

1 Introdução

O rápido crescimento das tecnologias relacionadas à Internet tem proporcionado o oferecimento de serviços cada vez mais elaborados aos usuários. Com o aumento da banda passante disponível na rede, através de tecnologias de acesso em banda larga tais como xDSL, rádio ponto-multiponto, *cable modem*, *Fast Ethernet* e *Gigabit Ethernet*, e a adoção de tecnologias de compactação e compressão de mídias contínuas, as aplicações puderam se beneficiar de diferentes tipos de conteúdo, tais como texto, imagens, sons e vídeos, na composição de documentos multimídia. Quando tais documentos agregam a possibilidade de interação com seus usuários, denominados documentos hipermídia, suas expressividades são ainda maiores.

Para possibilitar a exibição não linear e a interatividade, os documentos hipermídia estabelecem diversos relacionamentos entre os objetos de mídia que o compõem, tais como os relacionamentos espaciais, estruturais e temporais. Uma vez que os relacionamentos espaciais e estruturais não fazem parte do escopo deste trabalho, deste ponto em diante apenas os relacionamentos temporais serão referenciados.

Um ambiente de execução de documentos hipermídia pode ser representado como mostra a Figura 1. Nessa representação simplificada, o ambiente disponibiliza uma API para a interação com o usuário, assim como outra para a obtenção de objetos, que podem estar armazenados tanto localmente quanto em recursos remotos ou podem ser gerados simultaneamente à sua apresentação. A orquestração da apresentação de um documento hipermídia é realizada pelo elemento formatador, responsável por coordenar e assegurar a correta exibição dos vários objetos através do envio de informações de controle ao conjunto de exibidores (Rodrigues & Soares, 2003).

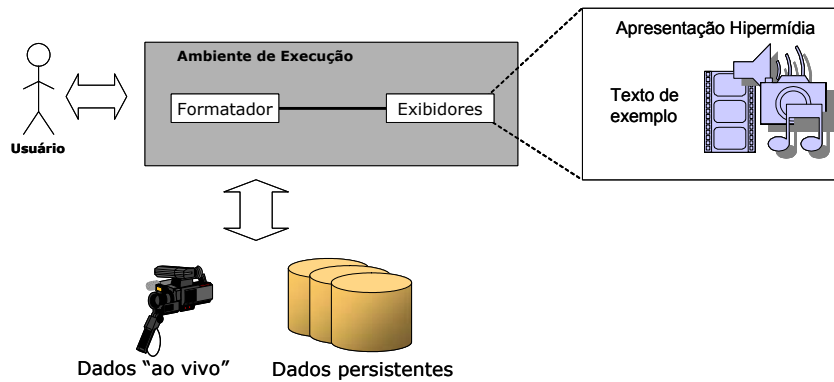


Figura 1: Ambiente de Execução Hipermedia.

Um dos aspectos críticos na apresentação de documentos hipermedia é a manutenção da sincronização intra e interobjetos. Cabe ao formatador tal controle, a detecção de problemas e a determinação das ações necessárias à garantia da sincronização.

A sincronização intra-objeto, ou intramídia, garante a continuidade e qualidade da apresentação de uma determinada mídia. Ela refere-se à manutenção do relacionamento temporal entre diferentes unidades de informação de um mesmo objeto de mídia. No caso de um vídeo, por exemplo, a sincronização intramídia refere-se à apresentação contínua dos vários quadros. Problemas na manutenção da sincronização intramídia em um vídeo podem causar efeitos como interrupções e cintilação durante sua apresentação, degradando sua qualidade. No caso de sinais de áudio, a sincronização intramídia previne a inserção de períodos de silêncio ou truncagem nos fonemas, ocasionando quebra na continuidade do sinal.

A sincronização interobjeto, ou intermídia, refere-se ao controle correlacionado do instante de exibição de cada unidade de informação dos vários objetos. Um exemplo de sincronização entre objetos é a garantia de simultaneidade no início da apresentação de um objeto de áudio e de um objeto de vídeo: a voz dos personagens de um filme tem que estar sincronizada com o quadro sendo exibido.

Embora a manutenção do sincronismo temporal possa ser realizada em tempo de compilação do documento, a fim de manter os relacionamentos de sincronismo especificados pelo autor, a qualidade da apresentação percebida pelo usuário depende da manutenção da sincronização em tempo de execução, ou apresentação, de forma a compensar possíveis descasamentos temporais

produzidos pelos ambientes de transmissão e execução, como atrasos na rede e nos sistemas operacionais, entre outros.

Os diversos fatores que podem ocasionar problemas na sincronização entre objetos de mídia podem ser entendidos como um problema de provisão de qualidade de serviço, que impõe a necessidade de atendimento aos requisitos de maneira fim a fim. Em relação à obtenção das unidades de informação de cada objeto no instante correto, podem-se utilizar mecanismos de *prefetch* de dados.

Uma das técnicas utilizadas para a manutenção da sincronização temporal é o ajuste elástico da duração de exibição de cada objeto, ou simplesmente ajuste elástico. Na literatura, o ajuste elástico (*elastic time adjustment*) também é chamado de alteração de velocidade de exibição (*timescale*), dilatação do tempo de exibição (*playback dilation*), controle da velocidade de exibição (*speed control*) e alongamento do tempo (*time stretch*).

Para a manutenção da sincronização temporal, o formatador pode utilizar alguma métrica para determinar qual o tipo de ajuste mais adequado e quais objetos devem ter sua duração de exibição alterada (Bachelet et al, 2004). Em relação aos objetos de vídeo, foco desta dissertação, o ajuste elástico deve prover um controle suave a fim de causar o mínimo impacto possível na percepção dos usuários.

Os efeitos do ajuste elástico no tempo de exibição de mídias contínuas dependem do tipo do dado audiovisual sobre o qual a adaptação é executada. De acordo com os testes realizados em (Omoigui et al., 1999), um aumento de 40% na velocidade de exibição de um vídeo ainda permite uma boa compreensão do conteúdo. Por outro lado, um aumento na taxa de exibição de um áudio acima de 25% deteriora sensivelmente a qualidade do sinal (Arons,1992). Segundo (Golubchik et al., 1995), evidências mostram que alterações de $\pm 5\%$ da taxa de exibição nominal de vídeos não são perceptíveis.

A qualidade da mídia audiovisual também é afetada pela perda de sincronização entre os fluxos de vídeo e áudio. Diferenças superiores a 160 ms já são perceptíveis aos usuários (Aly & Youssef, 2002).

Com o intuito de explicitar os problemas envolvidos no ajuste elástico do tempo de exibição de vídeos, a Seção 1.1 aborda o assunto como um caso particular de adaptação de conteúdo. A Seção 1.2 resume os aspectos importantes para o ajuste elástico em fluxos de vídeo MPEG-2 e descreve os objetivos desta

dissertação. A Seção 1.3 apresenta sucintamente os assuntos tratados nos capítulos posteriores.

1.1. Adaptação de conteúdo

O ajuste elástico do tempo de exibição de fluxos contínuos pode ser entendido como uma forma de adaptação de conteúdo, que se apresenta como uma ferramenta para ajudar na provisão de aplicações com qualidade adequada ao usuário, e pode ser realizada considerando-se o estado de cada componente de todo o sistema de provisão de serviço, desde o provedor, ou transmissor, até o usuário, ou receptor. A adaptação de conteúdo pode prover, também, melhor reuso dos diversos conteúdos, onde um mesmo conteúdo original pode ser disponibilizado para diferentes aplicações e usuários de forma diversa, de acordo com regras estabelecidas.

O processo de adaptação de conteúdo inclui a definição de quais os objetos deverão sofrer modificação, o local, a forma e o momento em que as adaptações devem ocorrer.

Os objetos podem ser adaptados em relação à qualidade, à forma de exibição e à semântica da informação (Rodrigues & Soares, 2003). A adaptação da qualidade de um objeto diz respeito à modificação de parâmetros como resolução, taxa de transmissão, duração do objeto, entre outros. Esta adaptação também inclui a substituição de um objeto por outro, denominada de adaptação *cross-media*, e pode ocorrer, por exemplo, quando um objeto de áudio pertencente a um documento necessita ser enviado a um dispositivo que não possui esta facilidade. Assim, a adaptação *cross-media* pode solicitar que o objeto de áudio seja substituído por um objeto texto, por exemplo.

A adaptação da forma de exibição altera o estilo da apresentação utilizado para o referido conteúdo, proporcionando que o mesmo conteúdo possa ser apresentado de forma diferente para cada usuário ou classe de usuário. A adaptação da semântica está relacionada com a quantidade de informação que compõe o conteúdo. Como um documento é composto por diferentes objetos e seus relacionamentos, a adaptação da semântica provê meios para a modificação dos objetos e seus relacionamentos.

Dois aspectos importantes na realização da adaptação referem-se ao momento e à localização em que deve ser efetuada. Em relação ao momento da realização da adaptação, o processo de adaptação de um objeto pode ocorrer antes da exibição ou distribuição do mesmo, quando a adaptação é dita estática e considera, por exemplo, as informações tais como perfil de usuário e de dispositivo, ou pode ser realizado durante a apresentação ou entrega do conteúdo, quando a adaptação é dita dinâmica e considera os fatores que influenciam a entrega e exibição do conteúdo, os quais podem variar durante a distribuição e/ou exibição. A adaptação dinâmica considera, entre outros, os fatores tais como a banda disponível da rede, latência e variação de latência, localização do cliente (em caso de ambientes móveis, por exemplo), os recursos computacionais disponíveis e, no caso de ambientes de exibição multimídia, as relações de sincronização intra e interobjetos.

Conforme pode ser visto na Figura 2, o ambiente de provisão de serviço inclui todos os recursos de hardware e software que compõem todo o sistema de comunicação, incluindo o transmissor e o receptor, eventuais servidores de adaptação, a rede de comunicação, os *proxies* e o ambiente do usuário.

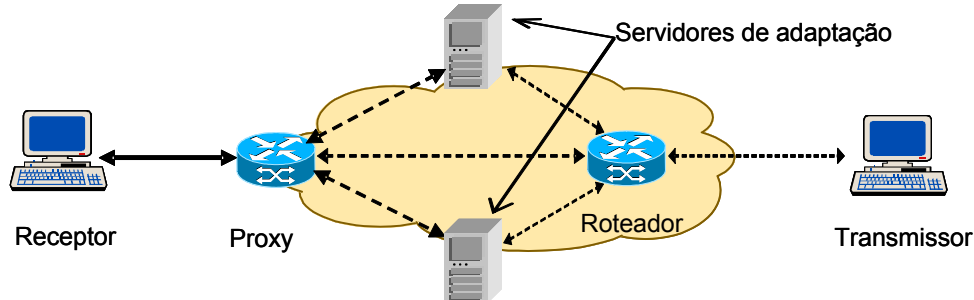


Figura 2: Ambiente de provisão de serviço.

O processo de adaptação pode ocorrer em qualquer dos pontos da rede onde haja a manipulação dos objetos. Assim, a adaptação pode ser realizada no ambiente do servidor de origem, no ambiente do usuário ou em algum recurso intermediário, tais como os servidores intermediários ou os *proxies* (Buchholz & Schill; Ma et al.). O local onde cada tipo de adaptação é efetuada influencia no desempenho do serviço, podendo prover reuso dos dados adaptados.

A adaptação de fluxos contínuos no transmissor ou em servidores intermediários normalmente é utilizada para adequar o volume de dados a serem transmitidos aos recursos disponíveis, tal como banda passante. Aplicações de

vídeo sob demanda também usam o ajuste elástico do tempo de exibição, realizado no transmissor ou em servidores intermediários, como ferramenta para evitar o desperdício de recursos de rede em operações de *fast-forward* ou *fast-rewind*, em arquiteturas cliente-servidor.

A adaptação de fluxos contínuos no receptor pode ser utilizada para a adequação da mídia aos recursos disponíveis no ambiente de exibição. Em sistemas multimídia/hipermídia, é essencial que o ambiente do usuário possa exercer controle sobre a exibição dos documentos, visto que, em última instância, é no ambiente do usuário que todos os problemas porventura ocorridos no documento são percebidos, tais como os problemas de sincronização entre os objetos que compõem o documento.

1.2. Objetivos

O serviço de controle do tempo de exibição de dados audiovisuais pode ser provido através do fornecimento de uma API entre o formatador e o respectivo exibidor, para que este último possa realizar o ajuste elástico em tempo real, ou seja, em tempo de exibição. Porém, os exibidores comerciais não provêm esse controle com a precisão adequada. Conforme será apresentado na Seção 3.6, o Windows Media Player Series 9, o RealPlayer, o QuickTime e o DivX Player provêm apenas a alteração da velocidade de exibição nas operações de *fast-forward* e *fast-rewind*. O Power DVD Versão 5 permite a utilização de alguns valores discretos como fatores de modificação da velocidade. O WinDVD 5 proporciona alterações mais em passos menores, entre os limites de 0,5 e 2,0. Finalmente, o Elecard MPEG-2 Player Versão 2.1 permite o ajuste elástico com passo de 0,1, através da modificação da taxa de renderização dos quadros.

A alternativa de modificação do código do exibidor resultaria em uma solução particular, impossibilitando o uso de outros exibidores comerciais disponíveis e restringindo o ambiente de execução hipermídia ao uso de um conjunto fixo de exibidores.

Por outro lado, a realização do ajuste elástico em exibidores, quando efetuado por meio da alteração da taxa de renderização de quadros, é adequada apenas para a adaptação efetuada no receptor. Caso seja necessário que a

adaptação ocorra no transmissor ou em algum ponto intermediário da rede, é importante que o fluxo resultado esteja codificado adequadamente para que seja transportado na rede de comunicação.

Esta dissertação propõe um adaptador de ajuste elástico para vídeos MPEG-2 capaz de realizar o ajuste elástico em fluxos de vídeos MPEG-2 e provendo, em sua saída, um fluxo MPEG válido, conforme apresentado na Figura 3, sem que seja necessária sua decodificação e recodificação.

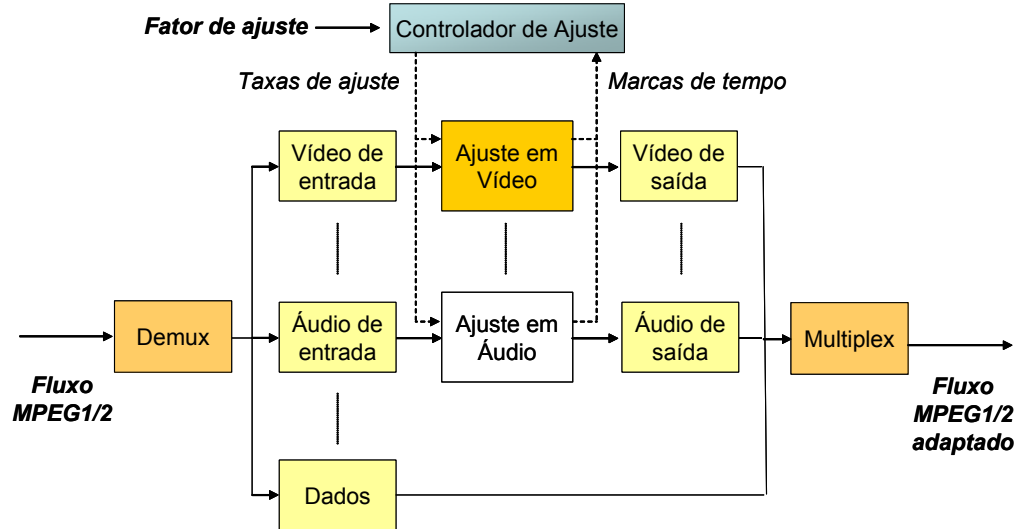


Figura 3: Sistema de ajuste elástico.

A forma trivial de execução do ajuste elástico é sua realização em fluxos não comprimidos, o que, para fluxos originalmente comprimidos, exige que o fluxo original seja inicialmente decodificado e, após a edição, seja novamente codificado. A edição dos dados audiovisuais diretamente em sua forma comprimida, no entanto, apresenta várias vantagens tais como menor memória necessária, menor tempo de acesso aos dados armazenados e menor requisito de poder processamento e tempo (Egawa et al., 2000). A maior vantagem, no entanto, reside sobre a qualidade durante a apresentação dos dados, devido às perdas inerentes ao processo de decodificação e recodificação, particularmente ao mecanismo de quantização da informação adotado em todos os padrões de codificação de mídias contínuas.

Como exemplo, foi estudado, em (Brightwell et al., 1997), o efeito de perda de qualidade apenas devido ao processo de decodificação e recodificação. Esse efeito pode ser visualizado, para um fluxo de vídeo, na Figura 4, onde foram utilizadas seqüências de teste padronizadas de fluxos MPEG de Vídeo (*Mobile &*

Calendar, *Basketball* e *Horseriding*). O gráfico também mostra que as perdas no processo de codificação são menores caso o tipo de cada figura, em relação ao fluxo original, seja mantido. Os possíveis tipos de figuras de um fluxo MPEG são discutidos na Seção 2.2.

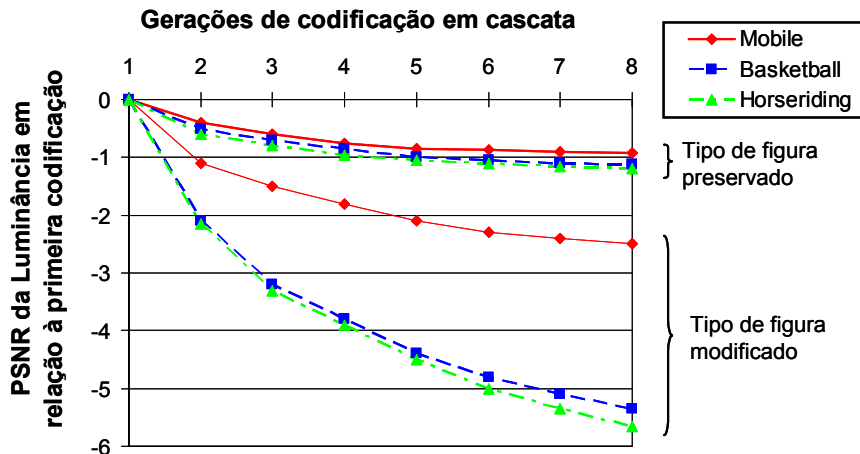


Figura 4: Efeito do processo de decodificação e codificação em cascata.

O objetivo desta dissertação é a apresentação de um adaptador capaz de realizar o ajuste elástico em dados audiovisuais no formato MPEG-2 (ITU, 2000a, b). Embora a ferramenta considere a necessidade do ajuste elástico nos sinais de vídeo e áudio, mantendo o relacionamento temporal entre os dois fluxos, ênfase especial será dada ao ajuste do sinal de vídeo (ITU, 2000b). Operações de ajuste elástico em fluxos MPEG de Áudio não fazem parte do escopo deste trabalho e maiores informações podem ser encontradas em (Omoigui et al., 1999; Covell et al., 2001; Arons, 1992; Maranhão et al., 2005).

Para possibilitar a utilização de exibidores comerciais MPEG-2, o fluxo audiovisual de saída do adaptador deve estar em conformidade com os padrões MPEG-2 de sistema, de vídeo e de áudio. Na realização do ajuste, deve-se ter especial atenção à recuperação do relógio pelo decodificador e à sincronização intra e intermídia. Quanto ao fluxo de vídeo, maior enfoque deve ser dispensado aos aspectos relativos à interdependência temporal entre os diferentes quadros e ao controle do *buffer* de entrada do receptor. Para o fluxo de áudio, devem-se evitar alterações no timbre do sinal e a inserção de ruídos ou artefatos.

O procedimento de ajuste deve poder ser realizado em tempo de exibição, o que sugere que todo ajuste deva ser executado nos dados codificados. A

transcodificação é desaconselhável pela introdução de um retardo adicional que em muitos casos não é tolerável, além da perda de qualidade inerente ao processo.

A importância de um adaptador para MPEG-2 deve-se à utilização desse padrão em uma ampla gama de aplicações comerciais, tais como nos sistemas de vídeo sob demanda, TV digital e TV interativa, e em diversos formatos de vídeo, tais como o VCD, o SVCD e o DVD. Dependendo da aplicação, o processo de adaptação pode ocorrer em diferentes pontos da rede de comunicação.

Como exemplo de aplicação em ambientes de execução de documentos hipermídia, o adaptador pode ser utilizado entre o formatador e o conjunto de exibidores, cuja tarefa é a de receber os fluxos audiovisuais originais e, a partir das informações do formatador, realizar os ajustes necessários, conforme representado na Figura 5. O fluxo resultante é, então, entregue aos exibidores apropriados, não exigindo deles qualquer modificação de funcionamento.

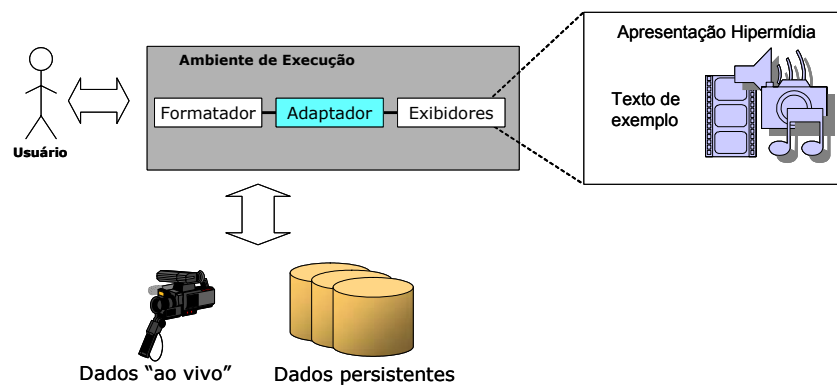


Figura 5: Ambiente de Execução com Ajuste Elástico.

1.3. Estrutura da Dissertação

Esta dissertação está organizada como se segue. O Capítulo 2 descreve os conceitos mais importantes do padrão MPEG-2 de Sistemas e de Vídeo, salientando os pontos em que operações de ajuste elástico causam impacto direto. O Capítulo 3 apresenta os trabalhos relacionados ao assunto. Nele, são mencionados sucintamente as principais técnicas e mecanismos de edição de dados audiovisuais e alguns sistemas de manipulação do tempo de duração de fluxos MPEG, anunciando as estratégias utilizadas e as questões não resolvidas de cada proposta. A escolha dessa abordagem baseia-se em que boa parte dos

trabalhos relacionados não tem como objetivo o ajuste elástico, mas apresentam técnicas que podem ser aproveitadas ou adaptadas para a realização deste tipo de adaptação.

O detalhamento de cada mecanismo de edição anunciado nos trabalhos relacionados, quando aplicável ao problema de ajuste elástico, é apresentado no Capítulo 4. Nele, é realizada a análise minuciosa dos problemas advindos do ajuste elástico diretamente em fluxos comprimidos e os principais algoritmos e técnicas aplicáveis para o ajuste em vídeos MPEG-2, salientando a autoria de cada um e as contribuições desta dissertação referentes aos algoritmos para o controle da ocupação do *buffer* do receptor e do ajuste correlacionado entre os fluxos de áudio e vídeo.

O Capítulo 5 descreve a implementação de um adaptador de ajuste elástico e discute, como exemplo de aplicação, sua integração a um formatador hipermídia. Os resultados obtidos através da ferramenta desenvolvida são apresentados no Capítulo 6. Finalmente, o Capítulo 7 discute as conclusões obtidas, as principais contribuições desta dissertação e os próximos passos a serem realizados.