ADAPTATION OF MULTIMEDIA DATA FOR CROSS-DISABILITY ACCESS

Vidhya Balasubramanian and Nalini Venkatasubramanian
Information and Computer Science Department
University of California, Irvine, Irvine California-92697-3425, USA
{vbalasub,nalini}@ics.uci.edu

ABSTRACT

Multimedia transmission over wide-area networks currently only considers the server and network resource constraints and client device capabilities. It is also essential that the accessibility of the multimedia content for different users with diverse capabilities and disabilities be considered. As a first step in building an adaptation framework we need to automatically predict the accessibility requirements of a user, given the abilities the user possesses to access the multimedia information. In this paper we develop an adaptation technique to present the multimedia content to suit diverse disabled user groups by using an ability based classification approach. Using ability-based prioritization, appropriate alternate modalities and quality levels are chosen, to replace the inaccessible modalities. The adaptation process allows for refinements to cater to specific types and degrees of impairments. Our performance results illustrate the benefits of the ability-based adaptation approach.

1. INTRODUCTION

Recent trends in ubiquitous computing focuses on providing information embellished with multimedia data to users anywhere and anytime. While this represents unprecedented opportunities, it is also essential to ensure that the resulting benefits also enrich the lives of the disabled, and the elderly, by making information more accessible. Specific applications like distance learning, video conferencing reduce the barriers of distance, physical constraints etc, and allow users with disabilities to obtain easy access to information. To date, application developers and service providers have focused on extensive functionality and optimal performance (possibly, driven by commercial needs) to ensure that customers can be delivered the best possible Quality of Service (QoS) with the least possible resource consumption overhead, often ignoring the goal of universal accessibility. However the best quality might not correspond to information that is accessible to a user with a disability. For example, increased network bandwidth and real-time video capture technologies may prompt en masse transmission of a live event using streaming video, but such a medium is of little use to the blind. The goal of the SUGA project [SUGA02, BV02] is to enhance existing infrastructures to account for user abilities (and disabilities) so that personalized content can be delivered in a cost-effective manner.

A major step in the realization of the "universal access" goal is that of a **content adaptation framework** that supports personalized (cross-disability) access to information. Content adaptation refers (a) selection of appropriate modalities and (b)modification of the parameters of a specific modality (for instance, an image can be encoded using various resolutions, color levels and intensities) in order to enhance usability and manage heterogeneity. There are multiple facets to developing such a framework foremost of which is a generalized representation of content to allow access to individuals with varying disabilities. Another critical issue is that of cost-effective mechanisms to achieve the adaptation. Specialized mechanisms built into adaptive user-interfaces and access devices for each application may not represent the optimal usage of system-wide resources. By taking advantage of the information about end-user ability prior to delivery of the information to the client, we can optimize resource utilization without sacrificing on the utility of the delivered content. Such functionality typically resides in a middleware layer. In the envisioned approach, discussed in the paper, user abilities are considered to provide suitable modalities of information to improve the end-user accessibility. For e.g., closed captioning could be provided in place of audio to a person with a severe hearing impairment. Furthermore, the alternate modalities often consume fewer network and server resources. For example, video content on distance-learning sites can be transmitted to blind users as a text summary, accessed via screen readers. The bandwidth saved can be used to send relevant modalities to other users, thereby admitting a larger number of users into the system.

Our focus is on developing a framework that provides "*cross-disability access*" to multimedia information the deals with a wide range of users with different disabilities, including the blind and visually impaired, the deaf and hearing impaired, cognitively impaired, motor impaired and the deaf-blind. In order to provide the relevant media types, and modalities, we use information about the ability that an end user possesses to determine the content to be delivered to specific clients. Using studies in the cognitive sciences domain, a rule-base is computed using the different user abilities based on which the adaptation is performed. Based on the user profile and system resource constraints, appropriate modalities are selected so as to maximize the utility of the information communicated to the user with disabilities. The remainder of this paper is organized as follows. Section 2 discusses related work. Section 3 describes the ability based transcoding framework. It addresses how the priority values for the multiple modalities are calibrated, describes the architectural layout of the transcoding system and introduces the priority based transcoding algorithm. We also present a refinement of the transcoding process for vision impairments using existing studies about perception of vision in Section 4. Section 5 evaluates the performance of the proposed transcoding mechanisms. We conclude in Section 6 with future research directions.

2. RELATED WORK

Problems of accessibility apply to both hardware and software systems. An individual who cannot use her hands or see cannot use a mouse, type on a keyboard, see a screen, or navigate easily through a user guide. Software components such as operating systems and application programs include many roadblocks to users with disabilities. There have been many different solutions proposed to tackle the problem of accessibility. Most universal accessibility solutions have been studied at the interface level, by improving the software interfaces to suit the users [SJ98, K98, G95, F98]. Other work addresses the use of assistive technology such as access equipment and input devices, and shows how such tools can be made more popular by alleviating some of the accessibility issues [IN, SGS93, SDD96, V98]. Universal Design techniques have been applied to information technologies to make all information perceivable, including keys and controls, facilitate navigation, and ensure compatibility with commonly used assistive technology [G00, JV01, SS01, FK99]. Adding the adaptation functionality in the interface and in specialized devices is very important to improving accessibility. However since information is increasingly being accessed over networks, it is desirable to push some of the adaptation functionality into the middleware layer, where user, application and system level information can be integrated.

Basic adaptation techniques have been addressed for specific disabilities. Blindness is a widely studied disability issue; tools such as screen readers, audio HTML interfaces and auditory navigation are proposed techniques to alleviate some problems faced by impaired vision [OY97, T94, W98, MP98, J98]. Image exploration and accessing graphical user interfaces are the focus of research in developing accessible interfaces for the blind. A guidance system outputs directional information to enable exploration of tactile graphics [RM96, WD99, K'98, B98]. Emacspeak has used the speech-enabling semantic approach to provide intelligent spoken feedback [R01]. Systems catering to users with hearing-impairments deal with recognition of sign languages, online generation of closed captions, and voice to text recognition capabilities [P98, HH99, AF98, B99]. Some of the proposed solutions to alleviate problems faced by users with motor impairments are use of gestures, force-feedback technology, eye tracking and vocal programming [JB98, KR98, HL01, AM00]. Cognitive Disabilities are the least studied with respect to computer access. For the benefit of cognitive and language impaired users certain solutions proposed are context-aware assistance, like hints, descriptions etc and natural language processing [CD98, L98]. With a huge user population users with varying impairments may need to access the same information. However most of these techniques described are tailored to suit individual disabilities and do not provide a generalized framework for the entire spectrum of users with different disabilities. In our proposed work we cater to cross-disability access, by providing a generic solution to a variety of disability groups, which can also be further tailored to suit individual preferences.

Accessibility to multimedia on the web is increasingly an important area of research [F97, ZV01, LM96, RG00] since most content is embellished with multimodal data. One way to make multimodal information accessible to users with disabilities is by transcoding the information appropriately [HS00, TA00]. Much of the transcoding and adaptation techniques on the web are text or image based; much work remains in the adaptation of more dynamic content such as video/audio for universal access. HMML is an increasingly popular technique for the adaptation of multimedia to meet specific user needs [P97]. Transcoding techniques have been applied for adapting to client device heterogeneity [MS99], to adapt to bandwidth heterogeneity [HS01] and to ensure that some level of QoS is maintained despite network fluctuations [SEV00, SEV99]. Object based transcoding of multimedia information [VW01] presents another approach to content adaptation where objects that constitute a multimedia scene are identified and prioritized. While many of these transcoding techniques help to adapt multimedia information for varying conditions, they do not consider the capability of the user to access the information. Our framework aims precisely at doing that, by adapting these transcoding techniques to suit the problem at hand - that of adapting multimedia information for accessibility over a cross disability platform.

Designing a transcoding system includes many design challenges. Many papers have discussed the issue of multimedia server design that is essential when considering server transcoding [VG94, BP94, SGV95]. Techniques have been proposed to address placement of media on disk to ensure real-time retrieval [AO92, YC92], admission control procedures to maximize server throughput [VR93] and buffer management policies to minimize memory requirements [GC92, LS93]. Replication and striping strategies or optimizing storage across disk arrays is described in [KK93, TP93]. Data placement techniques help to efficiently distribute the multimedia data without imposing undue load on the system [BS00, MW97]. A dynamic segment replication scheme for partial replication of video objects is proposed in [DK95]. This mechanism permits the executing requests to be dynamically migrated to a replica on a less loaded device. The trade-offs between storage space and transfer bandwidth are highlighted in [DS95]. More recent work studies the impact of dynamic replication techniques for continuous media servers [CG98]. A number of broadcast and multicast techniques have been studied for effective utilization of network bandwidth [AY96, HS97, GK98, AA96] for video-on-demand systems. Batching mechanisms group closely spaced requests for the same videos, thereby minimizing disk I/O and network bandwidth [DK95]. The above-mentioned design considerations address resource usage keeping quality of service in mind. In order to effectively build the adaptation framework these design techniques should be appropriately modified so as to also cater to user accessibility. For e.g. data placement in the context of our problem should consider the pattern of user accesses to efficiently support adaptation for cross-disability access. This paper specifically caters to translating the capabilities of the users to providing the information that is best accessible, given the resource restrictions of the system.

3. AN ABILITY DRIVEN APPROACH TO MODALITY SELECTION

The proposed multimedia adaptation framework should identify the user abilities and decide the appropriate format of the information so that the accessibility is maximized, while also efficiently allocating system resources. Hence the framework should perform the primary mapping operations as shown in the Figure 1.

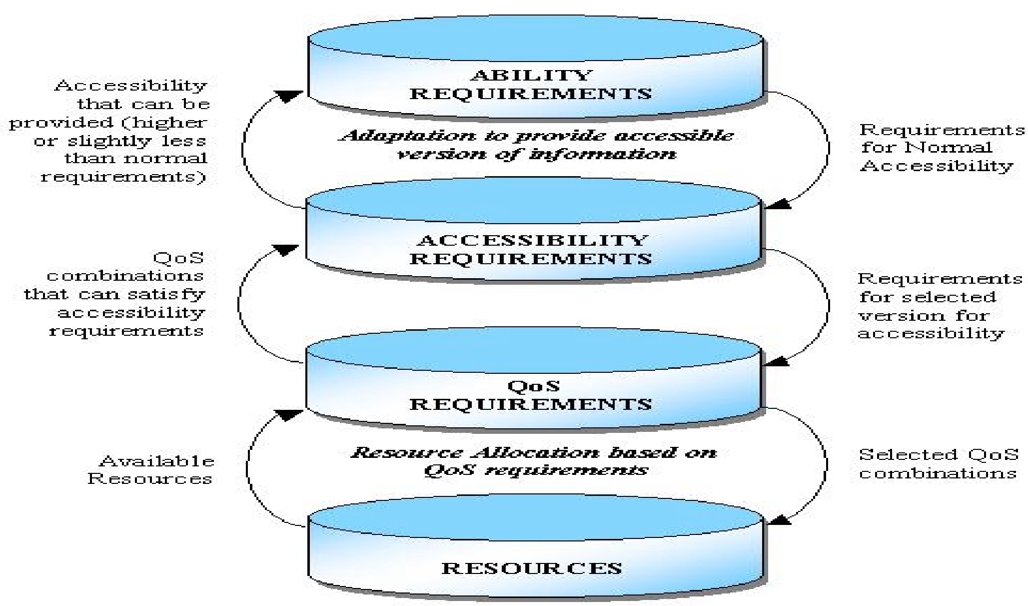

Figure 1: 3 layer mapping required for developing the adaptation framework

- ***Ability to Accessibility mapping***: This step requires that the system automatically identify how the information will be accessible to the user, based on the user's ability. It should be able to identify the modalities that are poorly accessible and should posses the intelligence to adapt that information in a form that is accessible. For instance if a user has poor visual acuity, accessibility of an image could be improved by either increasing image size, increasing contrast, or changing brightness or doing all three at some appropriate level.

- ***Accessibility to QoS mapping***: Once parameters for accessibility are calculated, it is essential to identify the QoS parameters that can satisfy the given accessibility requirement. For instance the accessibility requirement of an image might specify a particular image height, usable colors, required brightness level etc. These should be mapped to the QoS parameters of color levels, image resolution etc. For a given accessibility requirement there might be several combinations of the QoS parameters that might satisfy the requirements.
- ***QoS to Resource mapping***: Quality based Resource allocation is a well-studied problem[RL98,LL99] and this will be the final step in our transition from user ability to the problem of resource allocation.

In this paper we develop a novel technique for automatically doing the first step of mapping from the user ability to the accessibility requirement. Given the ability of a user we try to quantify how accessible a particular modality is to that user and using this accessibility specification for effective resource utilization. Section 3.1 describes a method to do this ability accessibility translation. In order to effectively select the appropriate modalities, a suitable architecture is required. We describe this specialized architecture for adapting information in section 3.2 and then describe an algorithm to select the appropriate modality based on this mapping in section 3.3. In section 4 we discuss a fine-grained transcoding technique for different vision impairments using analytical models of the human vision system.

3.1 Ability to Accessibility Translation

Several approaches can be used for the purpose of quantifying the accessibility of the information provided to the user. We use two well-studied approaches to model user abilities and analyze their relative advantages and disadvantages. The two approaches are a) the abilities requirements approach and b) the analytical approach

- ***Ability Requirements Approach***: The cognitive sciences community has conducted abundant studies for characterizing the abilities of humans. One approach is the Ability Requirements approach adopted from Fleishman and Quaintance [FQ84]. Several abilities, required for performance of tasks have been identified and classified into cognitive, perceptual, physical and psychomotor abilities as shown in Table 2. In this approach tasks are described, compared and contrasted in terms of the abilities that each of them requires of an individual performer. Ability rating scales have been developed and validated after a lot of experiments and systematic studies. (see Ability Requirements Approach, Chapter 12 in FQ84] . Human tasks are classified based on a seven-point scale using definitions of high and low ability requirements as scale anchors. It has been proved that this method of rating tasks provides a statistically reliable tool for assessing amount of ability requirement. Table 1 gives one such example of tasks representing the ability ‘Near Vision ‘, adopted from FQ84-Appendix C.

Task Items for the Selected ability– Near Vision	Mean	Standard Deviation
Read the fine print of legal journals	5.89	0.90
Cut and mount color film transparencies	4.67	0.97
Plug in a TV set	1.44	0.62

Table 1: Sample Ability levels required for different tasks (Near Vision)

Cognitive Abilities	Perceptual Abilities	Physical Abilities	Psychomotor Abilities
1.Oral Expression	1.Speed of Closure	1.Static Strength	1.Control Precision
2.Written Expression	2.Flexibility of Closure	2.Dynamic Strength	2.Multilimb Coordination
3.Oral Comprehension	3.Spatial Orientation	3.Explosive Strength	3.Response Orientation
4.Written Comprehension	4.Visualization	4.Trunk Strength	4.Rate control
5.Fluency of Ideas	5.Perceptual Speed	5.Extent Flexibility	5.Reaction Time
6.Originality	6.Near Vision	6.Dynamic Flexibility	6.Arm-Hand Steadiness
7.Memorization	7.Far Vision	7.Gross Body Coordination	7.Manual Dexterity
8.Problem Sensitivity	8.Visual Color Discrimination	8.Gross Body Equilibrium	8.Finger Dexterity
9.Mathematical Reasoning	9.Night Vision	9.Stamina	9.Wrist-Finger Speed
10.Number Facility	10.Peripheral Vision		10.Speed of Limb Movement
11.Deductive Reasoning	11.Depth Perception		11.Selective Attention
12.Inductive Reasoning	12.Glare Sensitivity		12.Time Sharing
13.Information Ordering	13.General Hearing		
14.Category Flexibility	14.Auditory Attention		
15.Speech Hearing	15.Sound Localization		
16.Speech Clarity			

Table 2: Ability Classification adopted from Fleishman and Quaintance

- ***Analytical Approach***: Another approach for quantifying the ability to assimilate information would be to use the individual analytical models of human faculties [SW98, SW99, EVV, VL97, CCRMA]. The vision models help calibrate visual acuity, based on illumination, distance, size of object, etc. These models help quantify the distortion of image as viewed by the user, giving better view of the perceptual distortion a naive Mean Square Error Metric.

The ability requirements approach gives a more wholesome approach, applying to not only information assimilation, but also interactive computing, which require more motor skills. This approach is also helpful when considering people with multiple impairments, a factor missing in other models. Furthermore, the ability-requirements approach also accounts for the abilities that are primarily affected as well as other abilities that are affected as a result. For instance, even a mild hearing loss of (25 – 40dB) might result in mild language retardation [SD84]. Hence this approach is useful for selection of modalities based on user abilities as a first step to content adaptation. On the other hand specific analytical models of the human faculties like vision and auditory system help to further adapt the selected modalities thus increasing accessibility, since they contain significant details about the physiology and functionality of the human faculties. In this paper we use a ***Hybrid approach*** that combines the two approaches described above. We use the ability requirements approach for a coarse grained adaptation of information, specifically modality selection. The analytical approach is used for fine-grained adaptation, i.e. to fine-tune the selected modalities to improve accessibility. This approach for fine tuning information for specific vision impairments is discussed in Section 4.

Prioritization of Modalities: The ability model requires specific tasks for which the required abilities will be determined. In our case, the task at hand is to assimilate multimodal information. The first step is to find the ability required to perform the task of assimilation of each of these modalities, by a normal user. Specifically we consider the following modalities - audio, video, text and image. In addition we maintain a set of alternate modalities of the multimedia information in case the original modalities are inaccessible. The alternate modalities are, closed-captioned audio, captioned video, audio description of video, tactile text, text summary, summary of audio and video, text description of image and audio description of image. For each of these original modalities, we select the abilities that are necessary for retrieving the respective information. For e.g., in the case of video information, sample abilities considered include near vision, visual color discrimination and reaction time; we do not consider abilities like trunk strength and dynamic strength. For the abilities selected, we decide the closest matching task given in the sample task classification in Fleishman and Quaintance [FQ84], to the application at hand. For e.g. consider the ability 'Near Vision' in Table 1. It is defined as "the capacity to see close environmental surroundings" [Appendix B in FQ84]. Considering the small size of fonts of text data, we correlate the assimilation of text data to that of reading fine print of legal journals. This gives us the required ability of near vision to read text as 5.89. Similarly the values for other abilities for normally accessing that modality are calibrated.

Let there be *m* number of different modalities. Let *G* be the set of user groups {blind, low-vision, deaf, hard of hearing, motor impaired, cognitively impaired, deaf-blind} and let $usergp \,\varepsilon\, G$. Let n_j be the total number of abilities selected for the particular modality *j*. $<A_{j1}, A_{j2},, A_{jn}>$ denote the subset of the abilities selected. The values of each of these abilities required to access modality *j* is given by $<W_{j1}, W_{j2},..W_{jn}>$, where $0 < j < = m$. Then the total ability required by a normal user to access a modality, *j* is given by

$$Ability_j^{normal} = \sum_{i=1}^{n_j} W_{ji} A_{ji} \qquad 0 < j < = m$$

The next step is to calculate the total ability an impaired user possesses to access the above information. Based on different statistics on the different impairments, and how these different levels of impairments affect the information processing [LH02,WHO2,SS01], we correlate the tasks selected as required for the normal user to the maximum possible task a user with a specific disability can perform. Consider the same ability as before, i.e., near vision. The required task level was to read fine print. For a blind person (no light perception), even the minimum task possible cannot be performed, relying on this ability. Hence this ability is absent and is assigned a value zero. A vision-impaired user might be able to perform the task of plugging a television set (considering the worst case), but with difficulty. We calculate the value of that ability the user possesses using the mean and standard deviation values from [FQ84] (see sample values in Table 1). This yields a value of 1.3 as the amount of near-vision a severely vision impaired user possess to access fine print. For specific vision impairments we can refine these values, based on the vision models and visual acuity values of different levels of vision impairments (see Section 4).

For a given disability category *usergp*, the ability possessed to access modality j is calculated as follows. For the selected abilities $< A_{j1}, A_{j2}, \ldots ,A_{jn}>$ for modality *j*, the weights denoting the corresponding amount of that ability the user possesses (the upper limit being that required by that of a normal user) is given by $<U_{j1}, U_{j2}, \ldots U_{jn}, >$, where $0 < j < = m$. Hence the total ability the user possesses to access the particular modality j is given by

$$Ability_i^{usergp} = \sum_{i=1}^{n_j} U_{ji} A_{ji} \qquad 0 < j < = m$$

The accessibility of the modality *j* to the user group *g* is defined by the accessibility ratio (AR) value that is calculated as below.

$$AR_i^{useergp} = \frac{\sum Values\ of\ abilities\ the\ user\ possesses}{\sum Values\ of\ abilities\ required\ to\ access\ the\ \mathrm{mod}\, ality} = \frac{Ability_i^{usergp}}{Ability_i^{normal}}$$

Once the accessibility ratio values are calibrated, an accessibility matrix is developed where the cells indicate the accessibility of the corresponding modality to the person with the particular disability in the corresponding row. Table 3 below gives a sample accessibility matrix. An extended version of this accessibility matrix is also created for the alternate modalities like closed captioned audio, tactile text and auto-summarization.

Disability	Audio	Video	Text	Image
Blind	0.990	0.063	0.172	0.145
Low Vision	0.990	0.608	0.554	0.669
Deaf	0.090	1.000	0.996	1.000
Hardof Hearing	0.575	1.000	0.996	1.000
Cognitive	0.685	0.830	0.776	0.888
Motor	0.903	0.833	0.751	0.870
Deaf blind	0.090	0.063	0.172	0.145

Table 3: Accessibility Matrix showing accessibility of the particular modality to different user groups

3.2 System Architecture

The architecture of the proposed transcoding system is illustrated in Figure 2. For the purpose of transcoding the following information is needed- a) user profile information b) available server resources like server bandwidth, CPU resources, storage etc and c) available client resources like client CPU, available buffer at the client etc. The essential components of the architecture are the transcoding meta-data that stores information to assist transcoding, the content source that stores the data content, and the coarse grained and fine-grained transcoders, which perform the actual transcoding function.

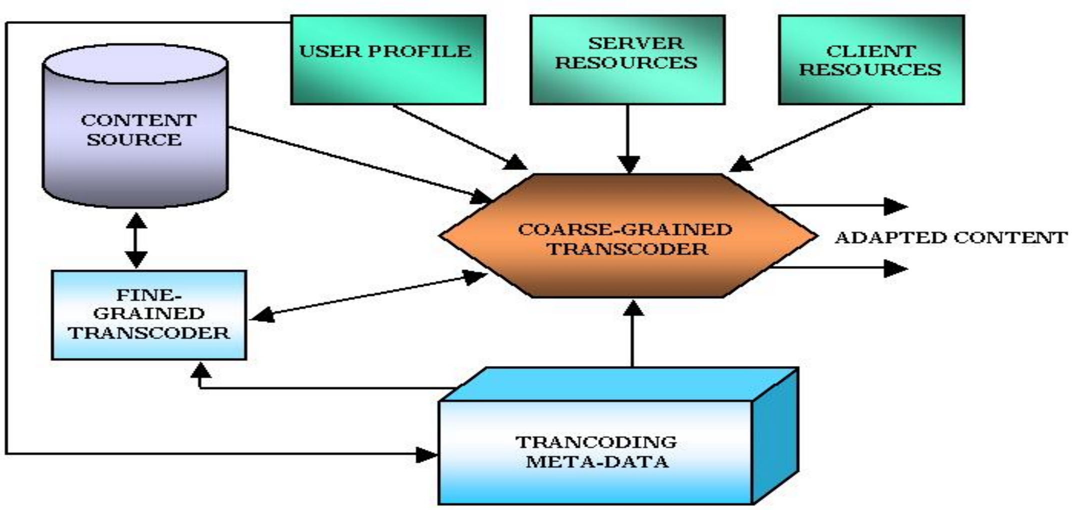

Figure 2: Architecture of the Transcoding System

Transcoding Meta-data: The transcoding meta-data represents the following information:

- Accessibility information: An accessibility matrix containing the accessibility of each of these modalities to each of the user groups is created based on the application and the abilities required to access the information. This accessibility matrix encodes the utility of each modality *(and quality level)* to each user class for that specific application.
- Media related information: This part of the transcoding meta-data represents the information about the media at different quality levels. Specifically, it stores information about the different quality parameters for each media-type (e.g. color, intensity, resolution for images, sampling rate for audio, frame rate for video etc.) and encodes the importance of each parameter to that media type. The selection of appropriate information content for the particular user is based on both the media information and the accessibility value – the accessibility value defined earlier is refined based on the quality level.
- Resource requirements of different media types at different quality levels: This component represents the resource requirements for the different media types at the different quality levels. Resources required include storage, network bandwidth, memory and CPU resources. To represent resource requirements of different modalities at different quality levels, we will leverage work on efficient multidimensional data structures [ZM99] to accurately represent resource information and derive the resource needs for a particular user session. A simple example is a bandwidth matrix, which represents the bandwidth required for the media type, say an image at different resolutions and different color encoding schemes as shown in Table 4. The different versions represent different resolution levels. For e.g., version 1 of audio, is single channel audio, sampled with 8 bits at a rate of 8Khz. We can also use other parameters to efficiently utilize system resources. Detailed resource models have been studied [LL99, RL97]. Since it is beyond the scope of this paper to be developing extensive resource models, we will use simplistic resource considerations. Detailed models will be used for future work.

Modality	Version1	Version2	Version3	Version5
Text	10bps	640bps	1.8kbps	3.8kbps
Image	614.4kbps	2.458mbps	7.373mbps	18.88mbps
Audio	64kbps	96kbps	224kbps	1.412mbps
Video	184.32mbps	300mbps	1.33gbps	1.88gbps

Table 4: Sample bandwidth matrix for the uncompressed modalities

Content Source: The content source stores the necessary information content. All the other modules of the transcoding system access the information from the content source. As and when the content adaptation is done on-line or offline by the transcoders the information in the content source is updated. The information in the content source can be accessed via the transcoding meta-data where different media types at different qualities are represented.

Coarse-Grained Transcoder: This module performs coarse grained-transcoding, which is primarily modality selection. Here the appropriate modalities are selected from the content source based on the meta-data information and accessibility matrix. The modalities are selected so that they maximize utility to the user while minimizing system load and resource requirements.

Fine-Grained Transcoder: This is the module where the finer granularity of transcoding is performed. The transcoding could result in the creation of new modalities such as close-captioned audio, video summarization (if they do not already exist) or new quality levels e.g. adapted images or video with changes in resolution, frame rate and deletion of irrelevant components. Adaptation of quality levels of specific media types is largely dependent on the nature and extent of the disability. A specialized treatment of transcoding for the vision-impaired is discussed in Section 3.3. We assume that content for fine-grained transcoding is created offline.

The operational dynamics the architecture mentioned are given below. When a request for the information comes in the, user profile is checked with the existing user profiles stored. Any new profiles are added. The client profile is checked to calculate the client capabilities and available client resources are recorded for that client. The coarse grained transcoder picks the most appropriate modalities at best possible resolution from the different modalities and resolutions available at the content source for that data requested, based on the accessibility values in the meta-data calculated, and the resources in the client and server. More fine-grained adaptation is provided by the fine-grained transcoder. The fine-grained transcoder adapts the modalities using the analytical models of the human faculties, thus improving the accessibility of the content. The content source is updated with more fine-grained content. The coarse grained transcoder finally transmits the adapted content to the client.

3.3 An Ability Driven Transcoding Algorithm

To effectively transcode the information we use a priority based transcoding algorithm that uses accessibility information from the accessibility matrix described above along with resource constraints within the environment to prioritize content that may be delivered for each incoming request. The transcoding technique employs an iterative procedure to select the appropriate modalities. When a request comes in, the following factors are considered – (i) User profile (consisting of the user abilities); (ii) Server Resources, (iii) Client Resources and (iv) Resources required by the request. Based on the user profile the appropriate accessibility is determined from the accessibility matrix. The resource consumption for each of the modalities at various resolutions is represented in a resource consumption matrix. The value Rjk at each cell represents the value of the resources required for the particular modality j, at the resolution k. The resources can include server bandwidth, link bandwidth, storage, CPU, etc. The primary bottleneck resource we have considered is server bandwidth since this paper does not focus on CPU-intensive online transcoding and compression techniques. The QRAM model provides heuristics for optimal allocation of resources to applications with different QoS requirements [RL97, RL98]. The ability driven resource allocation problem at hand is similar to the QoS Based resource allocation, however with accessibility requirements instead of QoS requirements. The QRAM model works on the fact that the quality and hence utility increases with increased allocation of resources. However accessibility does not increase with increased resources allocated, (infact defining an analytical relation between accessibility and resources is not straightforward). Hence adapting such detailed resource allocation schemes in the context of maximizing accessibility is a complex problem and is beyond scope of this paper. Here we consider a simple resource allocation scheme with the primary goal of providing maximum accessibility to the user that is based on a greedy strategy. According to the accessibility requirement of the user, the resources are allocated. If the resources are constrained the modalities with the best accessibility requirements that is satisfiable is provided.

```
While there is a user request:
  Select the best accessible modalities at best accessible resolution with P_ij > P_ij^min from priority matrix
    If admissible by server // first come first served admission control and  within client resource availability
       admitted = true;
    Else
     Repeat
       Repeat
          If  (R_s > 0) && (R_c >0) && (R_jk > min(R_s, R_c)
             Reduce the resolution of the least priority modality.
             If  (R_jk <R_s) && (R_jk<R_c)
                     admitted = true;
             End If
          Else
             admitted = false;
          End If
       Until accessibility of modality with reduced resolution P_ijk < P_ij^mink
        Reduce the priority of selected modalities by ε
        Select the new combination of modalities(with revised accessibility ) so that  accessibility is maximized
     Until the request is admitted or P_ij< P_ij^min  for some some modality.
     If  (admitted)
       Send requested data that is adapted
     Else
       Reject request
    End If
End While
```

Figure 3: The Resource Aware Ability Driven Transcoding Algorithm

Given n distinct user groups , m distinct modalities and K resolutions for each of these modalities

- M_{jk} represents the index to the modality j at resolution k , $0<=j<=m$; $0<=k<=K$
- P_{ijk}, which gives the ability based priority value for the particular user group i for accessing M_{jk}, $0 <= i <= n$; $0<=j<=m$; $0<=k<=K$
- P_{ij}^{min} is the minimum accessibility value required for modality j as required by user group j. For now we assume the user provides the minimum requirements. Calculating the minimum accessibility requirements will require detailed user studies that will be done as part of future work.
- P_{ij}^{max} is the maximum accessibility beyond which there is no significant improvement in user experience of accessing modality j. This value also requires user studies and for now we assume the maximum possible accessibility as that for a normal user (which is 1.0 per modality).
- R_{jk}, the resource requirement for M_{jk}, $0<=j<=m$; $0<=k<=K$
- Available server resource R_s, and
- Client resource R_c.

The objective of the transcoding process can be stated as a maximization problem with the following specifications:

Choose M_{jk} for user group i so as to Maximize $\Sigma (P_{ijk})$ subject to the following constraints

$$\Sigma(P_{ijk}.R_{jk}) <= R_s$$

$$\Sigma(P_{ijk}.R_{jk}) <= R_c \quad 0 <= i <= n; 0<=j<=m; 0<=k<=K;$$

In other words the objective is to maximize the total priority value so that the total resource consumption is less than available server resources R_s, and Client resources R_c.
A greedy generic, algorithm for the maximization problem is given in Figure 3. The algorithm is greedy in that it attempts to maximize priority for a single request at a time without considering further requests. From the meta-data, initially the best accessible modalities at maximum accessible resolutions with accessibility values above the threshold value P_{ij}^{min}, for that particular user group is chosen. Admission control is then performed at the server end where the request can either be admitted or rejected, based on resource availability. If the required resources exceed the available resources, we reduce the resolution of the least priority modality, till the resolution can no more be reduced without affecting accessibility significantly. Let P_{ij}^{mink} be the accessibility of the least tolerable resolution k, of modality j for usergp i. P_{ij}^{maxk} denotes the maximum accessibility of resolution k of modality j for usergp i , beyond which any higher resolution will not be of significant help. If still not admissible then we change the combination of the modalities chosen. To do this we recompute the priorities of chosen modalities by reducing the priorities by ε where ε is the factor by which priority value is reduced in consecutive iterations. ε can be calculated based on the accuracy requirements. One simple assignment would be to partition the interval between P_{ij}^{min} and P_{ij}^{max} into small intervals then assign the size of the interval to ε.In our experiments we set $\varepsilon = (P_{ij}^{max} - P_{ij}^{min})/100$. With the new priority values we again pick a combination of modalities that provide maximum accessibility and repeat this admission process control. If we reach a point where the threshold value P_{ij}^{min} for some modality is reached then the request is rejected.

4. Specialized Transcoding for Vision Impairments – A Case Study

In this case study, we consider different vision disorders that exclude legal blindness, where transcoding techniques are applied to change the quality of the image and text information. The quality in this context encompasses resolution, color contrast, brightness, and font size. The transcoding is done to adapt to the varying perception levels of vision that a user possesses. For effective transcoding it is necessary to have a good understanding of the human visual system and its functions. According to the different studies on the human visual system, [GW95, AO02,CS02,VC02-1,VC02-2,WH01,CH93,MC92,IP02] effective visual perception is dependent on a number of factors. These factors are as follows:

- **Visual Acuity** (V): It is the ability to resolve a spatial pattern separated by a visual angle (α) of 1 minute of arc. In the context of textual information, it is the ability to resolve contrast on a white background that subtends 5 minutes of arc. Using the common Snellen fraction, visual acuity is 20/20 for a normal user.

$$V = 1/a \quad \text{where} \quad \alpha = h/d * 3438 \text{ minutes}$$

  where *h* and *d* are the height and viewing distance of the media information.

- **Contrast Sensitivity** (S): It is a measure of the limit of visibility of low contrast patterns before they become indistinguishable from a uniform field. It is the function of the coarseness/fineness of image features or spatial frequency. It is defined as the reciprocal of contrast.

$$S = 1/\text{contrast where contrast} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

where I_{max} corresponds to the maximum intensity level and I_{min} the minimum intensity level of the media.

- **Brightness** (B): Brightness is the subjective evaluation or interpretation of the amount of light reaching the visual system. For simplicity it can be related to the luminance of the image, which is the product of light, reflected and light reaching the surface. It varies logarithmically with luminance. For a RGB image where, R, G, B are the pixel intensity values for red, green and blue pixels brightness is calculated by averaging the luminance for all the pixels.

  Luminance = 0.3R +0.59G+0.11B

- **Color Perception** (C): For an arbitrary image, the three responses are X, Y, Z (also called tristimulus values) for red, green and blue , can be calculate from the RGB color values as follows

  X = 0.6067 R + 0.1736 G + 0.2001 B

  Y = 0.2988 R + 0.5868 G + 0.1143 B

  Z = 0.0000 R + 0.0661 G + 1.1149 B

- **Visual Field** (θ): the visual field is the total area where objects can be seen in the peripheral vision while the eye is focused on a central point. Each eye has a horizontal field of view of 150 deg (60 deg towards the nose, 90 deg to the side) and a vertical field of view of 120 deg. A visual field of 20 deg indicates blindness. Visual field affects the visual angle of the user, which influences the amount of information received by the eye.Hence the visual accessibility VA can be defined as a function of the above factors.

  $$VA = f(V,S,B,C,\theta) = norm(V) + norm(S) + norm(B) + norm(C) + norm(\theta)$$

  where norm(i) is a normalized value for each of these factors, and 0<=norm(i)<=1. For simplicity, we first normalize each of these factors to fit into a uniform scale and sum the normalized values to obtain a measure of visual accessibility.

Our goal is to maximize visual accessibility for users with varying degrees and types of vision impairments. The normalization process is specific to the parameter; we reduce each factor to a numerical value between 0 and 1 where 0 represents a high degree of impairment and 1 represents the value of that factor for a normal unimpaired user. While visual acuity (low in users with refractive errors like myopia) can be directly represented without further calibration, the normalized value of contrast sensitivity is calculated as 1-1/S, where S is the measured contrast sensitivity for a user. The normalized value of color sensitivity is the average of individual sensitivities to red, blue and green colors.

There are also intricate dependencies between the different factors. We model the dependency of brightness on contrast - i.e. as contrast increases the perceived brightness increases. Though other factors like luminance also affect brightness we do not model these dependencies due to complexities involved in changing luminance. To simplify the transcoding process, we only model those dependencies that have a significant impact on visual perception. Furthermore, the transcoding decisions may conflict in the case when multiple visual factors are impacted. For instance, a low visual acuity requires enlargement of image/text while a low field of vision will trigger a size reduction. The transcoding system must contain rules to handle such conflicts; for instance, in the above case, a bounded enlargement factor to suit the feasible field of view must be calculated.

The effect of different impairments with respect to the above factors is given in [SAP, LH02,RR98,RE02,RN02,RI02]. Based on the information we decide the level of transcoding to be performed for text and image data. Table 5 gives the general level of transcoding required for different impairments. For simplicity we consider only a limited number of resolution levels and a limited number of color-coding representations. For e.g. we consider the maximum size of an image to be 1024x768, beyond which a magnification is not feasible. Similarly the minimum size of the image or text is also defined. Similarly the color levels considered are Black and White images, 16 colors, 256 colors and 24 bit colors. From the available representations the appropriate quality level for the specific impairment is chosen based on the rules defined in the table 5.

Impairment	Affected Part of eye and effect	Factor affected	Transcoding required
Refractive errors (near vision, far vision, presbiopia)	Lens which results in poor focus and hence blurred image	Visual acuity (v) Norm(v) = v Other factors unaffected by transcoding	For given acuity v and size of image or text y Y = size /v New norm(v) = y*v/size
Cataract and Corneal Pathology	Affects lens resulting in hazy, distorted vision	Visual acuity (v), contrast sensitivity Norm(v) =v; Norm(s)= 1-1/s Other factors unaffected by transcoding	For given acuity v and size of image or text y Y = size /v For given contrast sensitivity s, original contrast c1, the required new contrast is 1/s. Choose appropriate contrast level c2 closest to 1/s. New norm(v)=y*v/size; New norm(c)= c2;
Diabetic Retinopathy and Macular Degeneration	Macula, retina and fovea resulting in loss of detail, distortion and reduced color perception	Visual acuity (v), contrast sensitivity (s), color(c) Normalization for v and c same as above Norm(c) = (X+Y+Z)/3	Transcoding for factors v and s same as above For given responses X,Y,Z, RGB color values calculated and closest color level chosen. New norm(c) = Total sensitivity of Selected colors/no of colors
Glaucoma, Retinis Pigmentosa	Retina, optic nerve, resulting in poor peripheral vision and brightness perception	Field of View (θ), brightness (b) Norm(θ)= θ/60; Norm(b) = 1-1/b	For field θ and size of image in x dimension, new size = $\frac{\theta * x}{60}$; Norm(θ) = ((x*θ)/ new size)/60 For brightness b the selected intensity level is e^b; For selected intensity I new norm(b) = 1-(1/logI)
Color Blindness	Cones, can affect red, green or blue cones	Color Perception (c) Norm(c)= (X+Y+Z)/3	For given responses X,Y,Z, RGB color values calculated and closest color level chosen. New norm(c) = Total Sensitivity of selected colors/no of colors

Table 5: Transcoding for vision impairments using detailed vision studies.

5. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the ability driven transcoding scheme for a variety of disabilities and under varying resource constraints. We first explain the simulation environment (system model, user model and request model) following which we discuss the different experiment results for the ability driven transcoding scheme in Section 5.1. The results of the specialized transcoding for vision impairments, and their implications are described in Section 5.2. Finally in Section 5.3 we discuss the adaptation of a sample distance-learning portal.

System Model: The basic server configuration consists of a single data source, with a storage of 50 Gbits. We assume that the server has enough storage to hold all the required modalities for the users considered. The server has a transfer bandwidth of 100Mbits/second in the basic configuration. For simplicity, CPU and memory resources of the data sources are not considered as bottlenecks. The data source consists of a single video with resolution – 640x480, and a frame rate of 30Frames per second. The audio has the following characteristics, 44KHz, sampling rate, and 32 bits per sample with two channels. The data source additionally has text associated with each clip of video. The average duration of each session is 30 to 60 minutes. The alternate modalities associated are, video captioning and summarization (text and tactile), audio captioning (text, and tactile), summary of text, tactile representation of text and summary. We assume that these representations are generated and already stored along with the original piece of information. In this simulation of coarse-grained transcoding we consider only a single resolution level for the modalities, since computing accessibility values of different modalities at different resolution requires additional studies. Our future studies of different models of human faculty will address this issue. For the specialized transcoding we modify this model by adding more resolutions of data. The data source now contains image stored at different resolutions (640 x 480, 800 x 600 etc), and different color levels (black and white, gray scale, 24-bit color etc. Since the specialized transcoding is for specific vision impairments we do not consider audio.

Request Model: We assume the requests for the data follows a Zipf-like distribution and model the requests using the same. Accordingly, the popularity for the video follows Zipf's law, with the request arrivals per day for the data v given by:

$$\text{Probability } (v_i \text{ is requested}) = K_V / i \quad \text{where, } K_V = \left(\sum_{1}^{V} \frac{1}{i} \right)^{-1}$$

The requests arrivals per hour are assumed to follow a Zipf-like distribution [BP95]. From the request arrivals-per day for the data and the probability distribution of requests each hour, the number of requests that arrive in each hour for the information is computed. The user category is integrated with the request model, hence modeling the user arrival as discussed in the following paragraph.

User Model: We also model the user arrival, based on the demographics of the different types of user categories. Classification of the different disabilities, and the different demographic statistics has been based on the reports from the World Health Organization [WHO1, WHO2]. The visual acuity values and hearing values, that help in deciding the user groups are derived from the different studies on these impairments. [ACI99, WHO3, RL01]. Some studies on motor and cognitive impairments have been researched, but more towards these areas have to be studied. [CPA, SCI, CUD]. (Note, motor impairments here more specifically concerns spinal cord injuries, as it deals with limited use of upper limbs, other motor impairments are being studied to perform efficient transcoding for motor impairments). Once all the above-mentioned information is collected, the user model is developed. The user profile includes, visual acuity, auditory perception, motor ability and cognitive ability, based on which the user is grouped into one of the existing classifications. A random mix of users is generated based on the demographic information. We assume that all the users fall into any of the categories defined.

Metrics: The metrics considered for analyzing the performance of the ability based transcoding scheme are a) Accept Rate b) Total Accessibility and c) Visual Accessibility. The accept rate determines the total number of users admitted into the system during the period of simulation. The Total accessibility defines the total accessibility achieved for the information delivered to the user, which can be a summation of the calculated accessibility ratio values for the different modalities presented to the user. The visual accessibility determines the accessibility of the transcoded information as presented to the user with specific vision impairment.

5.1 PERFORMANCE OF THE ABILITY DRIVEN TRANSCODING SCHEME

We study the performance of the proposed algorithm under varying bandwidth restrictions and the metric of comparison is accept rate. We compare the performance of the following three cases: a) when the information is sent without any transcoding b) where the best accessible information is sent (e.g., for a blind person we send audio, captioned video, and tactile text) and c) when the resource aware ability based transcoding is performed. The relative performances of these techniques are given by the graph in Figure 4. Note that this result is based on the general user population based on the demographics as previously discussed.

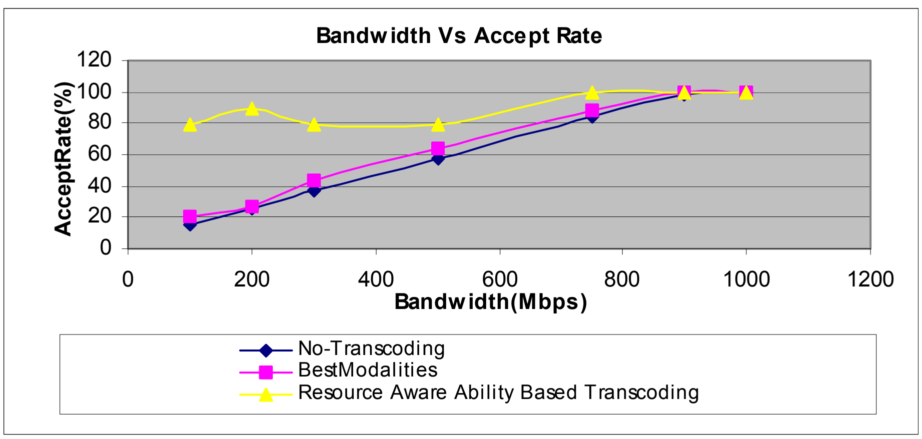

Figure 4: Graph showing acceptance rate for different bandwidths

The results illustrate that by considering accessibility information alone (given by the Best Modalities policy) we can improve the overall bandwidth utilization; hence, the best modalities policy performs marginally better than the case with no transcoding with respect to the acceptance rate. On the other hand, the resource aware ability based transcoding algorithm significantly outperforms the other strategies, especially under low bandwidth conditions. The irregularity in the curve is due to the random arrival of user groups, and the fact that the accessibility of information is not forsaken while considering the resource restrictions. Note that, if the accessibility drops below a minimum defined value, the request is rejected. However this result is not indicative of the improvements in the accessibility to the users. The following set of experiments indicates a qualitative measure of improvement in accessibility when using the resource aware ability based transcoding algorithm.

Effects of Load on Accessibility: In this section we study how the accessibility is affected as the number of requests in the system scales up. The set of graphs in figure 5 show the accessibility values, plotted as a function of user requests. We can see from the graphs, that, the accessibility is very low when information is not transcoded, and drops to zero soon, due to the resource intensive nature of information provided. On the other hand, due to resource aware transcoding the accessibility is maximized, and the maximum number of users that receive acceptable service is significantly increased. When total accessibility value of the modalities falls below the threshold P_{ij}^{min}, which varies for different user groups, the request is rejected. Hence we can see that the proposed algorithm tries to maintain accessibility even under heavily loaded conditions, while also reducing the number of rejects. We consider each disability category separately to analyze how the accessibility is affected due to load. The maximum accessibility attainable from given information is 3.0 in our experimental case.

- ***Accessibility for Blind Users:*** From the graph in Figure 5 we can see that by sending the original data without transcoding, the maximum total accessibility achieved is just a little over 2.0. On the other hand by transcoding we can reach a maximum accessibility of nearly 3. The lower resource requirement of the transcoded information (containing audio, and tactile captioning and summary) also plays a pivotal role in maintaining this accessibility, even if the load is high. After a point the accessibility suddenly drops to zero (indicating rejects), attributed to the limited number of alternate modalities available like captioned video and summaries that are accessible through only tactile means. Synthesized audio format of the visual data, and alternate choices for the audio (summarized audio) will increase choices for the transcoder, hence the reduction in accessibility will be more gradual.
- ***Accessibility for Vision Impaired:*** Although the accessibility follows a similar pattern as that for blind users, there are some minor differences to be noted. The maximum accessibility attained after transcoding is around 2.45. The accessibility of the information can be improved by using specialized techniques as described in the Section 4.
- ***Accessibility for Deaf and Hard of Hearing:*** The accessibility for both these classes of users exhibits a very similar trend, i.e., both these curves remain constant with high accessibility values. One reason might be due to the fact that most of these alternate modalities are visual in nature, which can be easily accessed by these users. Specialized transcoding techniques, as those for vision impairments however will improve the accessibility of the audio information, especially for those with low and moderate degree of hearing impairments. The drop in accessibility is more gradual than that for the blind. The reason is due the fact that the accessibility of the alternate modalities like text summaries is perfectly accessible for the user with hearing impairments. However for the blind these modalities have to be tactile in nature, which limits the options for selection of alternate modalities.
- ***Accessibility for Motor Impaired:*** We can see from the graph, that the originally sent modalities are almost as accessible as the transcoded information. This is because we have considered information access, which involves mainly perceptual abilities. Additionally modalities being highly textual in nature are less accessible compared to audio or video, which can be accessed without much motor involvement (assuming mot everyone have end devices to access information). These factors contribute to the drop in accessibility even at lower load levels.
- ***Accessibility for deaf-blind:*** The only information accessible is tactile. As we send only tactile text, maximum accessibility is maintained throughout constantly at any load. If we consider tactile images, we expect to see slight change in the trend followed, but this too depends on the available bandwidth and resources available. Generally the accessibility should not change much due to moderate loads, due to very little resources required.

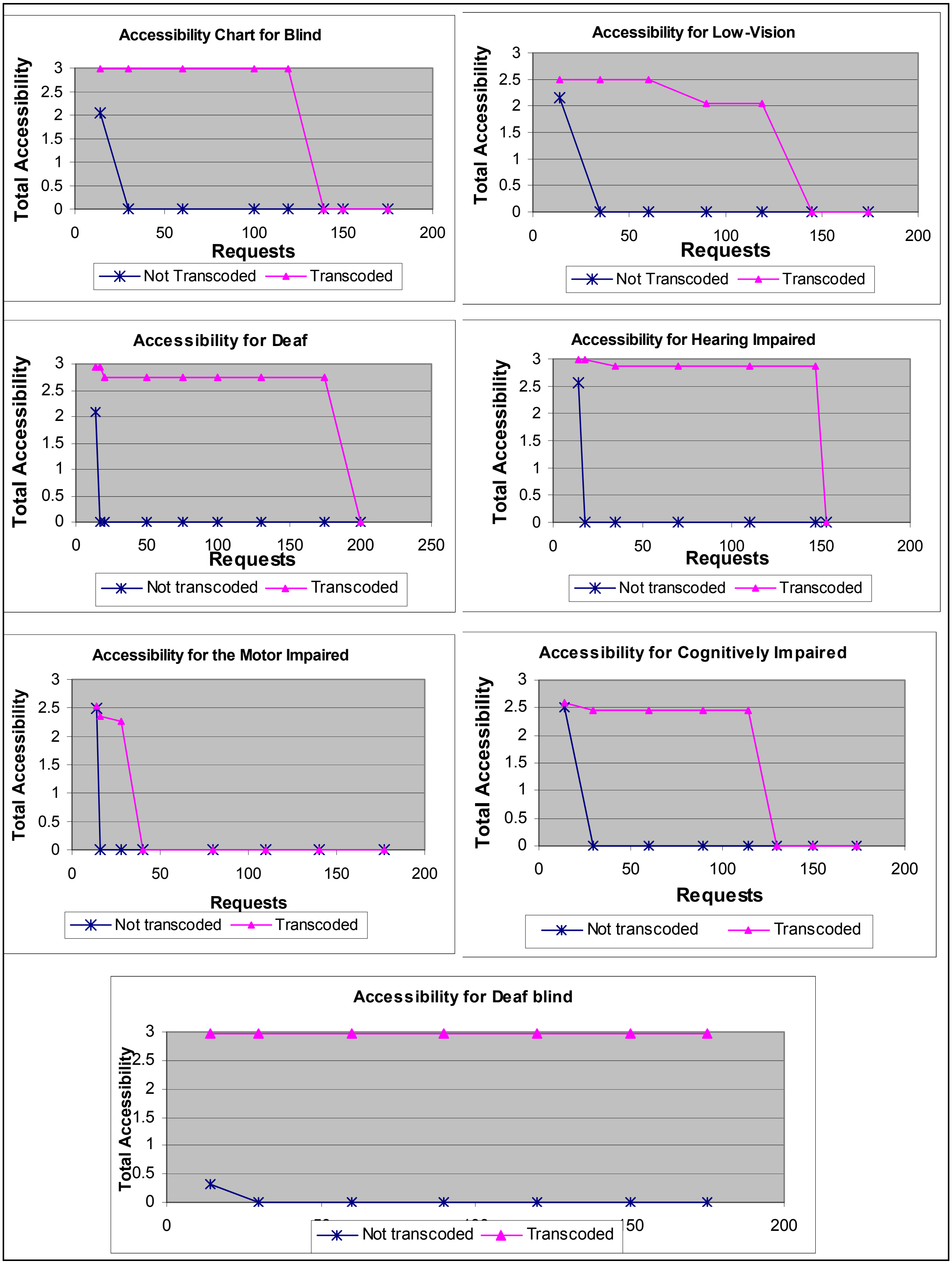

Figure 5: Effect of Load on Accessibility for the different impairment classes

- ***Accessibility for Cognitively Impaired:*** The accessibility here depends mainly on the type of data provided as an alternative. With the minimal understanding present we provide text summary and video summaries. But more detailed understanding of the cognitive impairments will give better transcoding techniques.

5.2 Results of the Specialized Transcoding for the users with Vision Impairments

In this section we study the effect of the specialized transcoding for vision impairments. Specifically, we illustrate the effect of transcoding on refractive errors (myopia, far vision, presbiopia) and visual field errors (e.g., glaucoma). We compare the visual accessibility values for varying degrees of impairments under 3 scenarios – no transcoding, sufficient transcoding and limited specialized transcoding. The sufficient level of transcoding for a specific user and a specific media-type is calculated as the accessible transcoding level for normal viewing with the best resource utilization. The specialized transcoding process is limited by the number of quality levels available in the content source – the transcoding technique picks the nearest feasible approximation to the sufficient quality desired. A visual accessibility value higher than 1, implies that the delivered quality is better than required for useful assimilation of the information. The increased quality may improve the visual accessibility (as calculated); however the price of the enhanced quality is additional resource usage.

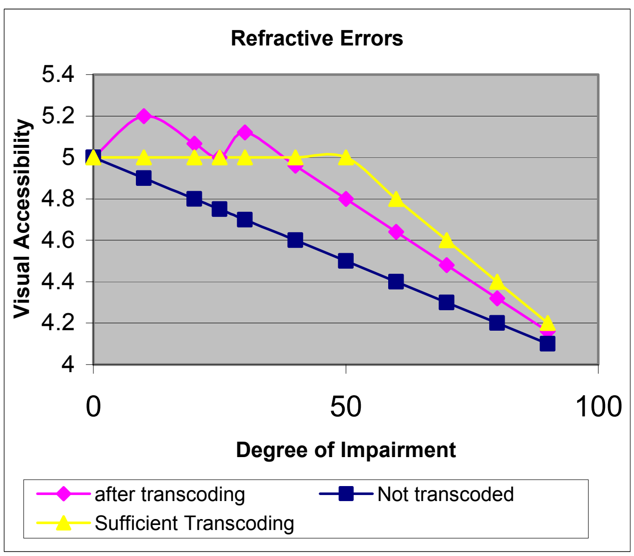

Figure 6: Effect of transcoding for refractive errors

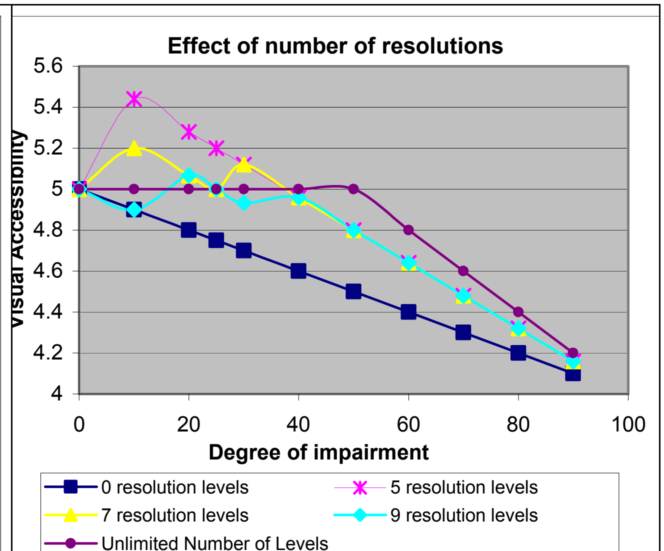

Figure 7: Effect of resolutions on visual accessibility

From Figure 6, we observe that as the degree of impairment due to refractive errors increases, the visual accessibility decreases dramatically in the untranscoded case. Note that the visual accessibility, especially for a medium level impairment, improves significantly as a result of transcoding. For example, we assume the maximum magnification attainable is limited by the screen size. Hence even in the sufficient case no more magnification is possible beyond the screen height. With higher resolutions, the visual accessibility may increase beyond the sufficient value (with higher resource consumption); however, the improvement in the user's visual experience is not clear. The second graph (Figure 7) shows the effect of increasing the available number of resolution levels in the data set. We see that as we increase the number of resolution levels the curve smoothens and gets closer to the sufficient case. When the number of levels is unlimited we can pick the required format for normal viewing without increasing resource consumption. However it is not realistic to have that many formats stored. Our experiments indicate that about 10 resolution levels are sufficient to obtain near ideal performance, which provides sufficient accessibility without compromising on resource usage.

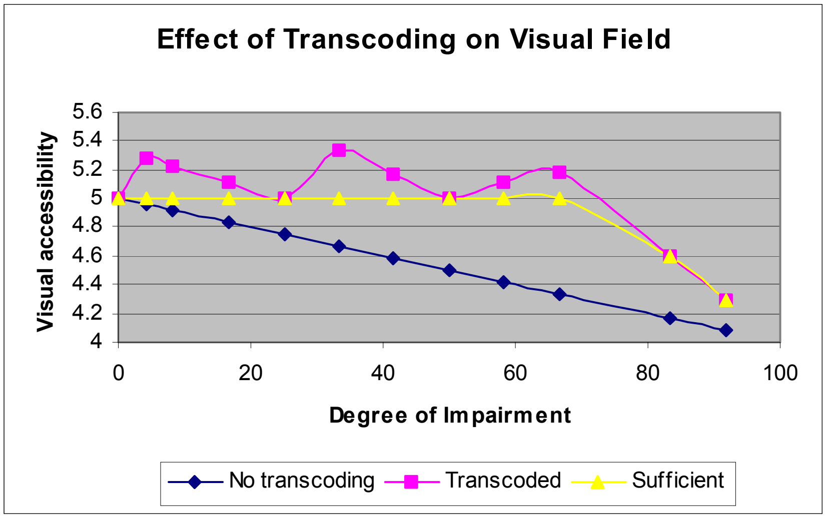

Figure 8: Graph showing the effect of transcoding on impairments affecting visual field.

Figure 8 shows the effect of visual field problems on accessibility. As the horizontal field of view decreases (as is the case in glaucoma patients), the area of the image seen is reduced. As a result, the size has to be reduced to fit the field of view. The effect of transcoding is similar to the refractive case. Here, we specify a minimum size, beyond which the image cannot be shrunk, keeping in view the effect of visual acuity.[1]

5.3. A Prototype Distance Learning Portal

In this section we will describe a distance-learning environment currently being developed under the SUGA project at UCI, that supports cross disability access. An initial prototype of such a distance-learning portal is generated using MPEG-4 specifications and SMIL[LM01,SM02,MP02]. The distance-learning environment is under development and the details of the actual framework are beyond scope of this paper. In the sample layout (See Figure 9) the distance learning session consists of the video session capturing the lecture, a slide session possibly showing animations or relevant images and the main course content immediately following. There are also panels for the navigating to different chapters, the video controls and also a search panel to search for different topics in the lesson.

[1] Very small images tend to impact the visual acuity of the user, as the visual acuity is dependent on the size.

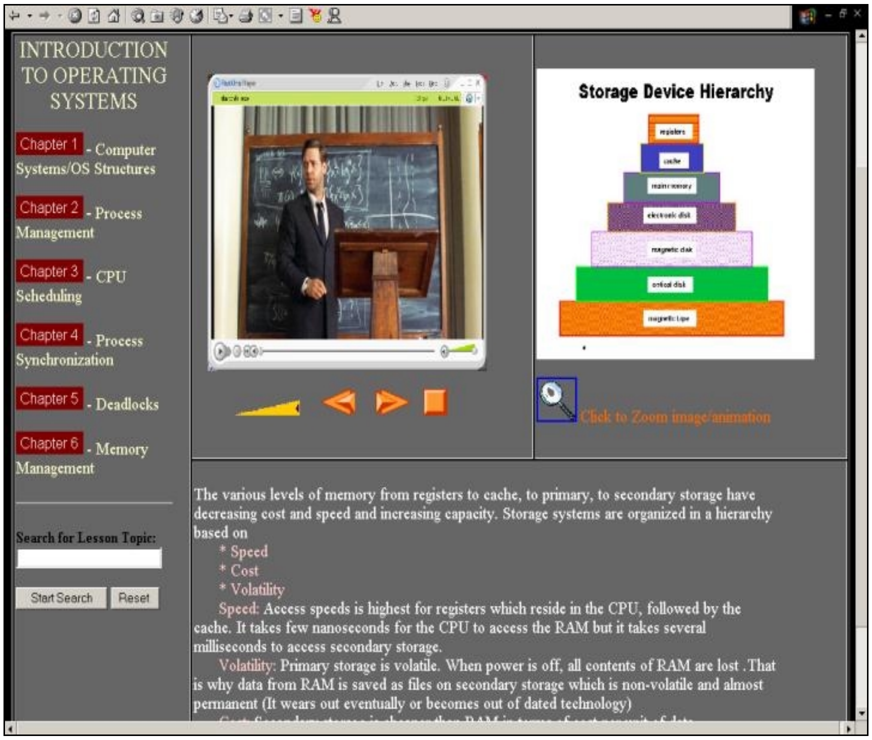

Figure 9: Sample screenshot of the distance-learning portal

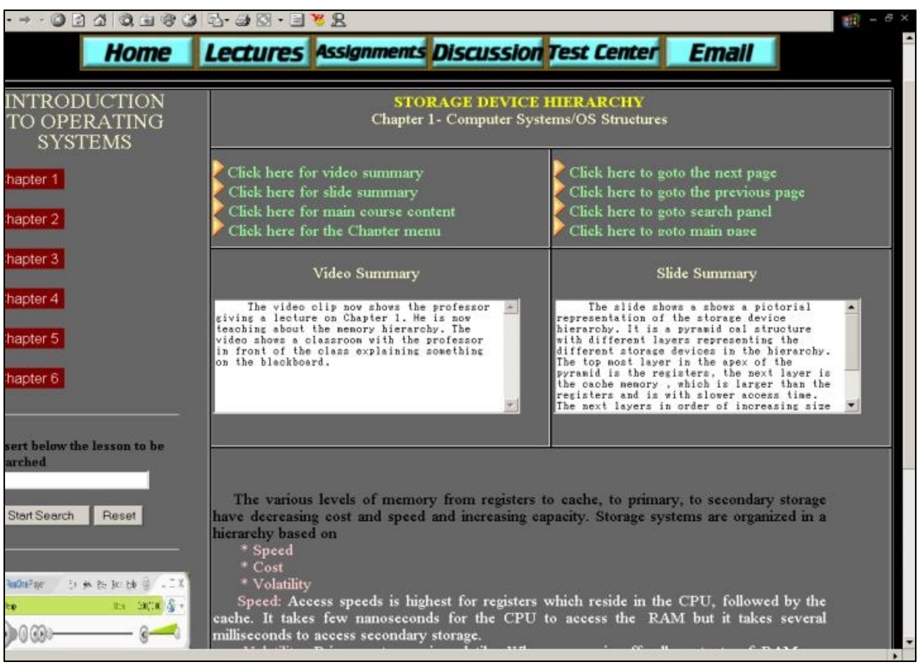

Figure 10: Transcoded portal for a user who is blind.

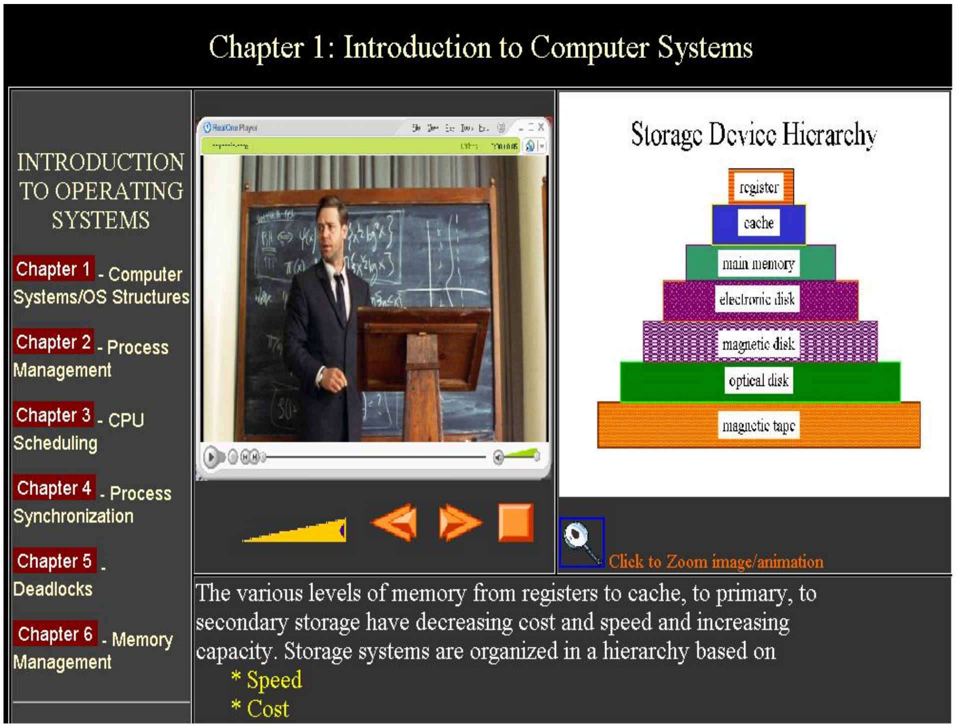

Figure 11: Transcoded portal for a person with a visual acuity 20/30

Figure 10 shows how the same portal looks when it adapted for a blind user. Since both the video and slides are not viewable, text summaries are given as alternative modality. Textual links to different parts of the page are given, so that they can be accessed through tactile devices or screen readers. There is also an audio panel for controlling the audio parameters. Thus the portal is suitably adapted for a blind user to access it with ease. When adapting such environments interface issues play a huge role [K98,JV01,SS01]; User Interface details are outside the scope of this paper.

For a user with refractive errors like myopia, hyperopia and presbyopia, the visual acuity is affected. To increase the accessibility magnification is necessary. Such magnification may require reorganization of the portal to accommodate the possibility of magnified parts affecting other media types in the portal. Figure 11 shows how the distance-learning portal is adapted to a user with a corrected visual accessibility of 20/30. Appropriate magnification of the video clip, image, and font are applied. Note that the background contrast is increased to improve the visibility of the white text against the background color. Magnifications are also applied to specific parts of the image clip (text labels in the image), video controls and the navigation buttons.

6. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this paper we discussed the ability based transcoding technique to improve accessibility of information to a disabled user. Our algorithm combines studies from the fields of cognitive sciences and multimedia to effectively develop a mechanism to select the modalities that are accessible to the user with the knowledge of the user's functional abilities. Our analysis shows that the proposed transcoding techniques provide increased accessibility to multimedia information. In addition, the results show that there is an improvement in resource utilization as a result of the adaptation. We also demonstrate how specialized transcoding for vision impairments can help improve upon the accessibility as provided by the basic transcoding. Such fine-grained adaptations for other categories of impairments like hearing impairments, motor impairments etc will further improve accessibility. We are currently working on specialized transcoding schemes for these other impairments using specific models of the human sensory, motor and cognitive systems.

Several challenges remain to be addressed in providing universal access to multimedia information. We are currently developing adaptive transcoding techniques to allow for dynamic reallocation of resources to ongoing streams to admit more users by allowing for simultaneous transcoding and load balancing. Other research topics currently being explored include ability-based transcoding in the presence of proxy servers in the network and dynamic creation of transcoded content at the server and proxies. A future challenge is to address the issue of cross disability access in highly dynamic mobile environments.

Advances in universal accessibility can dramatically impact the lives of the disabled by making information access easier, thereby simplifying many aspects of social and professional life, including communication and education. Our vision is to provide accessible information and enable communication across diverse user populations breaking the barriers of age, ability, language and literacy without the need for sophisticated and expensive technology. Such a human centric approach can achieve high quality of service. We expect that the middleware services for information management and distribution will provide cost-effective solutions to address this goal.

REFERENCES

[AA96] K.Almeroth and M. Ammar, "The use of multicast delivery to provide a scalable and interactive video-on-demand service", IEEE Journal on Selected Areas in Communication, 1996

[ACI99] ACI HEARING CENTER, Signs of Hearing Loss, http://www.aci-hearing.com//losses

[AF98] S. A. Su, R. K. Furuta, "VRML-based Representations of ASL Fingerspelling", ASSETS'98

[AM00] S.C.Arnold, L.Mark and J.Goldthwaite, "Programming by Voice, VocalProgramming" ASSETS'00

[AO02] American Optometric Association, http://www.aoanet.org, 2002

[AO92] D. Anderson, Y. Osawa and R.Govindan, "A file system for continuous media", ACM Transactions on Computer systems, 10(4): 311-337, Nov 1992.

[AY96] J.W.C.C.Aggarwal and P.S. Yu, "A permutation based pyramid broadcasting scheme for video-on-demand systems", IEEE Multimedia Computing Systems, 1996.

[B98] J;Baca, "Comparing Effects of navigational Interface Modalities On Speaker Prosodics", ASSETS'98

[B99] D.Bauer, "Assistive Technology for severely hearing impaired: Towards closing essential gaps", AAATE 99

[BJ99] A.C.Balcells, M.A.Jodar, et al, "A system for the analysis and scanning of tremor on handicapped people", AAATE 99

[BP94] M.M. Buddhikot, G.M. Parulkar, J.R. Cox, Jr., "Design of a Large Scale Multimedia Storage Server", Journal of Computer Networks & ISDN Systems, PP504-523, 1994

[BP95] Bisdikian,Patel,"Issues on movie allocation in distributed video-on-demand systems", IEEE Conference on Communication, 1995

[BS00] A.Brinkman, K.Salzwedel, C.Scheideler, "Efficient, Distributed Data Placement Strategies for Storage Area Networks", Twelfth annual ACM symposium on Parallel algorithms and architectures, 2000

[BV03] Vidhya Balasubramanian, Nalini Venkatasubramanian, "Server Transcoding of Multimedia Data for Cross-Disability Access", to appear in MMCN 2003 Conference.
[CCRMA] The Stanford University Center for Computer Research in Music and Acoustics, http://www-ccrma.stanford.edu
[CD98] E.Cole, P.Dehdashti, "Computer-based cognitive prosthetics: assistive technology for the treatment of cognitive disabilities"ASSETS'98
[CG98] C. Chou, L. Golubchik, and J. Lui, "A performance study of dynamic replication techniques in continuous media servers", Technical Report, University of Maryland, 1998
[CH93] Charles P.Halsted, " Brightness, Luminance and Confusion", Information Display, 1993
[CPA] Canadian Psychological Association, http://www.cpa.ca/factsheets/cognitive.htm
[CS02] ContrastSensitivity.net, http://www.contrastsensitivity.net, 2002
[CUD] CENTER FOR UNIVERSAL DESIGN , NCSU , http://design.ncsu.edu
[DK95] A. Dan, M. Kienzle, and D. Sitaram, "Dyanamic Policy of segment replication for load-balancing in video-on-demand servers", ACM Multimedia Systems, 3(3): 93-103, 1995
[DS95] A. Dan and D. Sitaram , "An online video placement policy based on bandwidth to space ratio (bsr)", SIGMOD'95, pages 376-385
[EV01] D.Eager, M.Vernon, J.Zahorjan, "Minimizing Bandwidth Requirements for On-Demand Data Delivery", IEEE Transactions on Knowledge and Data Engineering 2001, pp.742-757
[EVV] "Eye , Vision and Visibility", http: // www.saud.ku.edu/book/eye
[F97] Geoff Freed, "The Web Access Project – Access to Multimedia on the Web", CSUN 97
[F98] R.A.Foulds, A.W.Joyce III, "Expanded Interactions: Broadening Human- Centered Computing", ASSETS'98
[FK99] J.Fink, A.Kobsa and A.Nill, Adaptable and adaptive Information Provision for All Users, Including Disabled and Elderly People",webnet 99, www.iroe.fi.cnr.it/avanti/contents/contents/publ3.htm
[FQ84] E. A. Fleishman and M.K. Quaintance, "Taxonomies of human Performance: The description of Human Tasks", 1984
[G00] A.S.Gilman, Universal Design and the Grid, (Working draft for **Grid** Forum 5, Oct. 2000), http://trace.wisc.edu
[G95] H.W. Gellersen, "Modality Abstraction: capturing Logical Interaction Design from „ User Interfaces for All", UI4ALL-95
[GC92] J. Gemmell and S. Christodoulakis, "Principles of delay sensitive multimedia data storage and retrieval", ACM Transactions on Information Systems, 10(1): 51-90, 1992
[GK98] L. Gao, J. Kurose, and D. Towsley , "Efficient Schemes for broadcasting popular videos", NOSSDAV'98
[GW95] George M.Williams Jr, "Image Quality", Tutorial, http://www.pvinc.com/tutorial, 1995
[HH99] Harborg, Trym Holter, et al, "Subtitling of live broadcast TV-programs for the hearing impaired", AAATE 99
[HL01] Faustina Hwang, Patrick Langdon, Simeon Keates, John Clarkson, "Haptic Assistance to Improve Computer Access to Motion-Impaired Users", Eurohaptics2001
[HS00] Anita W.Huang and Neel Sundaresan, "A semantic transcoding System to Adapt Web Services for Users with Disabilities", ASSETS'00
[HS01] Richard Han, John R. Smith , "Transcoding of the Internet's Multimedia Content for Universal Access", Multimedia Communications : Directions and Innovations , Chapter 15,2001
[HS97] K.Hua and S.Sheu, "Skyscraper broadcasting: A new broadcasting scheme for metropolitan video-on-demand systems", SIGCOMM'97
[IN] INCLUDE Homepage, http://www.stakes.fi/include
[IP02] Image Processing Tool Box, http://www.mathworks.com/access , 2002
[IV01] S.Irani, and Nalini Venkatasubramanian, "Semi-Continuous Transmission for Cluster-Based Video Servers", IEEE International Conference on Cluster Computing 2001 (Cluster'2001)
[J98] Frankie James, "Lessons from Developing Audio HTML Interfaces", ASSETS'98

[Jb98] R.J.K.Jacob, "Eye Tracking in Advanced Interface Design", in Virtual Environments and Advanced Interface Design, pp. 258-288, 1995.
[JV01] J.A.Jacko and H.S.Vitense , "A review and reappraisal of information technologies within a conceptual framework for individuals with disabilities", UAIS 2001
[K'98] M.Kurze, "Tguide: A Guidance System for Tactile Image Exploration", ASSETS'98
[K98] S. Keates, P.J. Clarkson and P. Robinson, "Developing a methodology for the design of accessible interfaces", ERCIM'98
[KK93] K.Keaton and R. Katz, "The evaluation of video layout strategies on a high-bandwidth file server", NOSSDAV, 1993
[KR98] Simeon Keates and Peter Robinson , "The use of gestures in multimodal input", ASSETS'98
[L98] Jill Fain Lehman, "Toward the use of Speech and Natural language Technology in Intervention for a Language-disordered Population", ASSETS'98
[LH02] Lighthouse International, http://www.lighthouse.org, 2002
[LL99] C.Lee, J.Lehoczky, D.Siewiorik, R Rajkumar, J Hansen, "A Scalable Solution to the Multi-Resource QoS Problem", IEEE Real-Time Systems Symposium, 1999
[LM96] Lila F.Laux, Peter R.McNally, Micheal G.Paciello, and Gregg C. Vanderheiden , "Designing the World Wide Web for People with Disabilities : A User Centered Design Approach", ASSETS '96
[LM01] Liang Cheng, Magda El Zarki, "The Analysis of MPEG-4 Core Profile and its System Design", MTAC, 2001
[LS93] P.Lougher and D.Shepherd, "The design of a storage server for continuous media", The Computer Journal- Special Issue on Distributed Multimedia Systems, 36(1): 32-42, Feb 1993
[MC92] Blair MacIntyre and William B. Cowan, " A Practical Approach to Calculating Luminance Contrast on a CRT", ACM Transaction on Graphics 11(4): 336-347, October 1992.
[MP98] Sarah Morley, Helen Petrie, Anne-Marie O'Neil, Peter McNally, "Auditory Navigation in Hyperspace: Design and Evaluation of Non-Visual Hypermedia System for Blind Users", ASSETS'98
[MP02] http://www.mpeg4.net , 2002
[MS99] R.Mohan, J.Smith and C.Li, "Adapting Multimedia Internet Content for universal Access", IEEE Transactions on Multimedia, vol.1, march 1999, pp. 104-114
[MW97] Gin-Kou Ma, Chiung-Shien Wu, Mei-Chian Liu, Bao-Shuh P.Lin, "Efficient Real-time Data Retrieval Through Scalable Multimedia Storage", ACM Multimedia 97
[OY97] Yoshihiko Okada, Katsuhiro Yamanaka, Akio Kaneyoshi and Osamu Iseki, "Countervision: A Screen Reader with Multi-Access Interface for GUI", CSUN 97
[P97] Martin Prime, "Adapting multimedia information for internationalization and for users with disabilities", 3rd ERCIM Workshop, User Interfaces For all, 1997.
[P98] G.Pirelli, "Giving a VOICE to the deaf by developing awareness of VOICE to text recognition capabilities", TIDE-98
[R01] T.V.Raman, "Emacspeak- Toward The Speech-enabled Semantic WWW", www.cs.cornell.edu/Info/People/raman/publications, March 2001
[RE02] MEDLINEplus: Refractive Errors, http://www.nlm.nih.gov/medlineplus/refractiveerrors.html, 2002
[RG00] M.Rowan, P.Gregor, D.Sloan, P.Booth, "Evaluating Web Resources for Disability Access", ASSETS'00
[RI02] Retina International, http://www.retina-international.org/, 2002
[RL01] Robin Leonard , "Statistics on Vision Impairment" ,A Resource Manual, 2001
[RL97] R.Rajkumar, C.Lee, J.Lehoczky, D.Siewiorek, " A QoS Based Resource Allocation Model", IEEE Real-Time Systems Symposium, 1997
[RL98] R.Rajkumar, C.Lee, J.P.Lehokzky, D.Siewiorek, " Practical Solutions for QoS based Resource Allocation Problems" , IEEE Real-Time Systems Symposium, 1998
[RM96] C.Ramstein, O.Martial, A.Dufresne, M.Carignan, P.Chasse, P.Mabilleau, "Touching and Hearing GUI's: Design Issues for the PC-Access System", ASSETS'96
[RN02] Royal National Institute for the Blind, UK, http://www.rnib.org, 2002

[RR98] D. S.Ray, E.J.Ray, “Adaptive Technologies for the Visually Impaired: The Role of Technical Communicators”, Journal of the Society for Technical communication” , 1998
[SAP] http://www.sapdesignguild.org/resources
[SCI] http://www.spinalcord.org
[SDD96] A.Smith, J.Dunaway, P.Demasco, D.Peischl, “Multimodal input for Computer Access and Augmentative Communication”, ASSETS’96
[SEV00] S.Chandra, C.S.Ellis and A.Vahdat, “Differentiated Multimedia Web Services Using Quality Aware Transcoding” INFOCOM 2000
[SEV99] S.Chandra, C.S.Ellis and A.Vahdat, “Multimedia Web Services for Mobile Clients Using Quality Aware Transcoding”, WoWMoM’99
[SGS93] C. Stephanidis, P. Galatis, M. Sfyrakis, and G. Homatas, “A methodology for selecting interaction techniques for users with disabilities in the context of the B-ISDN environment”, FORTH-ICS’93
[SGV95] Prashant J.Shenoy, Pawan Goyal and Harrick M. Vin, “Issues in Multimedia Server Design”, ACM Computing Surveys, Vol.27, No. 4 (Dec. 1995), pp. 636-639
[SJ98] Richard S.Schwerdtfeger and Phillip D. Jenkins, “100% Pure Java ™ : IBM’s focus to Develop Accessibility for the Next Generation”, CSUN’98
[SM02] Synchronized Multimedia, http://www.w3.org/AudioVideo/ , 2002
[SS01] Constantine Stephanidis and Anthony Savidis, “Universal Access in the Information society: Methods, tools, and Interaction Technologies”, UAIS 2001
[SS99] S.Sahu, P. Shenoy, D. Towsley, “Design considerations for integrated proxy servers”, NOSSDAV’99
[SW98] Stefan Winkler, “A Perceptual Distortion Metric for Digital Color Video”, in *Proc.*5th International Conference on Image Processing, vol. 3, pp. 399-403, Chicago, Illinois, October 4-7, 1998.
[SW99] Stefan Winkler “Issues in Vision Modeling for Perceptual Video Quality Assessment” Signal Processing, vol. 78, no. 2, pp. 231-252, October 1999
[T94] Jim Thatcher, “Screen Reader/2 Access to OS/2 and the Graphical User Interface”, ASSETS’94
[TA00] H.Tagaki and Chieko Asakawa, “Transcoding Proxy for Nonvisual Web Access”, ASSETS’00
[TM96] Renu Tewari, Rajat Mukherjee, Daniel M. Dias, Harrick M. Vin, “Design and Performance Tradeoffs in Clustered Video Servers”, ICMCS '96
[TP93] F. Tobagi, J. Pang, R. Baird, and M. Gang, “Streaming RAID: A disk storage system for video and audio files”, ACM Multimedia’93
[V98] Gregg C.Vanderheiden, “EZ Access Strategies for Cross-Disability Access to Kiosks, Telephones and VCRs”, TIDE 98
[VC02-1] Vision Channel, http://www.visionchannel.net, 2002
[VC02-2] Visicheck , http://www.vischeck.com, 2002
[VF97] Oliver Verscheure, Pascal Frossard, “Perceptual MPEG-2 Syntactic Information Coding: A performance study based on Psychophysics”, in Picture Coding Symposium, 1997
[VG94] Harrick M. Vin, Alok Goyal, Pawan Goyal, “Algorithms for Designing Multimedia Servers”, ICMCS’94 pages 234-243
[VR93] H. M. Vin, P. V. Rangan, “Designing a multi-user HDTV storage server”, IEEE Journal on Selected Areas in Communications, 11(1): 153-164, Jan-1993
[VR97] Nalini Venkatasubramanian and Srinivas Ramanathan, “Load Management in Distributed Video Servers”, Proceedings of the IEEE Intl. Conference on Distributed Computing Systems (ICDCS '97), May 1997.
[VW01] A.Vetro, Y.Wang, and H.Sun, “Object-Based Transcoding for adaptable Video Content Delivery”, IEEE, transactions on Circuits and Systems for Video Content Delivery, Vol 11, No.3, March 2001
[W52] David Wechsler, “Range of human Capacities”, 1952
[W98] Gerhard Weber, “Programming for Usability in Nonvisual User Interfaces”, ASSETS’98

[WD99] Douglas Wakefield and Judith M. Dixon, "How Blind Persons Can Access the Audio and Video Formats of the WWW: I-wave, Mpeg, Jpeg, Gif—Potholes or Billboards", CSUN 99
[WH01] Jacob T. Wilensky, Anjali Hawkins, "Comparison of Contrast Sensitivity, Visual Acuity and Humphrey Visual Field Testing in Patients with Glaucoma", Tr. Am. Ophth. Soc. Vol. 99, 2001
[WHO1] http://esa.un-org/unsd/disability/
[WHO2] http://www.who.int/whosis
[WHO3] http://www.who.int/pbd/pbl/data.htm
[WY95] J.L Wolf, P.S Yu, and H. Shachnai, "Dasd dancing: A disk load balancing optimizing scheme for video-on-demand computer systems", ACM SIGMETRICS'95, pages 157-166
[YC92] P. Yu, M. Chen, and D. Kandlur, "Design and analysis of a group sweeping scheme for multimedia storage management ", Proceedings of Third International Workshop on Network and Operating System Support for Digital Audio and Video, San Diego, pages 38-49, Nov 1992
[ZM99] Yueting Zhuang, Sharad Mehrotra, and Thomas S. Huang, "A Multimedia Information Retrieval Model Based on Semantic and Visual Content", Proceedings of the 5th International ICYCS Conference, 1999.
[ZV01] Gottfried Zimmermann and Gregg Vanderheiden, "Modality Translation Services on demand Making the World More Accessible For All", RESNA 2001
[SUGA02] The SUGA Project, http://www.ics.uci.edu/~dsm/suga/, 2002