e. conversion and re-compression of content both between media types as well as within a media type, is depicted in the "InfoPyramid" data model of Figure 3 [Mohan99]. .... See textbook for further information ....

While the type of transcoding described in this section uses re-compression to adapt to links with different bandwidth characteristics, it has also been proposed that transcoding be applied to adapt to links with different loss/error characteristics [Swann98, de los Reyes98]. In this form of transcoding, the encoding format of the multimedia is adapted by a transcoder into a format that is more error-resilient to packet losses and/or bit corruption within a packet's payload.

1.2 Adapting To Client Heterogeneity

A second common motivation for applying transcoding is to resolve mismatches between a client's decoding capabilities and a content provider's encoding formats, so that multimedia can be conveyed to a variety of clients. Codec mismatches are often caused by hardware limitations on mobile clients, software restrictions at the client and/or content provider, and constraints imposed by the end user and/or environment in which the mobile client is operated. To compensate for codec incompatibilities, transcoding is applied to convert the original format of multimedia into a form that can be viewed and/or played back by a client. Consequently, transcoding expands the number and types of fixed and mobile clients with access to existing content. The terms translation, content repurposing, content adaptation, reformatting, data transformation, media conversion, clipping, universal access, format conversion, and filtering are often used as synonyms for transcoding when it is applied to adapt to client heterogeneity.

Transcoding is currently being applied in a variety of contexts to adapt to different client/user requirements. Speech recognition and speech synthesis represent two traditional examples of transcoding technologies in which a conversion process is required in order to resolve codec mismatches and convey useful information to the client/user. They can be classified as speech-to-text transcoding and text-to-speech transcoding respectively. In addition, transcoding is being applied to transform Web pages (images and HTML text) so that handheld PDA's can browse the Web [Gessler95, Fox98a, Han98, Smith98a]. Since PDA clients are constrained by their small screen size (e.g. 160x160 screen), often limited display depth (e.g. 2 bits of grayscale per pixel), limited memory, relatively slow processor, and limited support of media types (e.g. audio and/or video may not be supported), then transcoding provides a way to overcome these constraints so that images and text can be properly displayed on these handheld devices. Color GIF and JPEG images have been transcoded and pre-scaled to 2-bit grayscale bitmaps by a proxy [Fox98b, Han99].

In the future, we envision an expansion of the role of transcoding to cope with the rising number of codec mismatches caused by an increasingly heterogeneous world of Internet access devices. First, increasingly diverse client-side hardware for Internet access devices will likely drive the need for transcoding. .... See textbook for further information ....

A second likely role for transcoding will be adaptation of multimedia to suit human user preferences or needs as well as to suit environmental restrictions. .... See textbook for further information ....

A third category in which we anticipate the development of more transcoding solutions is based on resolving software codec mismatches between the client's decoder and the content provider's encoder. Web servers and Web browsers will likely need to transition to incorporate new software standards, e.g. XML, MP3, PNG, JPEG2000, MPEG4, etc. .... See textbook for further information ....

2. END-TO-END VS. PROXY-BASED TRANSCODING DESIGNS

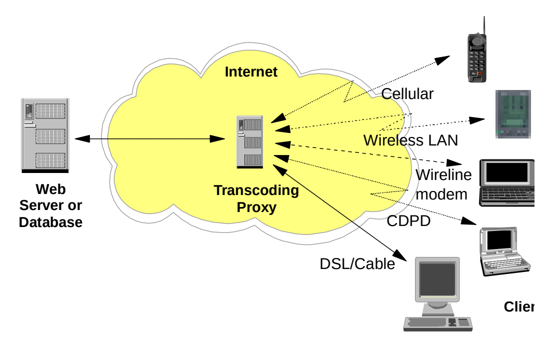

Figure 4. Proxy-based transcoding of multimedia content involves on-the-fly transformations of a server's content by an intermediate proxy for presentation to diverse clients connected to the Internet over diverse access links.

Multimedia content that needs to be adapted for presentation to diverse clients over diverse access links can be either transcoded on-the-fly by a proxy, or pre-transcoded off-line by a server. In on-the-fly or real-time transcoding, a request for content triggers the content provider to serve the original document, which is then transformed on-the-fly by a transcoding entity downstream to create a newly adapted version of the document. As illustrated in Figure 4, the real-time transcoding entity can be placed without loss of generality in a proxy that is located at some intermediate point between a server and a requesting client. The proxy is typically situated in the network on the opposite side of the access link as the client. In this way, compression-based transcoding can reduce download times over the often bottlenecked access link. For ISP's that wish to offer a Web transcoding service, the transcoding proxy can be placed within the ISP access cloud. For Web sites that desire to provide universal access to their content, the transcoding proxy can be placed adjacent to the Web server on the content provider's premises. Other terms for a transcoding proxy include mixer [Schulzrinne96], and gateway [Amir95, WAPForum99].

In contrast to the proxy-based approach, the end-to-end design philosophy for transcoding prepares multiple pre-transcoded versions of a multimedia document off-line, i.e. a server authors multiple versions of a multimedia document prior to receiving any requests for that document. Each version is specially tailored for a given client and/or access link speed. Consequently, no intermediate proxy is needed to transform the data. As illustrated in Figure 5, when a request for a multimedia document is received from a particular client over a given link speed, the content provider's server simply selects the appropriate pre-transcoded version to serve, thereby avoiding on-the-fly transformations. A simple example of server-side authoring is the support by certain Web sites for text-only or low-graphics versions of Web pages. .... See textbook for further information ....

Table 1 summarizes the tradeoffs associated with implementing either end-to-end server-side pre-authoring or proxy-based real-time transcoding. One of the primary motivations for proxy-based implementations of transcoding is that the existing infrastructure of servers and clients doesn't have to be modified in order to implement a transcoding service. The client browser merely needs to be redirected to access the Web through an HTTP transcoding proxy. Also, existing Web servers need not be changed. Ideally, any news or entertainment site on the Web could be accessed via transcoding technology without burdening content providers with the need to customize each site for display on multiple client devices. In contrast, end-to-end server-side authoring requires modifications to each Web server to generate multiple client-specific versions of a given Web page a priori.

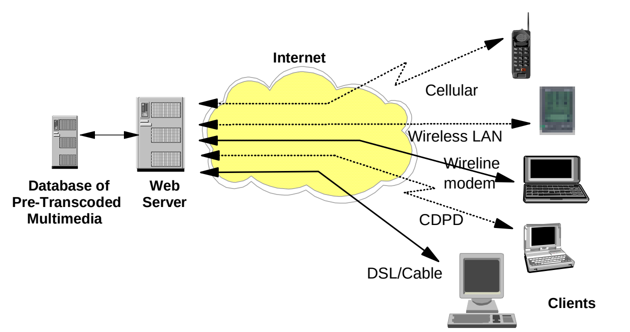

Figure 5. End-to-end transcoding generates multiple versions of a multimedia document prior to a client request. Each version is adapted for a specific client and/or access link. In response to a client request, the server returns the appropriate pre-transcoded version, thereby avoiding on-the-fly transcoding by a proxy.

While proxy-based implementations appear to offer an easy means to deploy transcoding into the existing Web server and Web client infrastructure, proxy-based solutions introduce their own problems. First, proxies often perform compute-intensive decompression and re-compression of images and/or video. The latency introduced by such transcoding is often non-negligible, e.g. on the order of a hundred milliseconds on an unloaded image transcoding proxy [Han98]. If the purpose of applying transcoding is to improve the response time over low-bandwidth links, then transcoding can under certain conditions worsen the overall response time. As we scale to many users, proxies may become quickly overloaded such that compute-intensive transcoding becomes completely impractical, i.e. the cost of maintaining a farm of proxies to meet transcoding obligations may become too high. However, several techniques have been offered to alleviate the burden of computation on transcoding proxies. Fast transcoding of video and images occurs in the compressed or frequency domain [Acharya98, Amir95, Assuncao98, Chong96]. .... See textbook for further information ....

Another drawback of proxy-based transcoding is security-related. If multimedia content is encrypted, then a proxy must be a trusted entity allowed to decrypt the content so that decompression and re-compression can be applied to the protected content. .... See textbook for further information ....

Another disadvantage of the proxy architecture is that the proxy has imperfect knowledge concerning the relative importance of various media objects within a multimedia Web page, which can lead to an imperfectly transcoded Web page. .... See textbook for further information ....

The process of decompression and lossy re-compression associated with transcoding proxies inherently accumulates quantization noise each time the cycle is invoked. This effect is most noticeable when there is a concatenation of transcoding proxies. Analog noise accumulates with each transcoding operation, so that the final image or audio suffers cumulative degradation even though individually each proxy can claim that the degradation it introduced was relatively minor [Chang94, Wilkinson98]. .... See textbook for further information ....

Transcoding proxies also lead to excessive information being sent over the backbone network from server to proxy. For example, the entire Web page is typically transmitted from server to proxy, only to have the proxy significantly filter the objects on that page via compression. A more efficient design would be to send only the images and text that are necessary for final reconstruction on the client, so as not to waste backbone bandwidth. This efficient approach is taken by server-side authoring, which sends images and text that have already been compressed to account for the destination client.

As shown in Figure 6, a chain of proxies will exacerbate many of the problems mentioned previously. .... See textbook for further information ....

Despite these numerous disadvantages, proxies have the advantage (in addition to ease of deployment) of offering transitional technology that allows new standards to be incrementally introduced and permits legacy standards to be indefinitely supported. When a server adds a new standard for images, audio, video, or text, clients need not immediately adopt the new standard in order to access the newly encoded content provided that there is a proxy that can translate the new format into a legacy format. It is unlikely that all Web sites that adopt a new standard, especially non-commercial ones, will retain full backward compatibility, i.e. retain multiple versions of an image (in JPEG2000 and GIF), or an audio stream (MP3 and WAV format). In the absence of a proxy, it would therefore be difficult to incrementally deploy new standards without isolating certain users with legacy decoding software.

Some additional tradeoffs in the design of general purpose proxy-based systems have been enumerated [Zenel97, Border99], including how proxies can break the semantics of TCP.

Table 1. End-to-end serving of pre-transcoded Web pages vs. proxy-based real-time transcoding of Web pages.

Properties	End-to-End Serving of Pre-Transcoded Content	Proxy-Based Real-Time Transcoding
Modifications to existing infrastructure	upgrade servers to generate multiple pre-transcoded versions of a Web page, server informed of client characteristics	no modification at server, client redirected to point to proxy, client needs to advertise its characteristics to proxy
Latency	little additional delay	compression reduces delay, but compute-intensive transcoding adds delay
Scalability in terms of processing	off-line transcoding can be performed when convenient	highly compute-intensive transcoding, often under real-time deadline pressure
Security	supports end-to-end encryption	proxy must be trusted to decrypt, decompress, re-compress, and re-encrypt
Semantic understanding	server knows semantic importance of each media object in a page/document	proxy has incomplete knowledge of page composer's intent, proxy can be aided by hints from server
Degradation	no additional degradation	noise/analog degradation accumulates with each decompression-lossy compression cycle
Efficiency	send appropriately sized images and text over backbone	excessive image/text/audio/video sent over backbone between server and proxy
Scalability in terms of upgrading to a new format	each *server* need to be upgraded to support N+1 encoders from N encoders	each *proxy* needs to be upgraded to add new encoder and decoder (i.e. 2(N+1) codecs)
Legacy support	backward compatibility may be limited	transitional role supporting legacy client hardware and compression standards, phased introduction of new standards
TCP semantics	unaffected	often broken by a proxy

3. ARCHITECTURE OF A TRANSCODING PROXY

4. TO TRANSCODE OR NOT TO TRANSCODE

4.1 A Store-and-Forward Image Transcoding Proxy

4.2 A Streamed Image Transcoding Proxy

5. TRANSCODING POLICIES FOR SELECTING CONTENT

5.1 Optimal Policies For Off-Line Pre-Transcoding

5.2. Policies For Real-time Transcoding

6. A SAMPLE SET OF TRANSCODING POLICIES

7. RELATED ISSUES

ACKNOWLEDGMENTS

We wish to thank David Messerschmitt as well as the MobiCom 98 panel for their insights regarding proxy vs. end-to-end issues.

REFERENCES

[Acharya98] S. Acharya, B. Smith, "Compressed Domain Transcoding of MPEG," *IEEE International Conference on Multimedia Computing and Systems*, 1998, pp. 295-304.

[Amir95] E. Amir, S. McCanne, H. Zhang, "An Application-Level Video Gateway," *ACM Multimedia*, 1995, pp. 255-.

[Assuncao98] P. Assuncao, M. Ghanbari, "A Frequency Domain Video Transcoder For Dynamic Bit Rate Reduction of MPEG 2 Bit Streams," *IEEE Transactions on Circuits Systems Video Technology*, vol. 8, no. 8, December 1998, pp. 953-967.

[Avaro97] O. Avaro, P. Chou, A. Eleftheriadis, C. Herpel, and C. Reader, "The MPEG-4 Systems and Description Languages: A Way Ahead in Audio Visual Information Representation," *Signal Processing: Image Communication, Special Issue on MPEG-4*, 4(9):385--431, May 1997.

[Bartlett94] J. Bartlett, "W4 - the Wireless World Wide Web," *IEEE Workshop on Mobile Computing Systems and Applications*, 1994, pp. 176-178.

[Bharadvaj98] H. Bharadvaj, A. Joshi, S. Auephanwiriyakul, "An Active Transcoding Proxy to Support Mobile Web Access," *IEEE Symposium on Reliable Distributed Systems*, 1998, pp. 118-123.

[Bickmore97] T. Bickmore, B. Schilit, "Digestor: Device-Independent Access to the World Wide Web," Sixth International World Wide Web Conference, *Computer Networks and ISDN Systems*, vol. 29, no. 8-13, September 1997, pp. 1075-1082.

[Bjork98] N. Bjork, C. Christopoulos, "Transcoder Architectures for Video Coding," *IEEE Transactions on Consumer Electronics*, vol. 44, no. 1, February 1998, pp. 88-98.

[Border99] J. Border, M. Kojo, J. Griner, G. Montenegro, "Performance Enhancing Proxies," *Internet Engineering Task Force Draft*, June 25, 1999 (see http://www.ietf.org/ under Internet Drafts).

[Bouthillier98] L. Bouthillier, "Synchronized Multimedia on the Web," *Web Techniques Magazine*, 3(9), September 1998.

[Candan96] K. S. Candan, B. Prabhakaran, and V. S. Subrahmanian, "CHIMP: A Framework For Supporting Distributed Multimedia Document Authoring and Presentation," *Proc. ACM Intern. Conf. Multimedia (ACMMM),* pages 329 -- 339, Boston, MA, November 1996.

[Chandra99] S. Chandra, C. Ellis, A. Vahdat, "Multimedia Web Services for Mobile Clients Using Quality Aware Transcoding," *ACM International Workshop On Wireless Mobile Multimedia (WOWMOM),* 1999, pp. 99-108.

[Chang94] S. Chang, A. Eleftheriadis, "Error Accumulation of Repetitive Image Coding," *IEEE International Symposium on Circuits and Systems*, vol. 3, 1994, pp. 201-4.

[Chong96] U. Change, S. Kim, "Wave let Transcoding of Block DCT Based Images Through Block Transform Domain Processing," *Proceedings of SPIE, Wavelet Applications in Signal and Image Processing IV*, vol. 2825, pt. 2, 1996, pp. 901-908.

[Christopoulos99] C. Christopoulos, T. Ebrahimi, V.V. Vinod, J. R. Smith, R. Mohan, and C.-S. Li, "Universal Access and Media Conversion," *MPEG-7 applications proposal*, number MPEG99/M4433 in ISO/IEC JTC1/SC29/WG11, Seoul, Korea, March 1999.

[de los Reyes98] G. de los Reyes, A. Reibman, J. Chuang, S. Chang, "Video Transcoding for Resilience in Wireless Channels," *International Conference on Image Processing (ICIP),* vol. 1, 1998, pp. 338-342.

[Eleftheriadis95] A. Eleftheriadis, D. Anastassiou, "Constrained and General Dynamic Rate Shaping of Compressed Digital Video," *International Conference on Image Processing (ICIP),* vol. 3, 1995, pp. 396-399.

[Fielding97] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, T. Berners-Lee, *HTTP/1.1*, RFC 2068, Jan. 1997.

[Fleming97] T. Fleming, S. Midkiff, N. Davis, "Improving the Performance of the World Wide Web Over Wireless Networks," *GLOBECOM*, vol. 3, 1997, pp. 1937-1942.

[Floyd98] R. Floyd, B. Housel, C. Tait, "Mobile Web Access Using eNetwork Web Express," *IEEE Personal Communications*, vol. 5, no. 5, October 1998, pp. 47-52.

[Fox96a] A. Fox, E. Brewer, "Reducing WWW Latency and Bandwidth Requirements By Real Time Distillation," Fifth International World Wide Web Conference, Computer Networks and. ISDN Systems, vol.28, no.7-11 May 1996, pp. 1445-1456.

[Fox96b] A. Fox, S. D. Gribble, E. A. Brewer, and E. Amir, "Adapting to Network and Client Variability Via On-Demand Dynamic Distillation," *ASPLOS-VII*, Cambridge, MA, October 1996.

[Fox98a] A. Fox, S. Gribble, Y. Chawathe, E. Brewer, "Adapting to Network and Client Variation Using Active Proxies: Lessons and Perspectives", *IEEE Personal Communications*, vol. 5, no. 4, August 1998, pp. 10-19.

[Fox98b] A. Fox, I. Goldberg, S. Gribble, D. Lee, A. Polito, E. Brewer, "Experience with Top Gun Wingman: a Proxy Based Graphical Web Browser For the 3Com PalmPilot," *Middleware '98, IFIP International Conference on Distributed Systems Platforms and Open Distributed Processing*, September 1998, pp. 407-424.

[Gardner99] W. Gardner, "Web Microbrowser Market Is Still Up For Grabs," *Portable Design*, October 1999, pp. 18-.

[Gessler95] S. Gessler, A. Kotulla, "PDA's as Mobile WWW Browsers," *Computer Networks and ISDN Systems*, vol. 28, no's 1 and 2, December 1995, pp. 53-59.

[Han98] R. Han, P. Bhagwat, R. LaMaire, T. Mummert, V. Perret, J. Rubas, "Dynamic Adaptation In an Image Transcoding Proxy For Mobile Web Browsing," *IEEE Personal Communications*, vol. 5, no. 6, December 1998, pp. 8-17.

[Han99] R. Han, "Factoring a Mobile Client's Effective Processing Speed Into the Image Transcoding Decision," *ACM International Workshop On Wireless Mobile Multimedia (WOWMOM),* 1999, pp. 91-98.

[HDML97] Unwired Planet, "Handheld Device Markup Language Specification," *Technical Report Version 2.0, Unwired Planet, Inc*., April 1997.

[Hori99] M. Hori, R. Mohan, H. Maruyama, and S. Singhal, "Annotation of Web Content For Transcoding," *Technical Report NOTE-annot-19990524, W3C*, July 1999.

[Joseph95] A. Joseph, et al, "Rover: A Toolkit For Mobile Information Access," *ACM Symposium on Operating Systems Principles*, 1995.

[Kan98] K. Kan, K. Fan, "Video Transcoding Architecture With Minimum Buffer Requirement For Compressed MPEG 2 Bitstream," *Asia Pacific Broadcasting Union (ABU) Technical Review*, no. 177, July/August 1998, pp. 3-9.

[Kling94] R. Kling and M. Elliott, "Digital Library Design For Usability," *Proc. Conf. Theory and Practice of Digital Libraries*, College Station, TX, June 1994.

[Li98] C.-S. Li, R. Mohan, and J. R. Smith, "Multimedia Content Description in the InfoPyramid," *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP*), Seattle, WA, May 1998. Special session on Signal Processing in Modern Multimedia Standards.

... For more references, order the textbook...