

1 Introdução

A modificação da duração de exibição de um objeto de mídia é realizada por uma técnica conhecida como ajuste elástico. A alteração pode ser quantificada através de um fator de ajuste, expresso por um número real maior que zero. Se um objeto de mídia M , cuja exibição dura d unidades de tempo, for exibido com fator de ajuste f , sua exibição ajustada deverá ter a duração de $d*f$ unidades de tempo. Quando $f=1$ tem-se a exibição normal da mídia, sem alteração de duração. Uma alteração para acelerar a exibição é expressa por $f<1$ e para retardar é expressa por $f>1$.

A técnica de ajuste elástico pode ser aplicada a uma mídia estática (como texto e imagem) ou contínua. Considerando que o ajuste da duração das mídias estáticas é trivial, este trabalho irá focar na descrição de algoritmos para mídias contínuas, mais especificamente o áudio. O termo ajuste elástico é conhecido em Inglês como *timescale*, *elastic time computation*, *time stretching/sketching* e *time compression/expansion*.

As subseções a seguir ilustram a motivação para realização deste trabalho, os objetivos a serem atingidos e a estrutura do restante do documento.

1.1. Motivação

Muitas aplicações são beneficiadas pelo uso de ajuste elástico (Amir et al., 2000; Arons, 1992; Lee et al., 2004; Omoigui et al., 1999). Estudantes de uma nova linguagem podem desejar escutar gravações com voz mais lenta para melhor compreender a pronúncia dos vocábulos. Músicos também podem desejar escutar uma música mais lentamente para entender melhor uma seqüência rítmica ou melódica. Aplicações de síntese de voz a partir de texto irrestrito podem utilizar ajuste elástico para modificar a duração de um segmento de voz sintetizado de acordo com a prosódia inerente ao contexto. A aplicabilidade principal do algoritmo proposto neste trabalho, entretanto, ocorre em cenários que não existe

interferência do usuário, pois o fator de ajuste é dinamicamente calculado por aplicações clientes.

Como primeiro exemplo, pode-se citar a manutenção da consistência temporal de um documento hipermídia, isto é, a garantia que todos os objetos de mídia serão exibidos no tempo obedecendo às restrições impostas às suas sincronizações relativas, especificadas pelo autor do documento. Por exemplo, em documentos hipermídia, a exibição de um fluxo contendo vídeo e áudio em ambientes distribuídos pode ser prejudicada por atrasos na rede durante a transmissão. Se a transmissão do fluxo de vídeo sofrer retardo maior do que o fluxo de áudio, pode-se modificar a taxa de reprodução de uma das mídias para que a sincronização com os movimentos labiais (*lip-sync*) não seja comprometida. Nesse caso, o ajuste precisa ser aplicado o mais próximo possível do exibidor de conteúdo (*player*) para compensar a variação estatística do retardo da transmissão na rede.

Em outros exemplos, o ajuste deve ser aplicado o mais próximo possível do transmissor. Por exemplo, ajuste elástico pode ser utilizado para permitir que emissoras de rádio e TV acelerem ou retardem a sua programação. Supondo que um determinado comercial precisa ser exibido em um horário fixo, pode ser preciso aplicar ajuste elástico em programas que antecedem a exibição do comercial a fim de compensar possíveis atrasos da programação. Outro exemplo de aplicação ocorre durante a transmissão de servidores de vídeo sob demanda (*Video-On-Demand* ou VOD) para múltiplos usuários. Existem várias técnicas de otimizar a alocação de fluxos de transmissão para usuários. Uma possibilidade é transmitir a informação audiovisual com diferentes velocidades até que os fluxos estejam transmitindo o mesmo trecho de informação em um dado instante. Nesse momento, diferentes fluxos podem ser unificados numa mesma transmissão *multicast*. É importante que a variação de velocidade seja pequena, tipicamente 5% no máximo, para não degradar a qualidade da informação recebida (Soares, 2005; Golubchik et al., 1995). Para implementar essa idéia, pode-se armazenar várias cópias do mesmo conteúdo audiovisual com diferentes velocidade de exibição (Lin et al., 2001) ou aplicar ajuste elástico nos conteúdos durante a transmissão.

Como as aplicações hipermídia envolvem diferentes tipos de mídia, que em geral estão comprimidas, é importante que o algoritmo de ajuste suporte formatos

audiovisuais comprimidos. Além do mais, tais aplicações precisam realizar o ajuste elástico em tempo de exibição e com possibilidade de variação do fator, uma vez que o sincronismo deve ser monitorado e mantido durante toda a apresentação de um documento hipermídia. Aplicar ajuste elástico em tempo de exibição significa processar o fluxo de mídia à medida que ele é recebido, gerando um novo fluxo de dados, com um retardo imperceptível para o usuário. A variação do fator aplicado em tempo de exibição permite modificar a velocidade de exibição das mídias contínuas de acordo com a necessidade corrente da aplicação. A fidelidade da mídia resultante da técnica de ajuste deve ser alta e o processamento deve ser realizado da forma o mais linear possível, de tal modo que a variação de velocidade na exibição da mídia resultante seja mínima.

É também interessante que a aplicação de ajuste seja capaz de monitorar o novo valor de um instante específico da mídia original durante a realização do ajuste. Por exemplo, considerando que documentos hipermídia utilizam âncoras¹ para indicar pontos de sincronismo, é importante conhecer quais são os novos valores dessas âncoras após o processamento.

Uma vez que a realização de ajuste não implica necessariamente que a mídia ajustada deva ser imediatamente exibida (podendo ainda sofrer outras transformações ou mesmo ser enviada pela rede) e, principalmente, porque o mecanismo de ajuste não deve ser dependente de qual exibidor será utilizado, é também importante que a aplicação de ajuste elástico seja independente da decodificação que ocorre no exibidor do conteúdo do fluxo de mídia.

Por fim, é importante que a aplicação cliente possa facilmente solicitar o ajuste elástico em qualquer local da transmissão, ou seja, no transmissor, receptor ou em um nó intermediário, e que seja possível manipular não somente arquivos locais, como também remotos, e até mesmo fluxos de mídia de dados “ao vivo” (dados de mídia que não estão armazenadas em arquivos, mas transmitidos como um fluxo de bits na rede).

¹ Uma âncora identifica um subconjunto de unidades de informação de um objeto de mídia. No contexto deste trabalho, uma âncora marca um instante de tempo em um fluxo audiovisual.

1.2. Objetivos

O objetivo principal deste trabalho é propor um algoritmo de ajuste que atue em áudio comprimido e que atenda aos requisitos das aplicações descritas na seção anterior.

O algoritmo de ajuste será então integrado a um algoritmo de fluxos de vídeos através de desenvolvimento de um terceiro algoritmo para fluxos de sistemas².

O algoritmo de ajuste de áudio irá manipular fluxos de mídia cujo formato de dados é dividido em quadros, que são unidades independentes (ou quase independentes) de decodificação. Como exemplo, o algoritmo será instanciado para os formatos MP2 e MP3, especificados pelos padrões MPEG-1 (ISO, 1993), MPEG-2 BC (*Backward Compatible*) (ISO, 1998), AC-3 (ATSC, 1995) e MPEG-2/4 AAC (ISO, 1997; ISO, 2001b).

O algoritmo de vídeo a ser utilizado foi desenvolvido por Cavendish (Cavendish, 2005) e contempla o formato MPEG-2 vídeo (ISO, 2000b). Finalmente, o mecanismo de ajuste de fluxos de sistemas será desenvolvido para o MPEG-2 sistemas (ISO, 2000a).

Como último objetivo deste trabalho, a ferramenta de ajuste deverá ser integrada a um formatador hipermídia, o HyperProp. O formatador HyperProp é um controlador de apresentações hipermídia que incorpora mecanismos de otimização para ajuste elástico das apresentações, a fim de manter o sincronismo temporal dos documentos exibidos (Rodrigues, 2003).

1.3. Organização do restante do documento

O restante deste documento está estruturado como se segue. O Capítulo 2 classifica os algoritmos de ajuste elástico de áudio existentes na literatura, apresenta alguns trabalhos relacionados e os compara com o trabalho proposto.

O Capítulo 3 descreve o mecanismo de ajuste elástico de áudio desenvolvido, assim como sua particularização para cada formato suportado e sua execução como fluxo elementar de um fluxo de sistemas. A ferramenta de ajuste é

² Um fluxo de sistema contém dados de áudio e vídeo multiplexados.

descrita no Capítulo 4. Esse capítulo apresenta como ferramentas clientes podem solicitar serviços de ajuste elástico através de APIs de ajuste e detalhes da implementação da ferramenta, o que inclui a implementação do mecanismo apresentado no Capítulo 3 e a integração do ajuste de áudio com o vídeo. O Capítulo 5 apresenta o mecanismo de integração desenvolvido para possibilitar a apresentação da mídia ajustada em exibidores de conteúdo.

O Capítulo 6 apresenta a integração da ferramenta de ajuste em sistemas hipermídia, em especial com o formatador Hyperprop. O Capítulo 7 apresenta medidas de qualidade, subjetivas e objetivas, extraídas do algoritmo de ajuste de áudio proposto por este trabalho. As análises subjetivas foram realizadas com base em um experimento também descrito nesse capítulo. Além disso, esse capítulo também realiza uma análise da implementação da ferramenta de ajuste descrita no Capítulo 4.

O Capítulo 8 apresenta conclusões e trabalhos futuros e o Capítulo 9 lista referências utilizadas no decorrer deste trabalho.