

Sincronismo de Vídeo MPEG-2 VBR Sobre Redes ATM

Luciano Junqueira Campos *+

luciano@telemidia.puc-rio.br
luciano.junqueira@tvglobos.com.br

Antônio Tadeu A. Gomes *

atagomes@telemidia.puc-rio.br

Luiz Fernando G. Soares *

lfgs@inf.puc-rio.br

* Laboratório TeleMídia
Departamento de Informática PUC-Rio
R. Marquês de São Vicente 225
22453-900 – Rio de Janeiro, Brasil

+ Departamento de Pesquisa e Desenvolvimento
TV Globo Ltda.
R. Von Martius 22
22460-040 – Rio de Janeiro, Brasil

Resumo

A transmissão de vídeos MPEG-2 VBR em redes de computadores comutadas por pacotes é uma tarefa bastante complexa. A sincronização entre os relógios do codificador e do decodificador é de difícil realização graças aos retardos variáveis introduzidos por estas redes. Este artigo propõe uma solução para a transmissão de vídeo MPEG-2 VBR em redes ATM, onde a variação do retardo é eliminada através de uma bufferização controlada, que visa moldar o fluxo MPEG de volta a um formato que esteja em conformidade com as especificações do padrão. Este trabalho é resultado de uma parceria entre o Laboratório TeleMídia da PUC-Rio e o Departamento de Pesquisa e Desenvolvimento da TV Globo.

Palavras chaves: sincronização, MPEG, ATM, transmissão de vídeo

1 – Introdução

Transmissão de vídeo em redes digitais de telecomunicações está presente nas diversas classes das chamadas *aplicações em banda larga*. Este termo, *banda larga*, vem do fato de tais aplicações exigirem do sistema de comunicação uma maior largura de banda do que as aplicações tradicionais.

Não obstante a classe, os requisitos de banda passante são dependentes da qualidade das informações transmitidas. Quando se fala da qualidade de um sinal de vídeo, refere-se a qualidade do sinal de saída de um dispositivo de captura do vídeo. É importante notar, embora seja óbvio, que a qualidade do sinal de saída nunca poderá ser melhor que a do sinal de entrada. Assim, de nada adianta um bom codificador de vídeo, quando se tem um sinal de entrada de baixa qualidade. Da mesma forma, de nada adianta um bom codificador quando o decodificador é de baixa qualidade. Em todo o texto é assumido que a qualidade do decodificador é pelo menos a mesma do codificador.

Este artigo enfoca a transmissão de vídeo, de qualidade igual ou superior a TV, em redes que apresentam retardo de transmissão variável, em especial redes ATM. Para introdução do ambiente de desenvolvimento, a Seção 2 inicia discutindo as várias qualidades de um sinal de vídeo e apresentando alguns cenários de transmissão, chamando a atenção para as dificuldades da transmissão de vídeo de taxa variável em redes com retardos variáveis, principalmente para transmissões em tempo real. Nesta seção, o cenário de interesse do trabalho é caracterizado. Como o foco do trabalho é vídeo com qualidade igual ou superior a qualidade TV, o padrão MPEG-2 [ISO94] tem um papel de destaque. A Seção 3 discute as características dos fluxos MPEG-2 no que tange a sincronização temporal dos codificadores e decodificadores, tema focalizado neste artigo. A sincronização propriamente dita e a solução proposta são assuntos da Seção 4. Finalmente, a Seção 5 é reservada para as conclusões e discussão de trabalhos futuros.

2 – Alguns Cenários de Transmissão de Vídeo Digital

A qualidade do sinal de vídeo pode ser medida em termos da resolução geométrica, da resolução de cor, da resolução temporal e das perdas introduzidas pelo processo de codificação. Todos esses parâmetros são importantes na determinação da banda passante necessária do sistema de comunicação que dará suporte à transmissão do sinal. Note que aqui foi introduzido o conceito de *qualidade de sinal* e *qualidade da informação* transmitida. Utilizando um caso extremo para exemplificar, um vídeo transmitido a 30 quadros por segundo certamente tem uma qualidade de sinal melhor do que se fosse transmitido a 10 quadros por segundo; por outro lado, se fosse filmada uma paisagem sem qualquer movimento, a qualidade da informação transmitida seria a mesma (nesse caso extremo bastaria a transmissão de um quadro).

A resolução geométrica é medida pela quantidade de pixels (resolução horizontal e vertical) de cada componente (RGB ou YCrCb) de um quadro de vídeo.

A resolução de cor é medida pela quantidade de bits utilizada para codificar cada componente. A resolução temporal é medida em termos de quadros/segundo.

A taxa de perda introduzida pela codificação depende dos algoritmos de compressão utilizados e do vídeo sendo transmitido. A tolerância a perdas depende da informação a ser transmitida. Por exemplo, em uma aplicação de telemedicina, pode ser que uma imagem sendo transmitida não possa tolerar perda de nenhuma informação, pois esta perda poderia representar exatamente a anomalia que se quer estudar. Nesse caso, um algoritmo de codificação sem perdas deve ser utilizado. No entanto, tal robustez implica na eliminação de muito pouca redundância do sinal de vídeo, gerando uma taxa de bits elevada e exigindo, por conseguinte, uma grande banda passante do sistema de comunicação.

Ainda com relação às perdas devido a compressão do sinal de vídeo, algumas observações ainda se tornam necessárias para determinar qual codificador utilizar.

O sinal de vídeo de entrada em um codificador é um sinal CBR (constant bit rate), isto é, apresenta uma taxa de bits constante. Na saída do compressor (parte do codificador) o sinal pode se tornar VBR (variable bit rate), devido a eliminação de redundâncias espaciais e temporais. Quando há muitas mudanças de cena a taxa de saída do compressor tende a ser maior do que quando há poucas mudanças.

Alguns codificadores foram projetados para utilizarem sistemas de comunicação onde a taxa de bits contratada é fixa. No caso de um sinal VBR, a taxa de bits (banda passante) contratada deveria ser a taxa de pico do sinal de saída do compressor. Neste caso, quando o sinal VBR não se apresentar em sua taxa de pico, banda passante estará sendo desperdiçada. Para melhor aproveitar a banda contratada, alguns codificadores implementam buffers na saída do compressor, de forma a se poder contratar uma banda menor que a exigida pela taxa de pico. Caso o buffer tenda a esvaziar, uma realimentação aumenta a qualidade (diminui as perdas de compressão) do sinal de vídeo, aumentando a taxa de entrada no buffer. Caso o buffer tenda a saturar, a qualidade do sinal de vídeo é deteriorada de forma a diminuir a taxa de entrada no buffer. Agindo dessa forma, a saída do codificador é CBR, embora a saída do compressor seja VBR (caso do codificador de vídeo H.261 [ITU90] e de codificadores MPEG CBR). Como o buffer tende a saturar quando são geradas muitas informações, isto é, quando há muita mudança de cena no vídeo, tais codificadores CBR se aplicam melhor a vídeos com pouco movimento, pouca mudança de cena.

Quando o sistema de comunicação não reserva uma banda fixa para o canal de vídeo, codificadores de vídeo com saída VBR podem tirar um melhor proveito do meio do que codificadores com saída CBR,

para uma mesma qualidade de vídeo. Codificadores baseados nos padrões MPEG e MJPEG geram saída VBR, sendo assim os preferidos para vídeos com muitas mudanças de cena.

O mesmo princípio do buffer com realimentação para o compressor pode ser utilizado não mais com a finalidade de gerar uma saída CBR, mas para adaptar a saída VBR do compressor à taxa disponível na rede. Estes compressores podem ter um bom desempenho em sistemas de comunicação que apresentam a qualidade de serviço de melhor esforço (best effort, como é o caso da Internet atual), mas não apresentam nenhuma vantagem em redes onde a banda passante contratada é fixa (exemplificado por LPs e pela RDSI-FE [ITU94]). Em redes de pacotes com garantias de qualidade de serviço (como redes ATM e implementações atuais da Internet II), tais codificadores apresentam um desempenho pior que todos os outros tipos de codificadores mencionados, pois não tiram proveito de uma banda passante garantida que pode ser negociada.

Vários codificadores permitem a escolha dos parâmetros resolução geométrica, resolução de cor e resolução temporal. A escolha desses parâmetros determinará a taxa de bits produzida pelo codificador. Como exemplo de codificadores parametrizados, pode-se citar os que seguem os padrões de codificação H.261 e MPEG-2 (embora seja possível encontrar codificadores que sigam um padrão mas não ofereçam opções de configuração).

O padrão MPEG-2 é organizado em diversos perfis e níveis. Perfis possuem um conjunto de características que atende a uma determinada classe de aplicações. Definem o espaço de resolução de cores e a escalabilidade do fluxo de bits. Níveis definem as resoluções geométricas máximas e mínimas, o número de amostras de luminância e crominância, número de camadas de vídeo e áudio suportadas por perfis escaláveis e a taxa máxima em bps para cada perfil.

O nível baixo (low level) é compatível com o padrão MPEG-1 [ISO93], com resolução geométrica e temporal dada por: SIF (Source Input Format) — 382 x 240 x 30. A qualidade de vídeo TV está baseada no Perfil Principal e no Nível Principal do padrão MPEG, MP@ML (Main Profile at the Main Level); CCIR 601 [ITU90] — 720 x 480 x 30. A qualidade de vídeo HDTV está baseada no Perfil Principal e no Nível Alto do padrão MPEG, MP@HL (Main Profile at the High Level).

Apesar das dificuldades, as técnicas para transmissão de vídeo digital em redes já encontram-se bastante evoluídas e, em alguns cenários, a tecnologia já apresenta soluções bastante satisfatórias.

Cenário 1: Vídeo de Baixa Qualidade em Redes com Comutação de Circuito

Entende-se no contexto deste artigo como vídeo de baixa qualidade, qualquer vídeo com as seguintes características:

- Baixa resolução geométrica – desde um quadro QSIF (144x176) até por volta de qualidade de VCR (280x330)
- Possivelmente sub-amostragem temporal – aplicações como vídeo de segurança, por exemplo, podem usar taxas tão pequenas quanto 2 quadros por segundo
- Normalmente as perdas introduzidas pela codificação não são relevantes. Portanto, deve-se usar uma representação eficiente para codificação do vídeo, de forma a aproveitar a alta correlação espacial e temporal existente na mídia de vídeo.
- Perdas introduzidas pela rede não são corrigidas (a rede é suposta com baixa taxa de erro, mesmo quando muitas vezes não é).

Como consequência do exposto acima, a característica mais marcante de vídeos de baixa qualidade é a baixa taxa de dados gerada (desde que se esteja usando uma representação minimamente eficiente).

A natureza do tráfego gerado, se CBR ou VBR, é importante na utilização eficiente da largura de banda. Para vídeos CBR, basta alocar um circuito com largura suficiente para a taxa de vídeo. Para vídeos VBR, a transmissão em uma rede com comutação de circuitos é feita sem problemas, desde que se aloque o canal pela taxa de pico, com desperdício de banda passante.

Cenário 2: Vídeo de Baixa Qualidade em Redes com Comutação de Pacotes

A natureza do tráfego gerado, se CBR ou VBR, é determinante na sincronização dos relógios de codificador e decodificador. A introdução de retardos variáveis por este tipo de redes torna este sincronismo um problema de difícil solução.

Para vídeos CBR, a sincronização é facilitada pelo fato de se ter uma taxa constante de geração (pode-se citar como exemplo de solução a camada AAL1 de uma rede ATM [ITU96]) e, na grande maioria dos casos, pelo fato de vídeos de baixa qualidade em redes de pacotes não exigirem transmissão em tempo real interativa e reprodução síncrona (exemplificado pelo uso de vídeos em sistemas de freeze-frame videoconferência, do tipo CU-see-me)

Para vídeos VBR, a sincronização dos relógios do codificador e decodificador é um problema complicadíssimo, fazendo com que estas redes sejam utilizadas apenas nos casos em que não é exigido transmissão em tempo real interativa e com reprodução síncrona.

Cenário 3: Vídeo de Qualidade TV, CBR, em Redes com Comutação de Circuito e Sem Correção de Erros

Um vídeo de qualidade de TV deve apresentar uma resolução espacial de, no mínimo, 320x240 e resolução temporal de 30 quadros por segundo. Desta feita, a taxa gerada por um vídeo com estas características é bastante alta. Em emissoras de TV, a produção de conteúdo em geral utiliza o algoritmo MJPEG, com taxas de 50 Mbps. Para jornalismo, MPEG entre 18 e 25 Mbps são preferidos. Mesmo para distribuição, onde a qualidade não precisa ser tão grande, utiliza-se um fluxo MPEG com taxa bastante alta, em torno de 8Mbps.

Este cenário, apesar da alta banda passante requerida da rede, não é problemático, uma vez que o vídeo é CBR. Basta, então, alocar um canal da rede com a taxa do vídeo para que a transmissão seja feita de forma eficiente, da mesma forma como foi feito para vídeos de baixa qualidade.

Cenário 4: Vídeo de Qualidade TV (ou superior) em Redes com Comutação de Pacotes

Este cenário pode ser subdividido em vários outros, conforme o vídeo seja CBR ou VBR, e se a correção de erros é necessária.

Além da taxa de bits ser alta, outro fator importante é que a reprodução deve ser síncrona (sincronismo intra mídia) e, normalmente, também síncrona com o áudio associado (sincronismo inter mídia), mesmo se tendo por infra-estrutura de comunicação uma rede com retardos variáveis.

Para se obter o sincronismo entre o codificador e o decodificador, a variação estatística do retardo da rede (jitter) deve ser compensada. Existem vários algoritmos para compensação propostos na literatura, mas que não mantêm a taxa de dados entregues ao decodificador exatamente síncrona com o relógio do

codificador [SoLC95]. Manter os relógios do transmissor e receptor sincronizados é muito importante em transmissões de vídeo em tempo real.

Quando a taxa de geração de vídeo codificada é CBR, existem técnicas que permitem não só compensar a variação estatística do retardo, como também sincronizar os relógios de saída do codificador com o de entrega de dados ao decodificador. Um exemplo de solução já citado pode ser encontrado na camada AAL1 de uma rede ATM.

Quando a taxa é, no entanto, VBR, a tarefa é bastante complicada e o codificador tem de informar de algum modo o padrão VBR gerado, devendo este padrão ser reconstituído na entrada do decodificador. Cabe a rede manter esta informação precisa. Baseada nestas informações, a variação de retardo da rede deve ser compensada e dados devem ser entregues ao decodificador na mesma taxa em que são gerados no codificador.

Não existe uma solução geral para sincronismo de vídeo VBR em redes comutadas por pacotes. Como o fluxo MPEG-2 é o padrão para difusão de vídeo qualidade TV e como este fluxo possui alguma informação sobre a taxa VBR gerada, existem algumas propostas para o tráfego deste fluxo em redes comutadas por pacotes, em especial redes ATM. A Seção 4 aborda algumas destas soluções.

A sincronização na transmissão de vídeo MPEG-2 (qualidade TV e HDTV) VBR sobre redes ATM é o foco do trabalho que vem sendo desenvolvido no Laboratório TeleMídia da PUC-Rio em conjunto com a TV Globo. Este artigo enfoca apenas o problema da sincronização entre codificador e decodificador, levando em conta que, embora o vídeo gerado pelo codificador seja VBR, o vídeo decodificado é CBR, podendo-se então fazer um ajuste mais fino do que o proposto nos outros algoritmos da literatura.

3 – Fluxo MPEG-2

O padrão MPEG-2 [ISO94] define apenas a sintaxe e a semântica do fluxo MPEG. O funcionamento do codificador e do decodificador é inteiramente livre. Esta peculiaridade é importante para esta discussão, pois implica que não se pode fazer pressuposições sobre o funcionamento do decodificador ou do codificador. Esta seção discute o fluxo MPEG-2 no que tange apenas às suas especificidades para o sincronismo temporal de codificador e decodificador. Uma descrição completa do fluxo pode ser obtida em [ISO94a] e [ISO94b].

Um fluxo elementar (*PES*) é a representação de uma mídia, podendo ser tanto áudio quanto vídeo, ou ainda dados. De maneira a combinar diversos fluxos elementares correlacionados, o padrão MPEG define a camada *Systems*, que é responsável por multiplexar os diversos fluxos formando um único fluxo. O resultado desta multiplexação é um *fluxo de programa*¹ ou um *fluxo de transporte*². Tanto os fluxos de programa quanto os de transporte provêm a sintaxe necessária e suficiente para a sincronização do decodificador e da apresentação da informação de vídeo e de áudio, assegurando que os buffers de dados no decodificador não sofram transbordo (*overflow*) nem esvaziamento (*underflow*). No entanto, para que tal sincronização aconteça, é assumido que o retardo entre a saída do codificador

¹ Um fluxo de programa carrega um programa, que é um conjunto de fluxos elementares MPEG (áudios, vídeos, informações auxiliares) que são relacionados entre si, e apresentam, desta forma, a mesma base de tempo. O fluxo de programa foi desenvolvido para uso em ambientes relativamente livres de erros.

² Um fluxo de transporte combina um ou mais programas com uma ou mais bases de tempo independentes em um único fluxo. O fluxo de transporte foi desenvolvido para uso em ambientes onde erros são prováveis.

e a entrada do decodificador é constante (vide Figura 1), e aí reside o grande problema no transporte destes fluxos em redes com retardos variáveis.

Cada fluxo, elementar ou não, tem um modelo de tempo no qual o retardo entre a entrada no codificador até a saída do decodificador é constante. Este retardo é a soma do retardo da codificação, da *bufferização* no codificador, da multiplexação, da transmissão (ou armazenamento) para o decodificador, da demultiplexação, da *bufferização* do decodificador, da decodificação e os retardos de apresentação, como mostra a Figura 1. Todo modelo de temporização é definido em termos de um relógio comum do sistema, denominado *STC* (*system time clock*). Nos fluxos de transporte, este relógio deve ter a mesma frequência (denominada *frequência_do_relógio_do_sistema*) das amostras de vídeo e de áudio; nos fluxos de programa uma pequena variação é permitida.

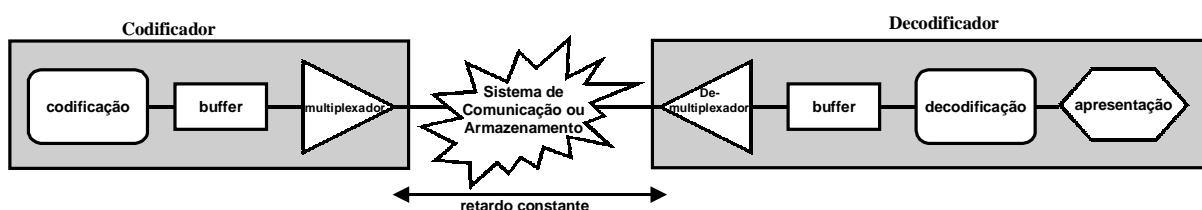


Figura 1 – Modelo de tempo do sistema

A sincronização entre múltiplos fluxos elementares é realizada pelo campo *presentation time stamps* (*PTS*) e faz parte do projeto do decodificador, fugindo ao interesse deste artigo. A sincronização do sistema de decodificação é obtida através dos campos *SCR* (*system clock reference*), nos fluxos de programa, e de seu análogo *PCR* (*program clock reference*), dos fluxos de transporte. *SCR* e *PCR* são “selos de tempo” codificando a temporização dos fluxos de bits. *SCR* e *PCR* são derivados da mesma base de tempo que o *PTS* do mesmo programa. Como cada programa pode ter sua própria base de tempo, existe um campo *PCR* separado para cada programa de um fluxo de transporte. Valores para o *PTS*, *PCR* e *SCR* são baseados em dois relógios, um de 90kHz e outro de 27MHz.

De acordo com o padrão, dados do fluxo de transporte e de programa são gerados e, portanto, devem entrar no decodificador³, a uma taxa variável, mas constante dentro de intervalos de tempo, conforme apresenta a Figura 2.

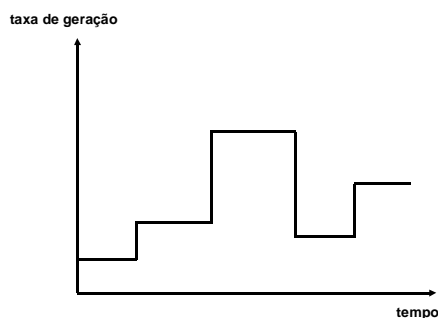


Figura 2 – Taxa de chegada de dados ao decodificador (ou de saída do codificador)

³ Uma vez que o retardo no sistema de comunicação é constante, a mesma VBR da saída do codificador é mantida na entrada do decodificador.

Dado um fluxo de transporte, seja:

- i' o índice do byte contendo a referência de relógio mais recente, denominada $PCR(i')$
- i'' o índice do byte contendo a referência de relógio anterior, denominada $PCR(i'')$
- i o índice de qualquer byte dentro de um fluxo de programa ou de transporte, tal que $i'' < i < i'$

Então, tem-se:

$$taxa(i) = \frac{(i' - i'') \times \text{frequência_do_relógio_do_sistema}}{PCR(i') - PCR(i'')} \quad (1)$$

Para um fluxo de programa, a taxa em um pacote, denominada *program_mux_rate*, é um valor especificado no próprio pacote. O formato de ambos os fluxos obedece ao gráfico da Figura 2.

Tem-se então dois cenários a discutir.

No primeiro, o que alimenta o decodificador é um vídeo armazenado. O decodificador pode, neste caso, trabalhar com um relógio diferente do codificador e ir buscar os dados de um fluxo quando precisar.

O segundo cenário é aquele que entre o codificador e o decodificador tem-se um sistema de comunicação sem qualquer armazenamento e com a geração de vídeo em tempo real. Poder-se-ia subdividir este cenário em duas situações.

A primeira é aquela em que o sistema de comunicação apresenta um retardo constante. Neste caso não há nenhum problema, pois é exatamente a situação que o padrão obriga os decodificadores a resolver, ou seja, retardo constante entre a codificação e a decodificação.

A segunda é aquela em que o sistema de comunicação apresenta um retardo variável. Uma solução que poderia ser pensada seria armazenar todo o programa de vídeo antes de entregá-lo ao decodificador, recaindo no primeiro cenário. Contudo, isto pode implicar na introdução de um grande retardo, impossibilitando aplicações interativas, como por exemplo, todas as aplicações de banda larga classificadas pelo ITU-T como *Serviços Conversacionais* [ITU94]. A solução deve então se concentrar em recuperar na entrada do decodificador o mesmo padrão VBR gerado pelo codificador, como propõe este artigo na próxima seção.

4 – Sincronização em Transmissões em Tempo Real

Várias soluções foram propostas na literatura para tratamento do problema da sincronização na transmissão de vídeos em sistemas de comunicação com retardo variável. Grande parte dessas soluções baseiam-se em idéias apresentadas em anexos do próprio padrão MPEG. Tais idéias são resumidas a seguir. Ao final da seção, será apresentada a técnica proposta neste trabalho.

Esta seção se limita, mas sem perda de generalidade, às técnicas desenvolvidas para fluxos de transporte, uma vez que é a forma mais adequada de multiplexação MPEG-2 para transmissão em redes.

As soluções sugeridas pelo padrão e em vários outros trabalhos partem, em geral, para a alteração do funcionamento do decodificador, mexendo no circuito PLL que é usado para recuperação do relógio de forma que ele ganhe robustez para suportar a grande variação no atraso de transmissão dos valores de amostra do relógio. São assim técnicas adequadas à implementação e alteração de decodificadores. Este

artigo, entretanto, parte no sentido oposto, assumindo que os codificadores não são alteráveis e que a variação do retardo no sistema de comunicação deve ser compensada externamente ao sistema de comunicação e ao decodificador. No entanto, as técnicas de alteração do decodificador são apresentadas, para uma comparação da complexidade de todas as técnicas.

Alteração do PLL

O anexo K do padrão MPEG-2 demonstra como o jitter da transmissão pode ser prejudicial ao sinal de vídeo gerado, partindo dos limites de tolerância dos padrões de sinal de vídeo (NTSC e PAL) e deduzindo os valores máximos de jitter que podem ser introduzidos pela rede, dadas as características do PLL usado pelo decodificador.

Este anexo é amplamente utilizado na literatura (exemplo [Adan96]) para cálculo em simulações, mas sempre parte do pressuposto de que o projeto do decodificador é aberto, limitando seu escopo de aplicação.

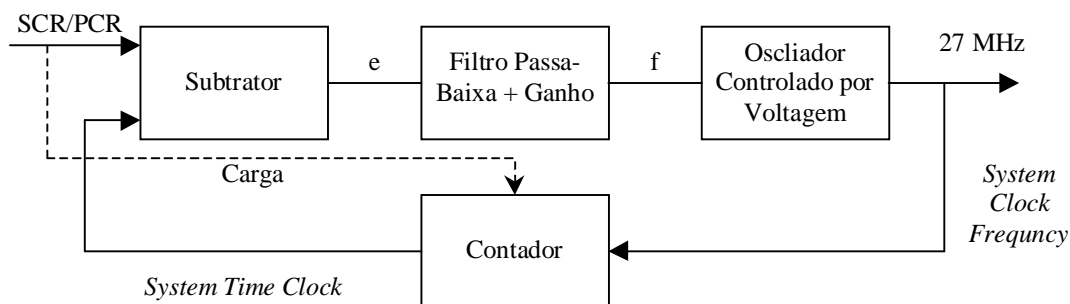


Figura 3 – PLL normalmente utilizado para recuperação de relógio

A redução do jitter é conseguida através de uma alteração no funcionamento do PLL de recuperação do relógio, apresentado na Figura 3, de forma que:

- Quando o erro e for muito grande, este não é usado para corrigir a frequência de oscilação;
- Manter sempre o buffer com um enchimento constante (note que este mecanismo só funciona para vídeo CBR);

Pode-se assim:

- Medir o tempo em que um dado entra no buffer, e controlar sua saída;
- Com esta informação, adaptar o PLL para prover atraso constante no buffer.

Relógio de Rede

Nesta solução, o padrão MPEG sugere o uso de um relógio independente do relógio do sistema, chamado relógio de rede. Uma camada adaptadora de rede é desenhada para compensar um jitter pico a pico máximo de J segundos. O agendamento do envio de bytes é feito usando-se o valor de CR (*Clock Reference*), que é conseguido a partir do TC (*Time Clock*). Numa analogia, pode-se comparar o CR ao PCR, e o TC ao STC.

Os pacotes enviados pela rede carregam um campo CR, com o valor corrente do TC. Após a recepção, os CR's são extraídos no decodificador, e usados para reconstruir o TC, por exemplo com o uso de um PLL.

Esta solução necessita do conhecimento do valor máximo do jitter e do atraso fim a fim, pois baseia seu funcionamento na enfileiramento de pacotes que chegaram muito rápido, e no envio imediato de

pacotes que chegaram atrasados. Para saber se o pacote chegou cedo ou tarde é necessário que se tenha o conhecimento do momento em que cada pacote deveria chegar, o que é impossível de ser feito, se os relógios do transmissor e do receptor não são sincronizados.

Esta idéia é aplicável, entretanto, no caso de se ter um relógio comum na rede. Este relógio seria então usado para sincronizar transmissor e receptor sem dificuldades.

Oscilador Livre

Pode haver situações em que é possível não se ter um PLL no decodificador, mas manter um oscilador livre, e sempre que uma referência de relógio for recebida, alterar o relógio corrente, provocando uma descontinuidade desta forma no contador de relógio. Esta descontinuidade impediria o uso do relógio do decodificador para gerar relógios para apresentação do vídeo e do áudio, além de provocar um efeito de repetição ou salto de unidades de apresentação⁴ de vídeo e áudio e complicar o gerenciamento dos buffers.

Algumas outras soluções

A referência [SoCL95] apresenta cinco algoritmos para compensação da variação do retardo em redes comutadas por pacotes. Cada algoritmo tem sua aplicação mais adequada, dependendo se é desejado privilegiar a interatividade do sinal (isto é, o retardo), a diminuição das perdas, ou ainda um compromisso entre as duas coisas. Os algoritmos partem sempre do pressuposto que perdas são toleráveis e que, portanto, os relógios dos transmissores e receptores não precisam estar exatamente sincronizados. O estouro ou esvaziamento do buffer de compensação devido a diferença dos relógios, é considerado uma perda aceitável. Os algoritmos propostos, conseqüentemente, não funcionam bem em sinais de vídeo onde nenhuma perda da informação é permitida, principalmente em sinais de vídeo comprimido, onde uma perda pode se propagar por vários quadros.

Vários outros algoritmos para sincronização de relógios para sinais CBR também são propostos na literatura, entre eles as duas opções oferecidas pelo padrão para a AAL1 do ITU-T [ITU96]. Infelizmente as soluções existentes não se aplicam bem a sinais VBR.

A Solução Proposta

A Figura 4 ilustra o que acontece com o fluxo MPEG gerado pelo codificador até o momento em que ele chega no decodificador, atravessando a rede. O codificador MPEG gera um fluxo VBR controlado (VBR_c), ou seja, a taxa do fluxo é variável, mas com variação controlada por alguns parâmetros funcionais, conforme foi exposto na seção anterior. Após a passagem pela rede, este fluxo passa a ser VBR descontrolado (VBR_d).

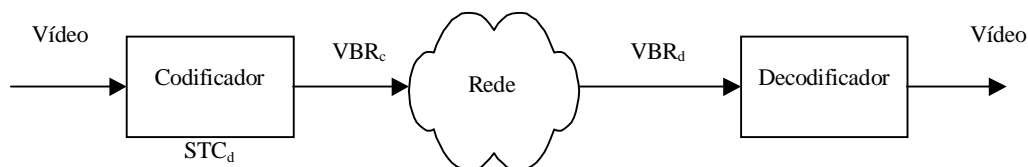


Figura 4 – Trajeto do fluxo MPEG

⁴ Unidade de apresentação seria a unidade da mídia mínima para ser apresentada ao ser humano – no caso de vídeo, trata-se de um quadro; no caso de áudio, é uma amostra

A técnica de transformar este fluxo descontrolado novamente em um fluxo controlado é a contribuição deste trabalho. A idéia é interpor entre a rede e o decodificador um elemento que armazena o fluxo proveniente da rede e entrega-o ao decodificador numa taxa igual, ou pelo menos muito semelhante, à taxa original VBR_c . O fluxo gerado por este elemento, VBR_c' , deve ser tal que obedeça aos requisitos de temporização definidos no padrão MPEG.

O funcionamento deste elemento é, portanto, regido pelas fórmulas de taxa e temporização apresentadas na seção anterior. A recuperação do relógio STC será feita via software, devendo o fluxo ser enviado para a entrada do decodificador segundo o padrão de saída do codificador. Note que, com os dados extraídos do fluxo e o STC, este objetivo é facilmente atingível. Dependendo ainda da situação, pode ser útil também alterar ligeiramente o posicionamento e os valores das referências de relógio.

Este elemento que se interporá entre o decodificador e a rede fará parte, em uma rede ATM, da camada AAL específica para MPEG-2.

O buffer de armazenamento deste elemento deve ser o menor possível, de forma a permitir uma comunicação interativa. Pela Equação (1), verifica-se que é necessário que haja dois valores de PCR disponíveis para que a taxa no trecho possa ser calculada. Desta forma, o atraso mínimo que se pode esperar com esta técnica é o intervalo de tempo entre duas referências de relógio. O padrão define que este intervalo, em fluxos de transporte, não deve ser maior que 100 ms.

Como mencionado nos parágrafos anteriores, para atingir o objetivo de recuperação do sinal VBR_c , basta recuperar a *frequência_do_relógio_do_sistema*, para então satisfazer a Equação (1), através dos outros parâmetros obtidos no próprio fluxo de transporte. Na Figura 5 pode-se ver que esta solução acrescenta mais um relógio no sistema, o relógio STC_a . O relógio do decodificador, STC_d , é automaticamente sincronizado ao relógio STC_a , por exigência de construção do padrão. Resta somente então, a sincronização entre o relógio do codificador e o relógio STC_a .

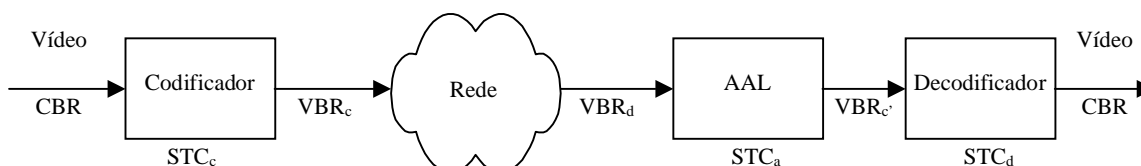


Figura 5 – Relógios do sistema proposto

Esta sincronização pode ser conseguida de duas formas. A primeira é baseando-se no fato de que o vídeo é uma mídia CBR por natureza, ou seja, sua taxa de quadros é constante. O relógio STC_a deve ser tal que a entrega de unidades de representação para o decodificador seja constante ao longo do tempo. Isto implica que o buffer de compensação deve ser tal que apresente ao longo do tempo um valor fixo de quadros (não de bits) armazenados. Assim, STC_a é acelerado ou desacelerado conforme o número de quadros no buffer.

A segunda forma de sincronização seria, de tempos em tempos, enviar uma unidade de representação mais de uma vez para o decodificador, ou então não enviar uma unidade de representação, de acordo com a necessidade de se acelerar ou desacelerar o relógio. Para que isto possa ser implementado, algumas alterações no próprio fluxo são necessárias.

A técnica empregada nesta última solução, apesar desta não ser realmente a melhor solução por introduzir alterações no vídeo original, abre uma outra perspectiva: a da elastização do vídeo. Entenda-se por elastização a apresentação do vídeo em uma duração diferente de sua duração original.

Elastização é muito importante na sincronização de várias mídias em uma apresentação multimídia. No caso específico da implementação em rede ATM, as aplicações poderiam informar sobre a duração esperada para o vídeo a ser exibido, cabendo à camada AAL alterar o fluxo para que o decodificador fosse “enganado” e reproduzisse aquele vídeo em uma duração diferente da original. Manipulando os campos SCR e PTS do fluxo pode-se fazer com que um quadro que tivesse que ser apresentado no instante corrente seja apresentado alguns instantes mais tarde, o que provocaria o efeito de “freeze-frame”, repetindo-se o quadro anterior. Igualmente, poder-se-ia diminuir o valor do PTS de um quadro para que ele fosse apresentado antes do esperado, provocando o efeito de salto de quadros.

Experimento

O ambiente de testes utilizado para transmissões de vídeo MPEG-2 VBR de alta qualidade tem sido a rede ATM que liga o Laboratório TeleMídia a TV Globo, conforme ilustrado na Figura 6. Foram realizadas, durante o Congresso da Sociedade Brasileira de Computação de 1999 (SBC'99), experiências de transmissão multicast envolvendo também o Planetário da Gávea e o Ginásio de Esportes da PUC-Rio, como indicam as setas pontilhadas na mesma figura. Versões dos algoritmos propostos no presente trabalho têm sido usadas para transmissões variando entre 8 e 12 Mbps.

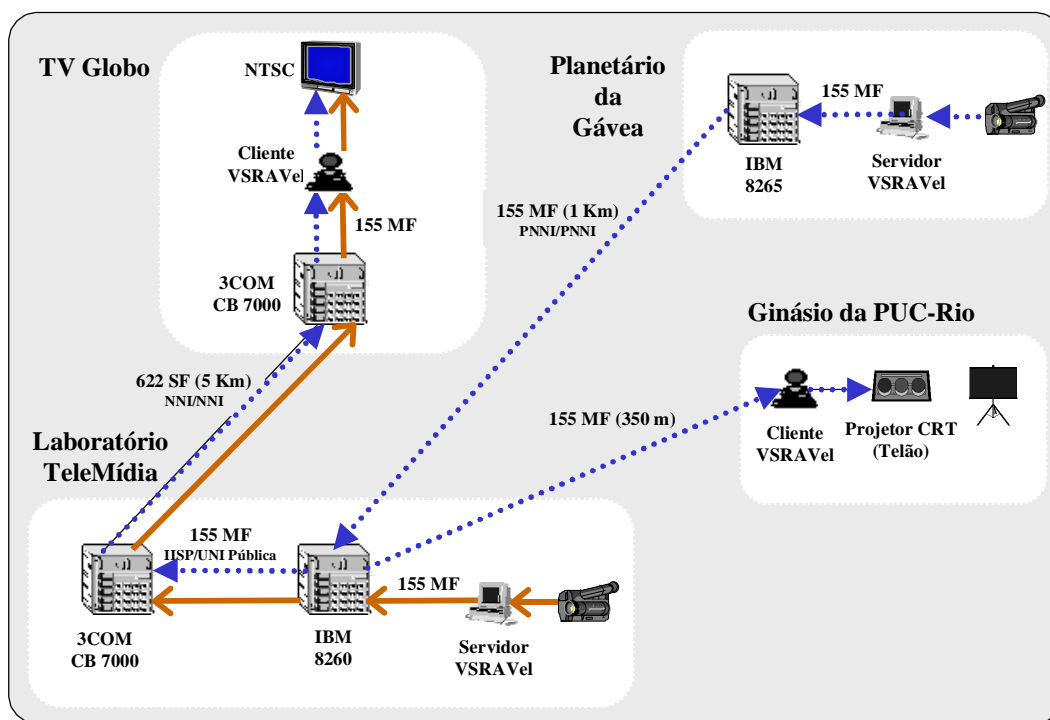


Figura 6 – Experiências de transmissão de MPEG-2 VBR sobre uma rede ATM.

Durante os testes tem sido também promovida a introdução proposital de erros na transmissão, para que o estudo sobre seu efeito possa ser direcionado.

5 – Conclusão

Este artigo apresentou o trabalho que vem sendo desenvolvido no Laboratório TeleMídia da PUC-Rio com relação a transmissão de vídeo VBR em redes com retardos variáveis. O foco deste trabalho foi a sincronização dos relógios do codificador e decodificador, mas existem vários outros tópicos de

interesse em investigação. A taxa de erros introduzida pela rede também pode ser determinante na qualidade final do vídeo. Embora existam estudos que indicam algumas diretrizes que devam ser tomadas em caso de erros, muito pouco se sabe da eficiência dessas diretrizes e da relação entre a taxa de erro e a qualidade do vídeo, tanto na adequação quanto na ignorância a essas diretrizes.

O objetivo final do projeto é implementar a transmissão de vídeo MPEG-2 sobre uma rede ATM. O controle da taxa do vídeo é necessário, para diminuir a variação estatística do retardo. Correção de erro também é desejada, uma vez que, devido a características do MPEG-2, um erro em um único bit em uma posição sensível pode provocar um erro na imagem que se propaga por vários quadros. Paralelo a este objetivo, é traçado um segundo: a elastização da apresentação do vídeo.

Este segundo objetivo relaciona-se com o primeiro no sentido em que é definida uma única camada, no receptor do vídeo, que terá como função prover o fluxo MPEG para a aplicação, e receber indicações sobre a o tempo de exibição do vídeo (sua elasticidade). Assim, se, por exemplo, um vídeo de 30 segundos for exibido em 29 segundos, nem todos os quadros serão exibidos, o que permite que sejam tolerados alguns erros, ou que haja uma realimentação para o emissor para que diminua o número de quadros enviados, aliviando a carga na rede.

Algoritmos de elasticidade serão muito importantes na introdução de interatividade a um sinal de vídeo. Vídeos interativos formarão a base da TV interativa e se constituem em uma de nossas linhas de pesquisa conjunta com a TV Globo. Neste sentido, o presente trabalho é um importante passo inicial.

Agradecimentos: os autores gostariam de agradecer a toda equipe do Laboratório TeleMídia e da TV Globo que viabilizaram os teste realizados, em especial ao colega Rogério Rodrigues.

Bibliografia

- [Adan96] Adanez, X. G. "Designing New Network Adaptation and ATM Adaptation Layers for Interactive Multimedia Applications". Swiss Federal Institute of Technology, Lausanne, 1996.
- [ISO93] ISO/IEC 11172. "Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s", 1993.
- [ISO94] ISO/IEC IS 13818 | ITU-T Recommendation H.222.0. "Information Technology – Generic Coding of Moving Pictures and Associated Audio", 1994.
- [ISO94a] ISO/IEC IS 13818-1 | ITU-T Recommendation H.222.0. "Information Technology – Generic Coding of Moving Pictures and Associated Audio - Part 1: Systems", 1994.
- [ISO94b] ISO/IEC IS 13818-2 | ITU-T Recommendation H.262. "Information Technology – Generic Coding of Moving Pictures and Associated Audio - Part 2: Video", 1994.
- [ITU90] International Telecommunication Union (Formerly CCITT Recommendation H.261). "Codes for audiovisual services at px64 kbit/s". Recommendation, Geneva, 1990.
- [ITU93] ITU-T Recommendation T.81. "Joint Photographic Experts Group", 1993.
- [ITU94] International Telecommunication Union. "Vocabulary of Terms for Broadband Aspects of ISDNs". Recommendation I.113, November 1994.
- [ITU96] International Telecommunication Union. "B-ISDN ATM Adaptation Layer Type 1 Specification". Recommendation I.363.1, 1996.
- [SoLC95] Soares, L. F. G., Lemos, G., Colcher S. "Redes de Computadores: das LANs, MANs e WANs às redes ATM". Editora Campus. 2ª edição, 1995.