

# Utilizando Rede Neural para o Ajuste da Qualidade do Vídeo em um Sistema sob Demanda

**Regina M. Silveira<sup>1</sup> e Wilson V. Ruggiero**

LARC – Laboratório de Arquitetura e Redes de Computadores  
PCS - Escola Politécnica da Universidade de São Paulo  
Av. Prof. Luciano Gualberto, Trav.3, nº 158, 05508-900 São Paulo, SP, Brasil  
{regina, wilson}@larc.usp.br

**Resumo:** Considerando a transmissão de vídeo em rede ATM a partir de um sistema de Vídeo sob Demanda, propomos um Servidor de Qualidade (SoQ) que tem como função ajustar a taxa de transmissão do vídeo à banda disponível na rede. O SoQ utiliza uma rede neural “backpropagation” para ajustar os coeficientes de quantização MPEG de forma não linear. Aqui é mostrada a especificação do SoQ e os resultados obtidos a partir da implementação de seu módulo principal.

**Palavras Chave:** MPEG, quantização, Vídeo sob Demanda, ATM, rede neural.

**Abstract:** Considering the video transmission over ATM network coming from Video on Demand System, we propose a Server of Quality (SoQ) that has the main function to fit the video transmission rate to the available network bandwidth. The SoQ utilizes a “backpropagation” neural network to fit the MPEG quantization coefficients in a no linear way. Here are shown the SoQ specification, and the results obtained from its main modulo implementation.

**Key words:** MPEG, quantization, Video on Demand, ATM, neural network.

## 1. Introdução

As redes ATM (*Asynchronous Transfer Mode*)[1][2] tem se destacado no panorama tecnológico atual por permitir a transmissão dos mais variados tipos de dados simultaneamente. Esses dados podem necessitar requisitos diferentes (como largura de banda, atraso e sincronismo) para serem transmitidos, como é o caso de vídeos, áudios, voz e dados. Consequentemente, a rede ATM torna-se uma das opção tecnológica para aplicações multimídia em tempo-real que utilize o conceito cliente/servidor e que oferece mecanismos para controle efetivo da qualidade de serviço (QoS).

O projeto RMAV-SP (Rede Metropolitana de Alta Velocidade de São Paulo) permitiu a expansão da área de abrangência de uma rede de alto desempenho, permitindo a interconexão do LARC (Laboratório de Arquitetura de Redes e Computadores) com outras instituições de ensino e pesquisa, permitindo a utilização de novas aplicações que requisitem qualidade de serviço (QoS).

Em um sistema de vídeo sob demanda (VoD), o usuário deve poder acessar de sua casa ou trabalho, uma base de dados remota e selecionar o material que deseja assistir. O vídeo selecionado será transmitido em tempo-real permitindo que o usuário que o requisitou possa usar controles que atuarão sobre o vídeo no sentido de pausar, retornar, ir a frente, ou parar a transmissão. Para isso é necessário que a rede ATM esteja apta a garantir a transmissão do vídeo solicitado com a qualidade desejada.

---

<sup>1</sup> Bolsista do CNPq-Brasil.

O objetivo aqui é determinar um método eficiente de diminuir, quando necessário, a taxa de bits dos arquivos de vídeos que serão transmitidos em tempo-real pela rede ATM. Para isso, considerou-se o sistema de Vídeo sob Demanda implantado no LARC, onde os vídeos a serem transmitidos já se encontram previamente codificados em MPEG-1 ou MPEG-2 (com qualidade máxima permitida pelo codificador) e armazenados em um servidor de vídeo ou em uma memória terciária. Durante a negociação para o estabelecimento da conexão ATM, é feita a verificação dos parâmetros de qualidade de serviço (QoS) que a rede poderá atender. Para que essa negociação seja otimizada, é feita a especificação de um “Servidor de Qualidade” (SoQ), que intermedia essa negociação, e reajusta os coeficientes de quantização do vídeo codificado em MPEG a partir de uma rede neural, possibilitando um ajuste não uniforme, levando em consideração as características perceptivas do ser humano e ainda se beneficiando do auto-aprendizado da rede neural para obter o ajuste mais eficiente. O benefício da utilização deste servidor de qualidade é o de dar ao usuário a opção de ter acesso ao material desejado, porém com uma qualidade diferente da original, em momentos em que a rede ATM não possa garantir a transmissão do material com a qualidade com que ele foi previamente codificado. Evitando assim, que o usuário do curso remoto fique desmotivado pela falta de capacidade de vazão da rede ATM e receba o material de vídeo com a melhor qualidade possível, face as restrições momentâneas de banda de transmissão. A validade de tal proposta foi verificada a partir da simulação da rede neural definida, utilizando os coeficientes da matriz de quantização da compressão MPEG.

## **2. Aspectos da Codificação e Transmissão de Vídeo**

As aplicações que utilizam informações "ao vivo", como vídeo-fone, vídeo-conferência, podem, com maior facilidade, utilizar a compressão MPEG variável (VBR MPEG), já que a banda disponível é conhecida no momento da compressão. Em um sistema de Vídeo sob Demanda, normalmente as informações (vídeos) são comprimidas previamente, impossibilitando a avaliação exata da banda de transmissão disponível no momento em que esta será efetuada.

O método de quantização especificado pelo padrão MPEG com taxa de bit constante, opera de forma adaptativa, onde a partir da alteração da escala (ou passo) de uma variável auxiliar na quantização (“quantizer scale”), é feita, de maneira linear, a alteração da tabela de quantização. Esse procedimento é utilizado durante a codificação MPEG dependendo da informação do controle da ocupação do “buffer” do codificador. Desta forma é possível manter a taxa de bit constante através da variação da qualidade da imagem. Este método de controle não leva em consideração a característica posicional dos coeficientes DCT, degradando a qualidade da imagem.

Muitos trabalhos já foram, e continuam sendo, desenvolvidos na tentativa de elaborar uma forma alternativa e eficiente de fazer o controle da taxa de transmissão, mantendo um compromisso com a qualidade da imagem, já que os dois fatores são mutuamente exclusivos, e, como consequência, obter uma otimização da banda utilizada pelas aplicação que envolvam transmissão de vídeo. Y. Saw[3] propõe um modelo prático de controle da quantização, por meio de funções não lineares (sigmoidal e unimodal), que apresentam um controle da taxa de transmissão superior ao utilizado pelo padrão MPEG, mantendo uma qualidade similar. Em [4], Saw propõe um outro método, agora para CBR MPEG, onde o controle da banda utilizada é baseado no perfil da atividade espaço-temporal do vídeo, usando uma função radial. E ainda

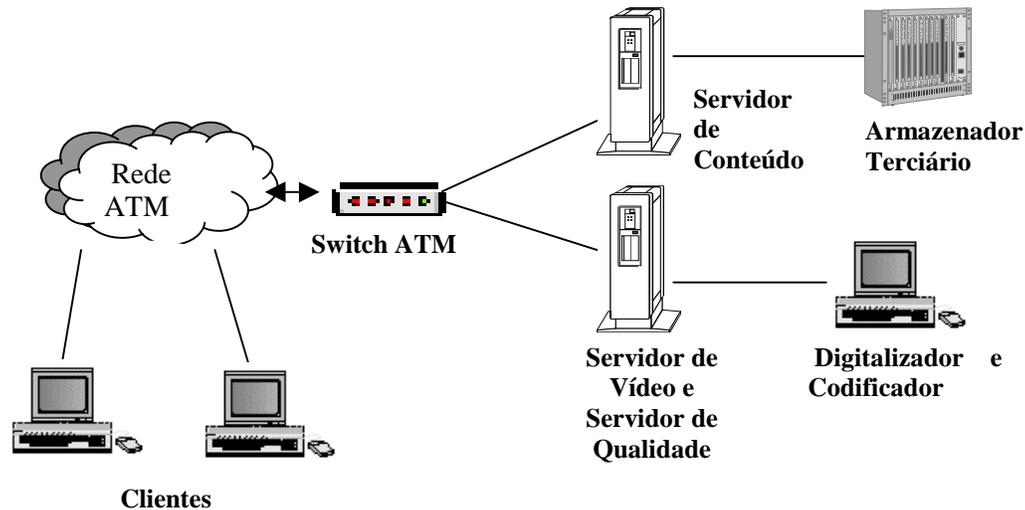
em [5], Reininger utiliza um método que verifica continuamente e parametriza o nível de ocupação de um buffer criado na interface de rede, e, a partir desta medida dinâmica é feita a alteração linear dos coeficientes de quantização MPEG.

No entanto, essas técnicas foram desenvolvidas para serem utilizadas durante a codificação, ou seja, para vídeos com transmissão ao vivo, baseados na resposta da rede. Para o caso de vídeos já previamente codificados, como ocorre em sistemas de vídeo sob demanda, poucos trabalhos foram desenvolvidos. Zhang[6] define um algoritmo de otimização da seleção de quadros MPEG ou JPEG, que devem ser descartados para que o fluxo de dados se adeque à banda de transmissão disponível na rede. Sabe-se que o descarte de quadros provoca degradação da imagem, causando a incomoda sensação visual de que a imagem treme ou pisca. Isto é normalmente agravado em seqüências de imagens que apresentam movimentos rápidos. Outra proposta é feita por Eleftheriadis[7], com a técnica denominada DRS (Dynamic Rate Shaping) onde é utilizado um filtro entre a saída do codificador e a interface de rede, que procura ajustar o perfil do tráfego à QoS oferecida pela rede. Esse ajuste é feito pela eliminação de alguns coeficientes da transformada discreta de cosseno. Trata-se de um método bastante rápido, pois não entra no mérito do processo de codificação, no entanto, a eliminação arbitrária desses coeficientes causará degradação da qualidade. Apesar de Eleftheriadis ter provado que a relação sinal-ruído do vídeo decodificado permanece adequada após a utilização do método, essa relação não é um parâmetro adequado e suficiente para garantir boa qualidade da imagem. Uma outra proposta feita por Reininger [8] se baseia em um esquema onde o servidor de vídeo armazena várias versões de um determinado vídeo, com diferentes resoluções. Esses vídeos, codificados com MJPEG, são submetidos a um controlador da taxa de bit do vídeo, fazendo o ajuste do vídeo de acordo com a banda disponível. Esse ajuste é feito alterando os coeficientes de quantização de forma linear.

Como podemos ver, nenhuma dessas propostas levam em consideração os aspectos de percepção do sistema visual humano. No método proposto aqui, esses aspectos são respeitados ao ser considerado as características posicionais dos coeficientes DCT, por meio da alteração, de forma não linear, dos coeficientes de quantização do padrão MPEG. Este método poderá ser aplicado à vídeos codificados tanto com VBR MPEG-1 como com VBR MPEG-2, já que ambos utilizam o mesmo esquema de quantização, permitindo a criação de “streams” de vídeo escaláveis. Da mesma forma, ele poderia ser aplicado também à vídeos codificados com H.261[9] e MJPEG[10] (Motion JPEG) que sejam baseados em taxas de bit variável.

### **3. O Servidor de Qualidade - SoQ**

O Servidor de Qualidade proposto aqui, deve ser implementado de maneira a otimizar a negociação da conexão para a transmissão de vídeo já previamente codificados em um sistema de VoD. A proposta prevê que durante a negociação, caso a rede não possa atender a descrição de tráfego (DoT) do vídeo a ser transmitido, o SoQ (Servidor de Qualidade) atue procurando alterar os parâmetros de descrição, até que a rede possa atender a requisição de transmissão, diminuindo o menos possível a qualidade do vídeo. Implantando o SoQ, teremos uma nova arquitetura para o sistema LDMD, como é mostrado a figura 1. O SoQ deve ser implementado na mesma máquina do servidor de vídeo, agilizando o acesso ao vídeo armazenado.



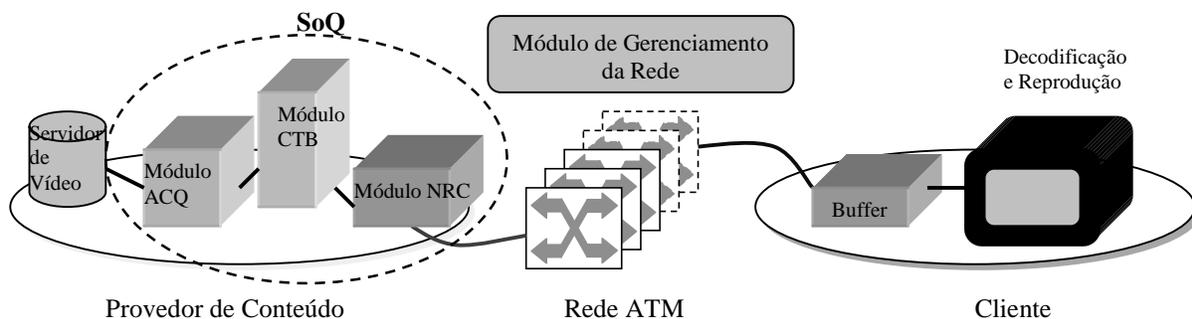
**Figura 1: Arquitetura do sistema LDMD com a implantação do SoQ.**

O sistema será acionado quando um cliente se conecta através de um computador (máquina cliente) ao servidor de vídeo, por meio da rede ATM. A máquina cliente deve ser basicamente formada por um “buffer” e pelos módulos de decodificação MPEG e reprodução do vídeo decodificado. O “buffer” deve armazenar os pacotes enviados pela rede até que o decodificador esteja pronto para processá-lo. Quando o sinal de vídeo é decodificado ele é passado ao módulo de reprodução para ser apresentado no monitor. A capacidade desses módulos variam consideravelmente de acordo com a arquitetura da máquina utilizada. Caso esta máquina tenha uma placa de decodificação e reprodução, seu desempenho deverá ser melhor se comparado as que fazem tais operações por software.

Os vídeos são armazenados no servidor de vídeo, já codificados com qualidade superior, isto é, com a melhor qualidade especificada pelo padrão de compressão utilizado. O controle da taxa de bit feita pelo SoQ permite uma escalabilidade da qualidade do vídeo a ser transmitido. Esse controle pode ser feito em diferentes níveis. Aplicando o método de controle somente sobre os quadros I teremos uma suavização (smooth) do perfil do tráfego. Desta forma, o parâmetro PCR (taxa máxima de células) de descrição do tráfego (DoT) da tecnologia ATM será diminuído, e conseqüentemente, os parâmetros MBS (tamanho máximo da rajada) e SCR (taxa média de células) também serão suavizados. Aplicando o método de controle sobre os quadros I e P ou I, P e B, obteremos uma redução mais significativa da banda necessária para a transmissão dos dados. Para que o ajuste seja feito adequadamente, o servidor de qualidade inclui um módulo de controle da taxa de bit (CTB), que deverá avaliar em que nível o SoQ deverá atuar para encontrar a taxa de bit desejada e o módulo de ajuste dos coeficientes de quantização (ACQ). O terceiro módulo que compõe o SoQ se destina a negociação e renegociação da conexão (NRC) para a transmissão pela rede ATM. O esquema da figura 2 ilustra os principais módulos de cada elemento do sistema.

### **3.1. Negociação da Conexão**

No esquema proposto aqui, o cliente deve inicialmente se conectar ao sistema por meio de uma conexão com categoria de serviço ABR ou UBR. Após escolher o vídeo a ser transmitido, o cliente se desconecta e uma nova conexão será estabelecida entre o SoQ e o cliente, onde, desta vez, o servidor solicita a conexão. Desta forma, o SoQ terá controle sobre



**Figura 2: Sistema onde opera o SoQ decomposto em módulos.**

a conexão. A conexão do sistema seguirá o esquema mostrado na figura 3, onde caso a solicitação não é atendida devido a falta de recursos ou de garantia de QoS, o SoQ entra em ação propondo valores reduzidos para os parâmetros de tráfego do vídeo pré-codificado. Ao se estabelecer a conexão, o vídeo será parcialmente decodificado e recodificado de outra forma, em tempo-real, afim de obter a taxa de transmissão máxima definida durante a negociação para o estabelecimento da conexão.

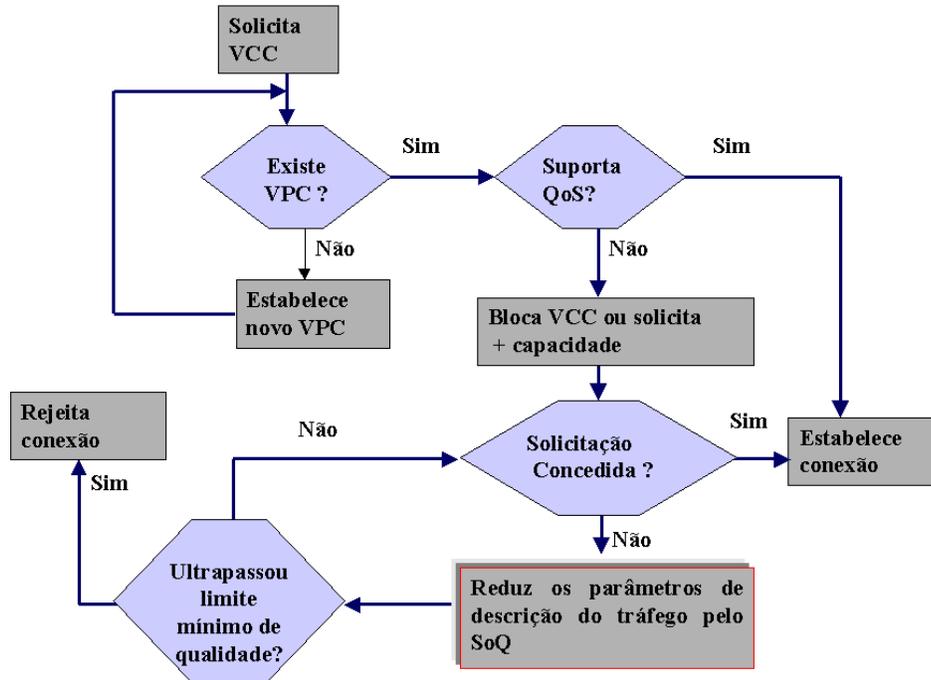
A partir deste esquema de negociação é possível elaborar propostas para a implementação de alocação dinâmica estatística e determinística, possibilitando assim a renegociação da conexão ATM[22].

### 3.2. Ajuste dos Coeficientes de Quantização

Apesar de haver muito menos flexibilidade no controle da taxa de transmissão do vídeo pré-codificado, é possível utilizar algumas técnicas para fazer esse controle. Tais técnicas são [5]:

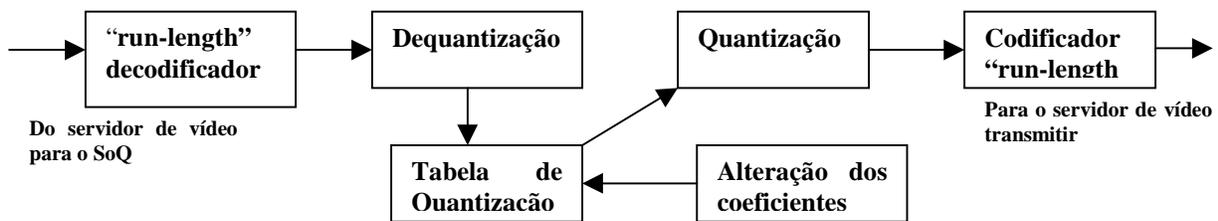
- Suavização do tráfego: redução dos picos máximos de transmissão introduzindo atrasos, e armazenamento em buffer entre os processos de codificação e decodificação [11].
- Escalonamento da resolução da frequência: redução da taxa de bits de vídeos precodificados através do descarte seletivo dos coeficientes DCT[4, 7].
- Escalonamento da resolução temporal: consiste em fazer o chaveamento entre múltiplas seqüências de vídeo precodificadas com diferentes qualidades. Em MPEG o chaveamento entre seqüências pode ser feito no final de um GOP (grupo de quadros), após a introdução de um 'header' da nova seqüência[8].
- Escalonamento da resolução espacial: consiste no descarte de alguns quadros (como por exemplo quadros B), reduzindo assim a taxa de transmissão[6].

Na implementação do SoQ todas essas técnicas podem ser consideradas. No entanto, a principal contribuição deste trabalho se dá com a proposta de um método ainda não explorado para a implementação da técnica de escalonamento da resolução da frequência. O processo básico consiste em desfazer as últimas etapas da codificação MPEG até obter os coeficientes resultantes da transformada discreta de coseno. Esses coeficientes são então "re-quantizados", utilizando-se coeficientes de quantização mais elevados, para obter a taxa de transmissão desejada. Os processos envolvidos são mostrados na figura 4. Resultados em [4] mostram que o processo básico consegue atingir redução da ordem de 20% da taxa de transmissão sem degradação significativa da imagem. Considerando que o sistema LDMD é destinado a prover



**Figura 3: Diagrama de estabelecimento da conexão utilizando o SoQ.**

informações técnicas e não conteúdo de entretenimento, podemos considerar que em benefício da transmissão, poderá ser tolerado alguma degradação da imagem, sem prejuízo da qualidade da informação. Por isso, nesse sistema prevemos a possibilidade de uma economia de banda de transmissão superior a mostrada por Reininger, ou seja, mais de 20%.



**Figura 4: Esquema de escalonamento da resolução da frequência.**

A inovação desta proposta está em utilizar uma rede neural, para fazer a alteração não linear, dos coeficientes de quantização. Esta proposta, teve como estímulo os grandes benefícios obtidos com sistemas que utilizam redes neurais. Dentre esses benefícios foram considerados o fato das redes neurais serem especialmente válidas para manipulação de dados não lineares e sua capacidade de auto-aprendizado.

As redes neurais tem sido largamente utilizadas para redução de ruído, reconhecimento, complementação e classificação de padrões[12], como voz e imagens, para previsão de comportamento de sistemas não lineares, para controle e otimização de sistemas, entre outras aplicações. O tratamento de imagens é uma das grandes aplicações das redes neurais. Particularmente em vídeo digital, as redes neurais são empregadas nas técnicas caracterizadas pela “codificação por previsão”[13]. Em vídeos codificados com MPEG, alguns trabalhos utilizam redes neurais para fazer a classificação dos macroblocos[14] e para prever o perfil do

tráfego do vídeo em tempo-real[15]. Mais recentemente, estão sendo desenvolvidas e publicadas pesquisas relacionadas a utilização de redes neurais para prever o comportamento e otimizar as redes ATM. Como exemplo, pode-se citar trabalhos que utilizam rede neural para o modelamento, em tempo-real, da descrição de tráfego de fontes heterogêneas. Obtendo o perfil da fonte de tráfego da transmissão em rede ATM, é possível otimizar o Controle da Admissão de Conexão[16, 17, 18].

#### **4. Redes Neurais Artificiais**

As redes neurais artificiais (RNA) procuram recriar o funcionamento de nossos neurônios cerebrais por meio da interconexão de nós denominados neurônios. Cada neurônio recebe um conjunto de sinais de entrada, e dependendo da amplitude (ou pesos) desses sinais, os neurônios obtêm um sinal de saída de determinado peso. Um neurônio, de posição genérica em uma rede, tanto aceita entradas de outros neurônios como emite saídas para outros neurônios. As RNA podem ser construídas com diferentes topologias e regras que juntas definem um modelo (ou paradigma) para a rede [19]. Algumas topologias estabelecem a forma de interconexão entre os nós, outras definem a disposição entre suas camadas de neurônios. Em uma rede de camadas, alguns neurônios são especializados em receber sinais e outros em emitir sinais de saída, e apenas os neurônios das camadas intermediárias (chamadas de “hidden”) é que mantém a generalidade.

As interconexões entre os neurônios são denominadas sinapses. Uma das principais características das RNA é sua capacidade de aprendizado. Durante o processo de aprendizado os pesos do sinal de entrada e de saída de um neurônio são ajustados. Os neurônios tornam-se capazes de tomar decisões futuras baseadas na experiência passada, isto é, no ajuste dos pesos.

Considerando todas as camadas existentes na rede os pesos vão se propagando ao ser aplicado uma função de soma que pode ser genericamente representada por:

$$\mu_j = \sum x_i w_{ij} - \theta_j$$

onde  $w_{ij}$  representa o peso da sinapse entre o neurônio  $i$  e o neurônio  $j$ , o  $x_i$  representa o sinal de saída do neurônio  $i$  e  $\theta_j$  representa o termo de polarização. Portanto, a entrada do neurônio  $j$ , vindo de  $i$  é indicado por  $x_i w_{ij}$ . A soma obtida é chamada de ativação do neurônio. Esta ativação pode ser positiva, zero ou negativa, já que tanto os pesos das sinapses como os valores de entrada podem assumir valores positivos ou negativos. Quando um peso fornece ao sistema uma contribuição positiva dizemos que a ativação representa um estímulo, em contra partida, quando a contribuição é negativa a ativação representa uma inibição.

Entre as principais vantagens da utilização de redes neurais podemos citar o aprendizado adaptativo, a auto-organização, operação em tempo-real, possibilidade de implementação tanto em hardware como em software, facilidade de inseri-la em tecnologias já existentes, capacidade de operar com processos não lineares.

Como mencionado anteriormente, as redes neurais podem ser construídas com diferentes topologias e regras que juntas definem um modelo para a rede. Um dos modelos mais utilizados em várias classes de problemas é a RN “Backpropagation”, por se tratar de uma rede caracterizada por treinamento supervisionado. O aprendizado supervisionado implica em

termos um conjunto de “bons” padrões associados ao treinamento da RN. Veremos mais a frente que este é o modelo que se ajusta melhor ao sistema aqui proposto. Nesta seção vamos detalhar o funcionamento deste tipo de rede.

#### 4.1. Aplicando Rede Neural para o Controle da Banda Utilizada

Uma RNBP (rede neural backpropagation) consiste de neurônios que formam uma soma de pesos em suas entradas e passam essa soma de forma não linear. O procedimento de BP minimiza os critérios de erro da rede. Um detalhamento maior do algoritmo deste tipo de rede pode ser visto em [12, 19 e 22]. As RNBP são largamente utilizadas para detecção de padrões de imagem e de áudio. Elas podem também ser aplicadas em previsões do comportamento de sistemas, onde dado um comportamento a rede é treinada para se ajustar aos variações dos seus fatores de entrada.

no SoQ, a rede neural BP será treinada de forma que dado uma situação de entrada em que houve uma alteração da banda disponível para a transmissão do sinal de vídeo, a rede deverá ajustar os pesos embutidos em cada transição da rede, para que se obtenha o melhor ajuste dos coeficientes de quantização MPEG. A aplicação de rede neural para a redução da banda de transmissão original do vídeo visa:

- Criar um método que leve em consideração as características perceptivas do ser humano, alterando os coeficientes de quantização de forma diferenciada dependendo de sua posição dentro da matriz de quantização, já que os coeficientes de quantização MPEG apresentam uma distribuição pré-definida de acordo com a sensibilidade do sistema de visão humana.
- Utilizando o auto aprendizado da rede, obter a redução da banda de forma eficiente. Isto é, obter o melhor ajuste da banda com a menor diminuição possível da qualidade.

##### 4.1.1. Formulação do Modelo

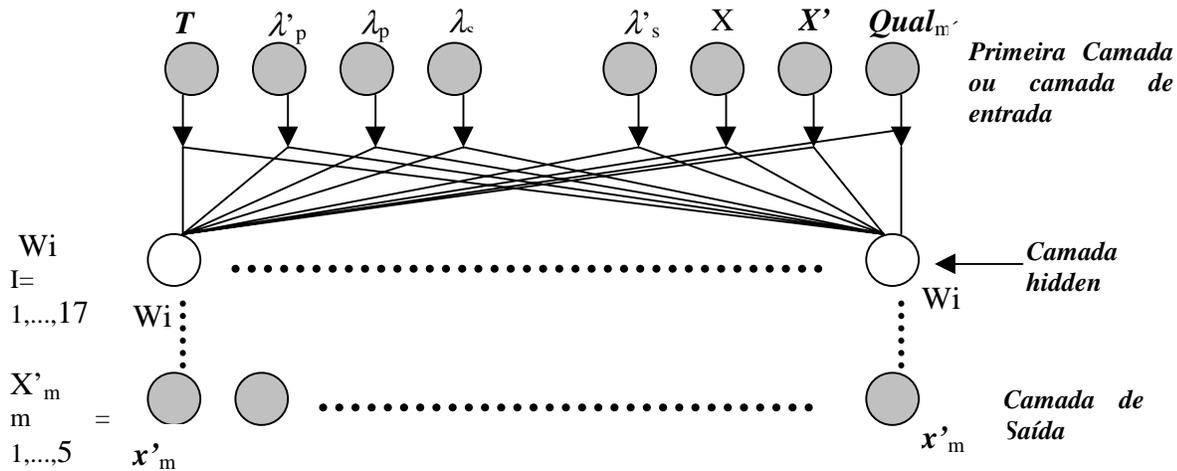
Para que seja possível implementar e treinar a rede neural “backpropagation” é necessário definir o modelo e a topologia da rede, assim como as características da rede neural que será implementada.

Para facilitar o dimensionamento dos valores de saída da rede, ela será modelada para prever os fatores multiplicativos a qual os coeficientes da matriz de quantização deverá ser submetida para que o vídeo comprimido se ajuste a banda disponível pela rede.

É necessário definir quais os fatores de entrada devem alimentar a rede para que ela nos forneça a porcentagem de variação dos coeficientes de quantização. Os fatores considerados para o treinamento da rede são:

- Um vetor de cinco elementos de entrada:  $[x_1, x_2, x_3, x_4, x_5]$ ;
- Um vetor de cinco elementos esperados:  $[x'_1, x'_2, x'_3, x'_4, x'_5]$ ;
- O tipo de quadro em questão: T (I, P ou B);
- Os valores de descrição original do tráfego:
  - Taxa Máxima de bit (representado aqui por  $\lambda_p$ );
  - Taxa média de bit ( $\lambda_s$ );
- Os valores de descrição de tráfego negociados pela rede:
  - Taxa Máxima de bit ( $\lambda'_p$ );
  - Taxa média de bit ( $\lambda'_s$ );

- Os parâmetros mínimos de qualidade estipulados ( $Qual_{\min}$ ).



**Figura 5: Topologia proposta para a RN do SoQ.**

A rede neural deve apresentar a topologia mostrada na figura 5 onde os sinais de entradas são inseridos na primeira camada. Os índices  $i$  e  $j$  indicam a posição do coeficiente na tabela de quantização. Os valores obtidos na camada de saída, indicados por  $Q_{mn}'$ , devem representar os fatores ao qual a matriz de quantização deve ser multiplicada, para que o tráfego do sinal de vídeo se ajuste a banda disponível na rede. Entre a primeira camada “hidden” e a camada de saída da rede podem haver múltiplas outras camadas “hidden” que permitam um ajuste mais refinado aos pesos que serão submetidos os coeficientes.

O processo de definição de uma rede neural inclui seu treinamento, teste e refinamento do modelo. Depois disso, a rede neural será inserida no código do SoQ, e será chamada por funções criadas especificamente para isso. O auto-aprendizado da rede é feito com o valor do resultado final da compressão. Isto é, após obter os valores dos coeficientes ajustados, o quadro passa novamente pelas etapas finais de compressão MPEG. O resultado final, ou seja, a quantidade de bits necessários para codificar o quadro, será enviado para a rede neural, que utilizará esse valor em seu aprendizado.

A verificação da validade de utilização da rede neural para o ajuste dos coeficientes de quantização MPEG, foi feito por meio dos resultados obtidos com um simulador de rede neural comercial desenvolvido por Craig Jensen, denominado QwikNet32 [20]. O treinamento da rede neural criada no simulador, demonstrou que os dados de entrada convergem para os dados de saída esperados[25].

Para garantir que, com a utilização do método de ajuste dos coeficientes de quantização, não haja uma degradação excessiva da imagem, deve ser definido um índice de qualidade. Esse índice foi estipulado a partir de um teste subjetivo da qualidade de seqüências previamente preparadas[24]. Assim, a definição da qualidade mínima para esse sistema foi feito a partir de um levantamento estatístico sobre a aceitação e tolerância da perda de qualidade em vídeos, utilizando clips de testes (que devem ser clips clássicos de testes de qualidade), de maior e menor complexidade computacional. Foram utilizados vários arquivos do mesmo clip, com qualidades diversas obtidas com o aumento progressivo dos coeficientes de quantização na codificação MPEG VBR. A partir deste estudo foi definido a mínima qualidade média

aceitável, criando assim um índice de qualidade variando entre 1 e 5, onde o índice 5 indica a melhor qualidade e 1 a menor.

## 5. Resultados Obtidos com a Implementação do ACQ

Após a implementação do módulo de Ajuste dos Coeficientes de Quantização, foram feitos teste para validar sua utilização no controle da taxa de bit do vídeo a ser transmitido, quando integrado a um sistema de vídeo sob demanda.

Durante esta tomada de dados, o sistema sofreu algumas evoluções que serão descritas à seguir. Outras otimizações ainda podem ser feitas, mas não foram implementadas por fugir ao escopo do trabalho, que visava principalmente validar a utilização da rede neural “backpropagation” para fazer o ajuste necessário.

Foram utilizados, para fazer os testes do ACQ, as mesmas seqüências usadas no teste subjetivo de qualidade, ou seja, os vídeos “Flowers”, “Bike”, “Siege” e “Tennis”, que apresentam as seguintes características:

Duração: 5 segundos;

Resolução: 352x240;

Número de quadros por segundo: 30 fps;

Número total de quadros: 150;

Amostragem utilizada: 4:2:0 (4 amostras de luminância e 2 amostras de crominância).

Os primeiros resultados obtidos indicam que a rede ajustou-se melhor a vídeos de melhor qualidade e com menos movimento. Isso pode ser justificado pela alta taxa de bit obtido para o vídeo “Flowers” quando este foi codificado com a melhor qualidade possível pelo codificador MPEG VBR. Esse fator foi o de maior valor inserido na rede neural. A variação de movimento do vídeo é um fator muito importante para o algoritmo de codificação do padrão MPEG. O perfil da atividade média dos quatro vídeos utilizados para esse estudo é apresentado na figura 6.

Esse gráfico mostra uma atividade bem linear e crescente em uma das curvas, que caracteriza a tomada panorâmica que aparece em “Flowers”. As demais curvas apresentam uma curva muito menos comportada, que representam os cortes e cenas dinâmicas.

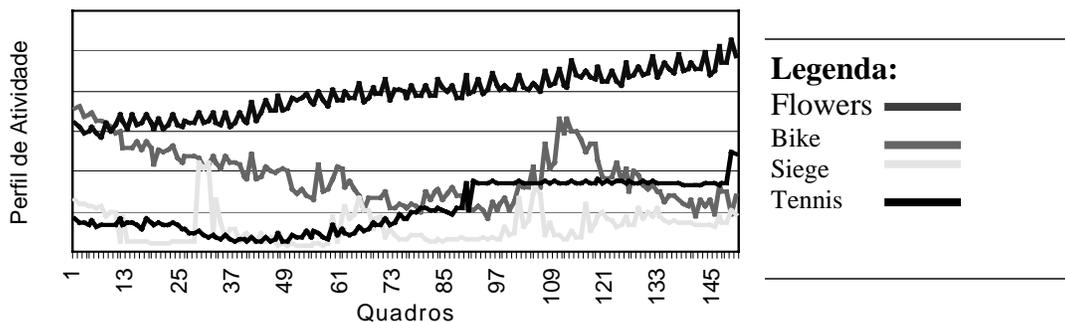
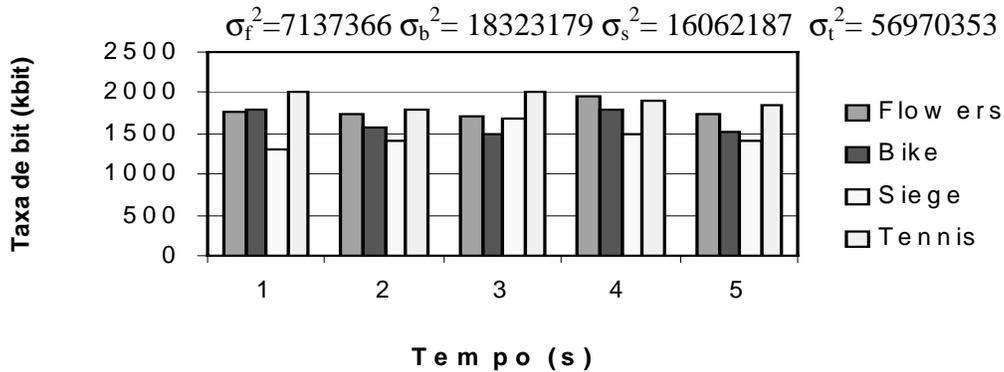


Figura 6: Perfil da atividade de cada um dos vídeos utilizados.

Para resolver o problema da falta de sensibilidade do ACQ a atividade do vídeo é necessário inserir como parâmetro da rede neural o grau de atividade do vídeo, e refazer o treinamento com essa informação adicional. É possível calcular a atividade espacial média de uma cena, comparando os valores dos pixels de um dado quadro, com os valores dos pixels de mesma posição do quadro anterior. Um novo treinamento da rede neural foi feita, acrescentando agora um fator que expresse o grau de atividade do vídeo. Todos esses treinamentos fazem parte do refinamento do modelo da rede neural, comum ao procedimento de teste de todo sistema que a utiliza.



**Figura 7: Variação da taxa de bit para os quatro vídeos quando recodificado para 2Mbps.todo sistema que a utiliza.**

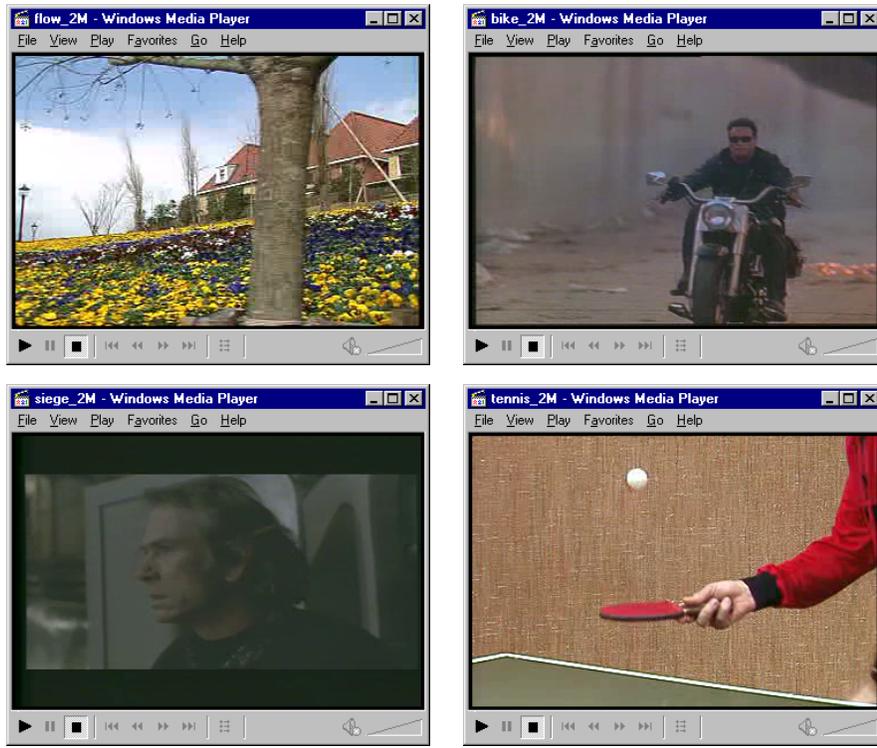
Os resultados obtidos com essa rede podem ser observados na figura 7, onde é mostrada a variação da taxa de bit para os quatro vídeos considerados. Nesta recodificação, a taxa de bit média inserida como parâmetro foi de 2 Mbps. Pode-se observar que os quatro vídeos apresentam taxas bem próximas da taxa solicitada ao ACQ. Considerando que trata-se da taxa média de bits, os erros encontrados são factíveis para um sistema que irá operar sobre uma rede ATM. As imagens da figura 8 mostram cenas desses vídeos onde podem ser observadas a qualidade de cada um deles.

Podemos observar que todos os vídeos foram ajustados para uma taxa de bit próxima da solicitada. Como já foi mencionado, estes resultados poderiam ser melhorados com o treinamento individual da rede para os parâmetros de cada vídeo. Porém, este treinamento não foi feito por considerar os resultados já bastante satisfatórios para mostrar a viabilidade da utilização da técnica proposta para fazer o ajuste dos coeficientes de quantização MPEG.

No esquema de alocação dinâmica de banda, proposto para a tecnologia ATM, após o estabelecimento da conexão, é possível fazer renegociações da conexão, com descritores de tráfego (DoT) e qualidade de serviço (QoS) diferentes dos indicados na negociação inicial. Esta renegociação é feita quando, devido a congestionamento ou liberação de recurso, o Gerenciador da Rede pode fornecer mais ou menos recurso para uma dada conexão. Este esquema de alocação é uma das principais inovações que esta tecnologia trouxe as transmissões de dados orientada a conexão, e esta inovação está sendo atualmente incorporada também, em protocolos de comunicação para tecnologias baseadas em comutação IP.

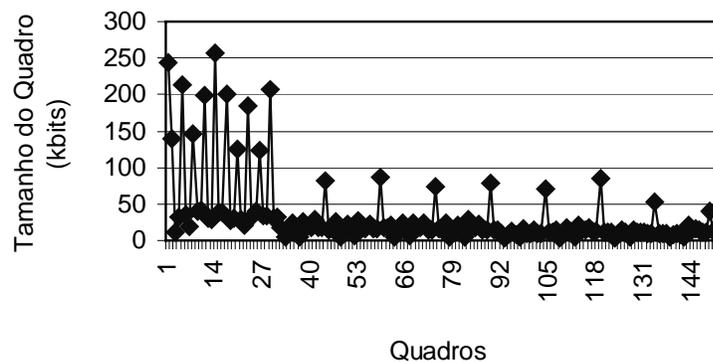
Fazendo algumas alterações no software ACQ, foi possível simular o comportamento do ajuste dos coeficientes de quantização para esta situação de renegociação de banda. Isto foi feito inserindo como parâmetro de entrada uma ou mais triplas de parâmetros que indicassem a taxa máxima, a taxa média e o momento de alteração para essas taxas. O momento de

alteração, por simplicidade, foi implementado como o número do quadro em que a alteração deveria ocorrer. Assim, antes do software recodificar o quadro indicado, ele roda novamente a rede neural, de forma a ajustar a matriz para as novas taxas máxima e média.

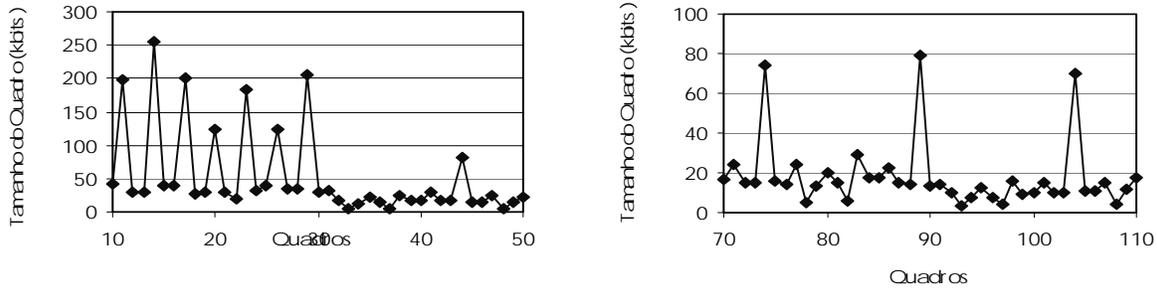


**Figura 8:** Cenas dos vídeos obtidos com recodificação para taxa de bit de 2 Mbps.

Uma simulação foi feita para o vídeo “Bike”, onde as taxas máxima e média iniciais eram ambas de 3,5 Mbps. Foi solicitada alteração da taxa à 1 segundo de início (30 quadros) do vídeo para as taxas máxima e média de 1,5 Mbps e 0,75 Mbps, e à 3 segundos do vídeo, para as taxas máxima e média de 0,75 Mbps e 0,42 Mbps. Pode-se observar na figura 9 alterações bruscas no perfil da curva do tamanho do quadro, nos instantes de alteração. A figura 10 mostra a ampliação de cada um desses instantes, mostrando o ajustes obtidos com os parâmetros a que o ACQ foi submetido.

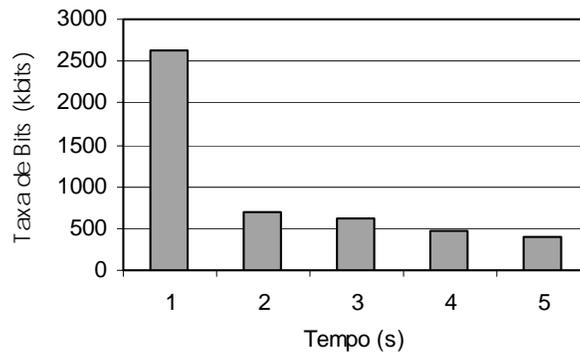


**Figura 9:** Tamanho dos quadros do vídeo “Bike” resultantes da simulação de uma alocação dinâmica de banda.



**Figura 10: Ampliação do gráfico da figura 9 para os instantes 1 e 3 segundos.**

A figura 11 mostra a variação da taxa de bit obtida com essa simulação, onde pode-se observar que a taxa média para o segundo e terceiro segundo foi de 650 kbps, e para o quarto e quinto segundo foi de 433 kbps.



**Figura 11: Variação da taxa de bit para a simulação de alocação dinâmica com o vídeo “bike”.**

Os dados obtidos com os testes do ACQ demonstraram ótimos resultados, permitindo uma grande variação de taxa de bits de vídeos já previamente codificados, possibilitando o ajuste a banda de transmissão disponível.

## 6. Considerações Finais

A técnica proposta neste trabalho, permite o ajuste da banda requerida para a transmissão de um vídeo digital previamente codificados, e conseqüentemente sua qualidade, quando a rede ATM, não apresenta no momento, recursos suficientes para estabelecer a conexão solicitada. Desta forma, é proposta uma aplicação auxiliar ao VoD, o Servidor de Qualidade (SoQ), que permite a adequação do sinal de vídeo, já previamente gravado, às condições da rede. No contexto das aplicações de “vídeo-em-casa”, principalmente para ensino a distância e distribuição de informação, essa adequação fornece flexibilidade no contrato da taxa de transmissão entre o provedor de conteúdo e o cliente. Portanto, o sistema permite, não só acesso ao vídeo em momentos em que a rede dispõe de uma quantidade menor de recursos, como também, acesso de menor custo, o que é extremamente interessante em um contexto globalizado.

Este trabalho trouxe como resultado várias contribuições. A proposta e especificação do Servidor de Qualidade, como aplicação auxiliar ao sistema de vídeo sob demanda, é uma das principais contribuições, que deu origem a várias outras.

A proposta de uma nova técnica para o ajuste dos coeficientes de quantização do padrão MPEG à partir de uma rede neural “backpropagation” é a contribuição mais significativa. Esta proposta foi concretizada com a implementação de um protótipo do módulo principal do SoQ, o módulo de “Ajuste dos Coeficientes de Quantização” (ACQ).

Os resultados obtidos com o software implementado, demonstraram a viabilidade da utilização da técnica proposta. Os testes foram aplicados, simulando a utilização do SoQ para negociação de banda inferior a da taxa de bit do vídeo original, mostrando bons resultados. Fez-se também, a simulação de uma situação de alocação dinâmica, com renegociação durante a transmissão do vídeo, o que impõe que o vídeo original seja recodificado, parte com os parâmetros negociados pelo descritor de tráfego da negociação, e parte com os novos parâmetros relativos a renegociação. Estes testes também mostraram ótimos resultados, com a obtenção de taxas de bits iguais ou inferiores à taxa de bit média esperada.

Ao elaborar a técnica de Ajuste dos Coeficientes de Quantização do Padrão MPEG utilizando rede neural, a expectativa inicial era de obter uma redução de até 40 % do vídeo original, com qualidade de imagem ainda tolerável. O ACQ mostrou que é possível obter reduções ainda maiores, como foi o caso do exemplo do vídeo “Flowers”, onde de uma taxa de 14 Mbps, foi possível obter um vídeo de 2 Mbps, com qualidade aceitável. Isso representa uma redução de 80 % do tamanho original do vídeo.

É importante ressaltar que, apesar de ter sido aplicado somente para MPEG-1, esta técnica pode ser utilizada para o ajuste dos coeficientes de qualquer padrão de compressão que utiliza quantização, como por exemplo MPEG-2 e H.261.

O término deste trabalho associado aos bons resultados obtidos, abre possibilidade para que o SoQ seja integrado a um grande leque de aplicações de vídeo.

## **7. Bibliografia**

- [1] ATM Forum, Technical Committee, Audiovisual Multimedia Services: Video on Demand, Specification 1.0, Dec. 1995.
- [2] Stallings, W. “High-speed Networks – TCP/IP and ATM Design Principles”, Prentice Hall, 1998.
- [3] Saw, Y-S.; Grant P.M. and Hannah J.M. “Rate-distortion analysis of nonlinear quantisers for MPEG video coder: sigmoidal and unimodal quantiser control functions”, IEE Proc. Vis. Image Signal Process., vol. 145, nº 4, August 1998.
- [4] Saw, Y-S.; Grant, P.M.; Hannah, J.M. and Mulgrew, B. “Video Rate Control Using a Radial Function Estimator for Constant Bit-Rate MPEG Coders”, Signal Processing: Image Comm., vol. 13, 1998.
- [5] Reininger, D.; Ramamurthy, G. and Raychaudhuri, D. “VBR MPEG Video Coding with Dynamic Bandwidth Renegotiation”, IEEE International Conference on Communications, pg. 1773-1777, 1995.
- [6] Zhang, Z-L.; Nelakuditi, S.; Aggarwal, R. and Tsang, R.P. “Efficient Selective Frame Discard Algorithms for Stored Video Delivery Across Resource Constrained Network”, <http://www-users.cs.umn.edu/~zhzhang/papers.html>.
- [7] Eleftheriadis, A. and Anastassiou, D. “Constrained and General Dynamic Rate Shaping of Compressed Digital Video”, Proceedings of 2º IEEE International Conference on Image Processing (ICIP’95), Arlington, Virginia, October 1995.

- [8] Reininger, D.; Raychaudhuri, D. and Ott, M. “A Dynamic Quality of Service Framework for Video in Broadband Network”, IEEE Network, vol. 12, n° 6, Nov./Dec. 1998.
- [9] Jack, K. “Video Demystified”, segund edition, HighText Publ., 1996.
- [10] Fluckiger, F. “Understanding Networked Multimedia – Application and Technology”, Printece Hall, 1995.
- [11] Joseph, K. and Reininger, D. “Source Traffic Smoothing and ATM Network Interfaces for VBR MPEG Video Encoders”, Proc. IEEE International Conference on Communications, Seattle, Washington, June 1995.
- [12] Grupo Comunicação Homem-Máquina e Cabral, E.F. “Redes Neurais Artificiais – Um curso teórico e prático para engenheiros e cientistas”, Editor: E.F. Cabral Jr., 1999.
- [13] Wang, L.C.; Rizvi, S.A. and Nasrabadi, N.M. “A Modular Neural Network Vector Predictor for Predictive Image Coding”, IEEE Trans. on Image Processing, vol. 7, n° 8, Aug. 1998.
- [14] Luo, L.; Lu, Y.; Zou, C. and He, Z. “Image Sequence Macroblock Classification using Neural Networks”, Signal Processing, n° 69, 1998.
- [15] Chang, P.R. and Hu, J.T. “Optimal Nonlinear Adaptive Prediction and Modeling of MPEG Video in ATM Networks using Pipelined Recurrent Neural Networks, IEEE Journal on Selected Areas in Communications, vol. 15, n° 6, Aug. 1997.
- [16] Cheng, R.G. and Chang, C.J. “Neural-network Connection-admission Control for ATM Networks”, IEE Proc. Commun., vol. 144, n° 2, April 1997.
- [17] Cheng, R.G.; Chang, C.J. and Lin, L.F. “A QoS-Provisioning Neural Fuzzy Connection Admission Controller for Multimedia High-Speed Networks”, IEEE/ACM Transactions on Networking, vol. 7, n° 1, Fav. 1999.
- [18] Youssef, S.A.; Habib, I.W. and Saadawi, T.N. “A Neural Network Control for Effective Admission Control in ATM Networks”, IEEE 1996, pg. 434-438.
- [19] Chester, M. “Neural Networks – Tutorial”, PTR Prentice Hall, 1993.
- [20] Jensen, C. “QwikNet32”, 1997, <http://www.kagi.com/cjensen>
- [21] Lin, F-H. and Mersereau, R.M. “Rate-quality tradeoff MPEG video encoder”, Signal Processing: Image Communication vol. 14, Fev. 1999, pg. 297-309.
- [22] Silveira, R.M. e Ruggiero, W.V. “Sistema de Ajuste dos Coeficientes de Quantização MPEG em Tempo-Real para Vídeo sob Demanda”, submetido à SBRC2000.
- [23] Ohta, N. “Packet Video – Modeling and Signal Processing”, Artech House, Inc; 1994.
- [24] R.M. Silveira and W.V. Ruggiero; “Servidor de Qualidade para Sistema de Vídeo sob Demanda”. SBMIDIA2000 – Simposio Brasileiro de Hipermidia e Multimidia, Junho/2000, Natal, Rio Grande do Norte, Brasil.
- [25] R.M. Silveira and W.V. Ruggiero, “Sistema de Ajuste dos Coeficientes de Quantização MPEG em Tempo-Real para Vídeo sob Demanda”, apresentado no XVIII Simpósio Brasileiro de Redes de Computadores, Maio/2000, Belo Horizonte, MG, Brasil.