

# **TVS - Um Sistema de Videoconferência com Documentos Compartilhados sobre a Arquitetura HyperProp\***

*Jauvane Cavalcante de Oliveira   Luiz Fernando Gomes Soares*

Laboratório TeleMídia - Departamento de Informática  
Pontifícia Universidade Católica do Rio de Janeiro  
Janeiro-1996, Rio de Janeiro, RJ

E-mail: jauvane@inf.puc-rio.br   lfgs@inf.puc-rio.br

## **Resumo**

Este trabalho apresenta o TVS (TeleMídia Videoconferencing System), um sistema de videoconferência com grande aderência aos padrões ISO e ITU-T para este tipo de aplicação. O TVS é um sistema que, além da transmissão básica das mídias de áudio e vídeo de modo síncrono, permite a manipulação de documentos multimídia/hipermídia através da máquina HyperProp. Adicionalmente o sistema oferece suporte a votação e envio de mensagens entre os usuários.

**Palavras-chave:** Videoconferência, Serviços Audiovisuais, Teleconferência, Manipulação de Documentos, HyperProp, Trabalho Cooperativo, H.261.

## **Abstract**

This paper presents TVS (TeleMídia Videoconferencing System), a videoconferencing system conforms to ITU-T and ISO standards. TVS is a system that, besides the usual transmission of audio and video, allows the multimedia/hypermedia documents handling, throught the HyperProp model. The system allows vote and text message sending facilities.

**keywords:** Videoconferencing, Audiovisual Services, Teleconferencing, Document Handling, HyperProp, Cooperative Work, H.261.

---

\* Este trabalho foi desenvolvido com o apoio do CNPQ através do projeto ProTeM II - HyperProp.

## 1. Introdução

A reunião de grupos de pessoas, em geral, acarreta um desperdício de tempo e dinheiro, principalmente devido à necessidade de locomoção dos participantes e ao pobre apoio de infra-estrutura. Na tentativa de melhorar este quadro, vários grupos de pesquisa iniciaram estudos com o objetivo de desenvolver um novo conjunto de serviços de comunicação denominados serviços de teleconferência.

Serviços de teleconferência são definidos [SoMB 88, Fluc 95] como um conjunto de facilidades de comunicação através de meios eletrônicos para possibilitar comunicações bidirecionais entre dois ou mais usuários ou grupos de usuários. A literatura em geral classifica os serviços de teleconferência em [H.200, SoMB 88]:

- *Áudio Conferência* - Aqueles em que somente sinais de áudio e controle são transmitidos entre os participantes
- *Conferência Áudio-Documentacional* - Similar à áudio conferência, havendo entretanto o tratamento de documentos textuais
- *Conferência Audiográfica* - Serviço com suporte a transmissão de áudio, sinais de controle, documentos e imagens estáticas
- *Freeze-Frame Videoconferência* - Serviço similar à conferência audiográfica acrescida do envio periódico de imagens estáticas dos participantes
- *Teleseminário* - Serviço que consiste da distribuição dos eventos ocorridos num local (áudio e vídeo) para todos os demais participantes, sendo o áudio o único sinal de retorno
- *Videoconferência* - Serviço similar à conferência audiográfica acrescida do envio on-line e em tempo real de sinais de vídeo entre os vários participantes

Além dos serviços de teleconferência, certos serviços audiovisuais [H.200], de características bastante próximas, são encontrados, Normalmente compartilhando os mesmos padrões e soluções desenvolvidos para teleconferência. Entre eles destacam-se:

- *Telefonia Convencional* - Incluída com o intuito de levantar considerações sobre interoperabilidade com os demais serviços
- *Videofonia* - Serviço com transmissão ponto-a-ponto bidirecional de sinais de áudio e vídeo
- *Telesurveillance* - Serviço de transmissão unidirecional de sinais com o intuito de efetuar observação de ambientes. Ainda não existe proposta de padrão para este tipo de serviço [H.200].

O serviço de videoconferência ocupa um lugar de destaque neste conjunto, uma vez que todos os outros serviços podem ser considerados casos particulares de videoconferência [F.730], o que facilita a interoperabilidade entre eles, como apresentado posteriormente.

Como resultado das pesquisas desenvolvidas nos últimos anos, vários protótipos de videoconferência foram apresentados e se encontram operacionais. Tais protótipos apresentam, no entanto, algumas limitações. Muitos não utilizam os padrões

acordados, além de se limitarem apenas às transmissões sem sincronismo das mídias áudio e vídeo e, mesmo assim, com qualidade duvidosa. Quanto à segurança em tais sistemas, alguns aspectos começaram a ser alvos de estudos apenas recentemente.

Para ser interoperável, o sistema deve seguir os padrões estabelecidos para transmissão das diversas mídias. Os organismos de padronização internacional (ISO e ITU-T) indicam a utilização do H.261 [H.261] para transmissão de vídeo on-line e em tempo real, MPEG, MPEG-2 ou MPEG-4 [MPEG, MPEG2, MPEG4] para armazenamento e recuperação de vídeo, JPEG [JPEG] para codificação de imagens estáticas, G.711 [G.711] obrigatoriamente para a codificação de áudio e opcionalmente G.722, G.728, G.723 [G.722, G.728, G.723].

A obediência a padrões de codificação das várias mídias é condição necessária, mas não suficiente, se se deseja um sistema de videoconferência aberto. Padrões para codificação de informações multimídia/hipermídia em sua forma final, incluindo aí o sincronismo espacial e temporal das diversas mídias e a descrição da apresentação, devem ser igualmente seguidas [MHEG]. Da mesma forma, o empacotamento dos dados multimídia deve seguir padrões adequados [H.221, H.223], bem como os mecanismos de segurança [Schn 95]. Recentemente uma proposta para a transmissão dos dados de controle foi também apresentada pela ITU-T [Draft H.245].

O uso de padrões adequados é muito importante, pois além de permitir a portabilidade do sistema, permite a interoperabilidade com sistemas afins, por exemplo videofonia [F.720, F.721], conferência audiográfica [F.710, F.711, T.120], outros sistemas de videoconferência, e até a própria telefonia convencional. Adotados e seguidos os padrões, a compatibilização pode ser realizada através da filtragem das mídias de modo apropriado. Sendo o áudio [G.711] a mídia básica, para fins de compatibilização, estando presente em todos os serviços audiovisuais [H.200, F.710, F.711, F.720, F.721, F.730].

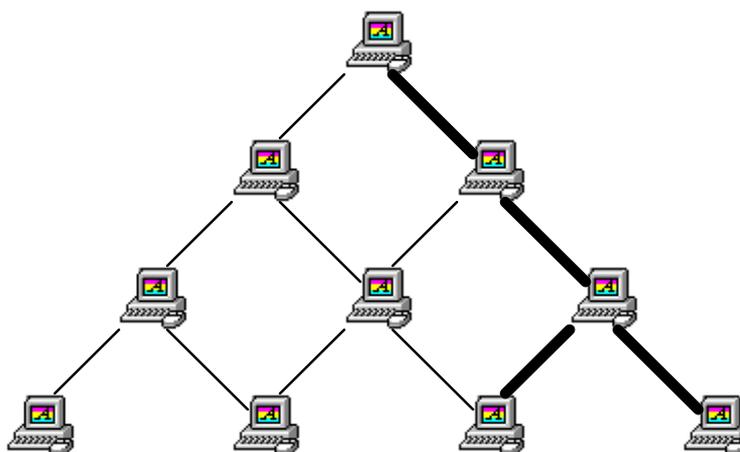


Figura 1.1: Minimização do Congestionamento no Sistema de Comunicação

Um grande problema na implementação de um sistema de videoconferência é a sobrecarga gerada no sistema de comunicação pela transmissão das várias mídias. O sistema deve prover mecanismos de diminuição no fluxo das mídias, em especial o áudio e vídeo, que são as que mais utilizam recursos de banda passante da rede. Tais mecanismos devem evitar, por exemplo, que um vídeo seja transmitido mais de uma vez em uma sub-rede (Figura 1.1).

Um sistema de videoconferência deve oferecer outras facilidades adicionais à simples transmissão síncrona de áudio e vídeo. Tais facilidades incluem uma etapa anterior à conferência (denominada pré-conferência) para agendamento e configuração do ambiente, facilidades para manipulação de documentos e trabalho cooperativo, suporte à votação, facilidades para troca de mensagens entre os usuários, possibilidade de gravação da conferência para posterior assistência e, em todas as suas funções, mecanismos de segurança, por exemplo, esquemas de autenticação que impossibilitem a falsa identidade de um usuário, entre outros.

Em qualquer sistema de teleconferência o controle de acesso ao ambiente (comumente denominado “floor control”), que gerencia quais participantes têm direito à fala, à manipulação de documentos, etc. em um dado instante, bem como o controle do período máximo de tempo que cada participante tem quando detém um controle específico do ambiente, é um mecanismo chave e, em geral, de difícil implementação. Um sistema ideal deve permitir a configuração completa do acesso às suas facilidades, e se encarregar de gerenciar as regras estabelecidas.

Este trabalho descreve uma implementação de um sistema de videoconferência, o TVS (TeleMídia Videoconferencing System), resultado dos estudos realizados no Laboratório TeleMídia do Grupo de Redes e Sistemas Multimídia do Departamento de Informática da PUC-Rio. Tal sistema busca cumprir os requisitos anteriormente mencionados, procurando estar em conformidade com as recomendações ITU e padrões ISO pertinentes.

O TVS é um sistema que possibilita, além da transmissão das mídias áudio [G.711, G.722] e vídeo [H.261, H.263] de forma síncrona e padronizada, a manipulação de documentos multimídia/hipermídia, baseada no Modelo de Contextos Aninhados (MCA) [MCA], em conformidade com a proposta de padrão MHEG [MHEG 95]. O sistema apresenta suporte à votação e envio de mensagens entre participantes, além de permitir uma ampla configuração do ambiente. Todo o controle de acesso ao ambiente é realizado por detecção de silêncio [Fari 92], por razões de eficiência com relação à interatividade em uma reunião.

Este trabalho está organizada como se segue. No capítulo 2 são apresentados alguns protótipos hoje disponíveis, incluindo o TVS, através de um estudo de suas características, pontos positivos e omissões. No capítulo 3 apresenta-se a arquitetura do sistema TVS. O capítulo 4 discorre sobre características da implementação do sistema em ambiente distribuído. Finalmente o capítulo 5 apresenta conclusões além de trabalhos futuros e tendências.

## 2. Trabalhos Relacionados

A terminologia utilizada no restante deste trabalho, uma adaptação daquelas apresentadas por Szypersky [SzVe 93] e Soares [SoMB 88, SoCR 95], pode ser resumida como se segue:

- *Participante*: usuário da conferência com direitos, controlados pelo coordenador, à fala e demais facilidades da conferência.
- *Organizador*: indivíduo que tem como tarefa agendar a conferência e, se necessário, divulgar aos participantes a existência da conferência. Pode ser um participante ou não. No âmbito do TVS o agendamento é realizado na etapa de pré-conferência, comentada posteriormente.
- *Coordenador*: participante com direitos especiais sobre todo o controle da conferência. Uma conferência pode ser realizada sem a presença deste indivíduo, quando o controle de acessos seria realizado pelo próprio sistema.
- *Interlocutor*: participante que detém, em um dado instante, o direito à fala e alteração dos documentos multimídia/hipermídia. O direito de alteração de documentos pode ser delegado a um secretário.
- *Secretário*: usuário da conferência para quem se delega o direito de escrita nos documentos multimídia/hipermídia da base compartilhada. Pode ser um participante ou não.
- *Assento*: dispositivo lógico que pode ser preenchido por um participante ou secretário.
- *Base Privada*: sessão de trabalho de um usuário, de acesso (controle) restrito a este usuário.
- *Hiperbase* ou *Hiperbase Pública*: depósito de documentos persistentes de acesso a todos os usuários da conferência, de acordo com seus direitos.
- *Base Compartilhada*: depósito volátil de documentos que possibilita o trabalho cooperativo entre os participantes da conferência. É visível por todos os participantes, mas com controle de alteração realizado pelo sistema. Através desta base as operações de trabalho cooperativo sobre documentos multimídia/hipermídia é suportada pelo TVS.

Vários são os grupos que estão estudando videoconferência, muitos protótipos já se encontram operacionais. Este capítulo apresenta resumidamente vários destes protótipos enumerando características desejáveis e omissões. Os sistemas considerados são o CU-SeeMe [Park 95] da Cornell University, Figura 2.1; o NV [Fred 94] da Xerox PARC; o VIC [McJa 95] do Lawrence Berkeley Laboratory e University of California, Berkeley; o IVS [Turl 93, Turl 95] do Inria, Figura 2.2; e o TVS do Laboratório TeleMídia, PUC-Rio, Figura 3.3.

### 2.1. Codificação das mídias básicas: áudio e vídeo:

- O Cu-SeeMe utiliza codificações proprietárias. O vídeo transmitido pode ter a resolução de 320×240 ou 160×120 com quatro bits por amostra. Os pixels são codificados em 16 tons de cinza e agrupados em blocos de 8×8 pixels. A compressão implementa detecção de movimento e proporciona

uma redução da ordem de 40%. O áudio é igualmente codificado por algoritmo proprietário.

- O nv utiliza a codificação MJPEG (Motion-JPEG), ou CellB da SUN Microsystems, ou ainda uma codificação proprietária, conforme configuração. O vídeo pode ser transmitido em cores ou preto e branco. O quadro é dividido em blocos de 8x8 pixels. Posteriormente é realizada detecção de movimento e aplicada a transformada de Haar [Turl 95] ou DCT (Discrete Cosine Transform) [STCN 92], dependendo do estado do sistema de comunicação. O nv não codifica sinais de áudio.
- O IVS é um sistema que funciona de acordo com as recomendações ITU-T, codificando vídeo H.261 [H.261], nos formatos padrão CIF e QCIF, e um novo formato, denominado SCIF [Turl 93], introduzido na recomendação H.263 com o nome 4CIF [H.263]. A codificação do áudio pode ser PCM [G.711], ADPCM ou VADPCM.
- O vic é um sistema que tem como meta a flexibilidade [McJa 95]. Neste sentido, ele permite codificação e decodificação de sinais de vídeo MJPEG, CellB, H.261 e nv, garantindo interoperabilidade com os sistemas IVS e nv. O tratamento de áudio é repassado a um outro aplicativo de teleconferência do mesmo grupo, o vat, que é um sistema de áudio conferência.
- O TVS é um sistema que funciona de acordo com as recomendações ITU-T, codificando vídeo H.261 [H.261] e áudio G.711 [G.711], G.728 [G.728].



Figura 2.1: Ambiente Cornell — CU-SeeMe

## 2.2. Empacotamento de dados:

- O Cu-SeeMe utiliza formato proprietário no empacotamento das mídias.
- O nv utiliza o protocolo RTP [RTP 94] do IETF (Internet Engineering Task Force).
- O vic utiliza o protocolo RTP [RTP 94] do IETF (Internet Engineering Task Force).

- O IVS utiliza um formato de quadro não padrão para envio de pacotes distintos de áudio e vídeo, o tamanho do pacote de áudio depende da codificação utilizada (PCM, ADPCM, VADPCM).
- O TVS utiliza o formato padrão ITU-T H.221 [H.221], quando da escolha do vídeo H.261.

### 2.3. Transferência segura de informações:

- O vic permite transmissão segura fim-a-fim através de encriptação com o algoritmo DES [Schn 95]. A chave pública para a decriptação é distribuída externamente.
- O TVS permitirá (próxima versão em implementação) transmissão segura das mídias selecionadas pelos participantes ou configuradas pelo organizador.
- O IVS, nv e CU-SeeMe não fornecem dispositivos de segurança.

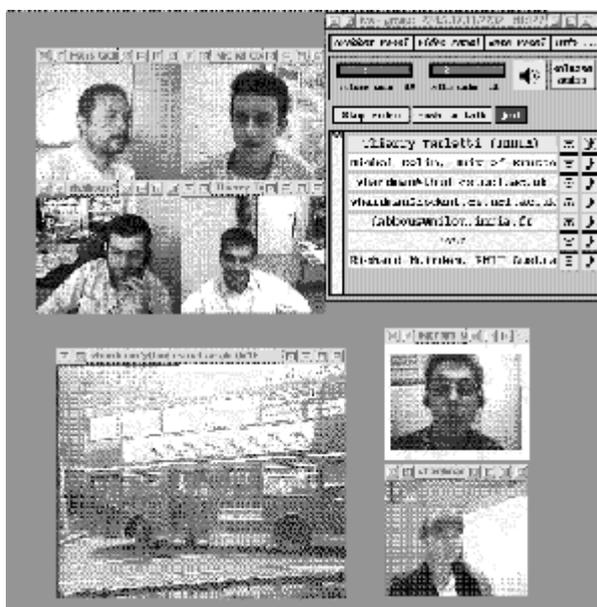


Figura 2.2: O Ambiente Inria — IVS

### 2.4. Sincronismo entre as mídias áudio e vídeo:

- Não há dados sobre dispositivos para sincronismo de mídias no CU-SeeMe.
- Não se aplica ao nv, uma vez que apenas a mídia vídeo é considerada.
- O IVS envia a mídia vídeo por uma porta UDP [Come ??] e a mídia áudio por outra. Em nenhum momento, a partir do envio, é reconsiderado o sincronismo entre as mídias que pode chegar ao destino dessincronizadas.
- O vic apenas transmite vídeo, sendo tarefa de outro aplicativo, o vat, a transmissão de áudio. Como funciona hoje, não há relação nenhuma entre as duas transmissões, o que pode acarretar no não sincronismo entre as mídias no destino. A documentação [McJa 95] promete a implementação de outro módulo, o ct, que coordenaria as transmissões do vic e vat de modo a garantir o sincronismo entre as mídias.

- O TVS empacota as mídias áudio e vídeo nos quadros H.221, que permitem a exibição sincronizada no destino, quando adequadamente tratados.

#### 2.5. Tratamento de Documentos:

- O vic não fornece tratamento. Do mesmo grupo, entretanto, existe um aplicativo de whiteboard, o wb. Este aplicativo funciona de modo independente do vic e permite o uso de uma área gráfica compartilhada entre os participantes.
- O TVS fornece a facilidade de manipulação compartilhada de documentos através do Modelo de Contextos Aninhados em conformidade com a proposta de padrão MHEG.
- O nv, IVS e CU-SeeMe não possuem tratamento de documentos.

#### 2.7. Suporte a Votações:

- Apenas encontrada no TVS.

#### 2.8. Envio de Mensagens Texto:

- O CU-SeeMe permite que mensagens sejam escritas sobre a janela de vídeo dos participantes.
- O TVS permite envio de mensagens para qualquer participante, ou grupo de participantes.
- O vic, nv e IVS não fornecem esta facilidade.

#### 2.9. Controle de Acesso:

- O vic não executa as funções de Controle de Acesso, porém existe a promessa de que o ct teria também esta função.
- O TVS controla o acesso a todos os recursos da conferência, podendo o organizador determinar as permissões de acesso de todos os participantes na etapa de pré-conferência e o coordenador alterar os direitos de acesso durante a conferência. O Controle de Acesso é implementado pelo próprio sistema através da técnica de detecção de silêncio [Fari 92].

### 3. Interface com o Usuário

O TVS prevê uma etapa anterior à conferência propriamente dita, denominada pré-conferência. Nessa etapa, o organizador agenda e configura o ambiente da conferência. Nesse agendamento, se disponibiliza o assunto principal a ser tratado, a data da conferência, bem como o número de assentos disponíveis. Cada assento pode ser preenchido de uma entre três formas: i) *assento cativo*, quando se indica exatamente qual usuário pode ocupá-lo (fulano@inf.puc-rio.br), ii) *assento para um domínio*, quando se deixa o assento disponível para uma comunidade de um determinado domínio (\*@inf.puc-rio.br) ou iii) *assento geral*, quando se deixa o assento disponível para qualquer participante (\*@\*). O organizador deve ainda indicar os direitos de

acesso de cada participante, podendo eleger um deles para ser o coordenador da conferência.

Ainda na fase de pré-conferência, vários outros aspectos, como tempo máximo dado a cada interlocutor e o formato do ambiente da conferência, devem ser determinados.

O ambiente TVS é completamente configurável, podendo o organizador otimizar a interface para o tipo de reunião desejada, excluindo janelas desnecessárias, reorganizando os vários diálogos, etc. O participante pode alterar o seu ambiente, se desejar, durante a sessão.

O TVS implementa as operações de pré-conferência através de interações do MIU – Módulo de Interação com o Usuário – com um *daemon* de controle de conexão. O *daemon* é responsável pela manutenção de um cadastro de conferências agendadas, dos assentos e seus tipos de ocupação, das votações ativas e apuração de votos de cada conferência, como será visto posteriormente. É o *daemon* quem controla o acesso dos usuários às conferências, sendo ainda responsável por realizar, através de informações obtidas dos vários usuários, o controle de acesso (floor control) de modo centralizado [SoMB 88].



Figura 3.1: Menu de Conferências

A primeira tarefa do MIU é requisitar ao *daemon* a lista de conferências agendadas. A lista é então apresentada ao usuário, que pode escolher a conferência que deseja participar (Figura 3.1). Selecionada a conferência, o MIU solicita ao *daemon* uma autorização de entrada para o usuário. Neste ponto, o acesso é negado ou o ambiente da conferência selecionada é apresentado ao usuário.

O ambiente da conferência é composto de duas janelas obrigatórias e dez janelas configuráveis. As obrigatórias são: Janela Principal e Console. As janelas configuráveis são subdivididas em quatro grupos: Bases de Informação, Vídeos, Controle e Votação.

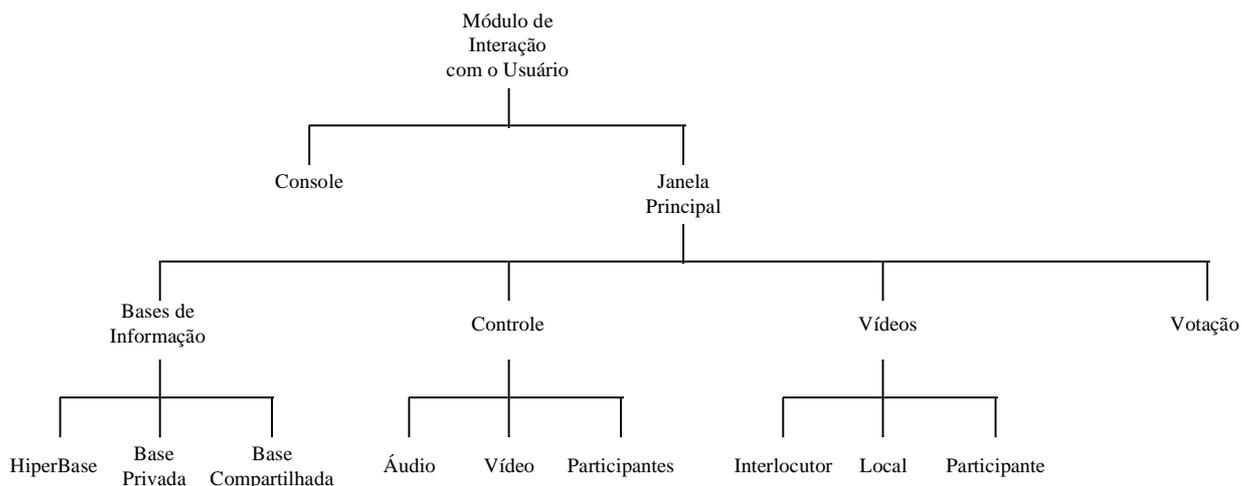


Figura 3.2: Estrutura do Módulo de Interação com o Usuário TVS

Nas janelas do grupo Bases de Informação encontra-se a janela de Hiperbase, Base Privada e Base Compartilhada. Nas janelas do grupo Vídeo encontram-se a janela do vídeo do interlocutor, a do vídeo local e a do vídeo de participante. No grupo janelas de controle encontra-se a janela de controle de transmissão de vídeo, de áudio e a lista de participantes. A janela de Votação completa o ambiente da conferência (Figuras 3.2 e 3.3).

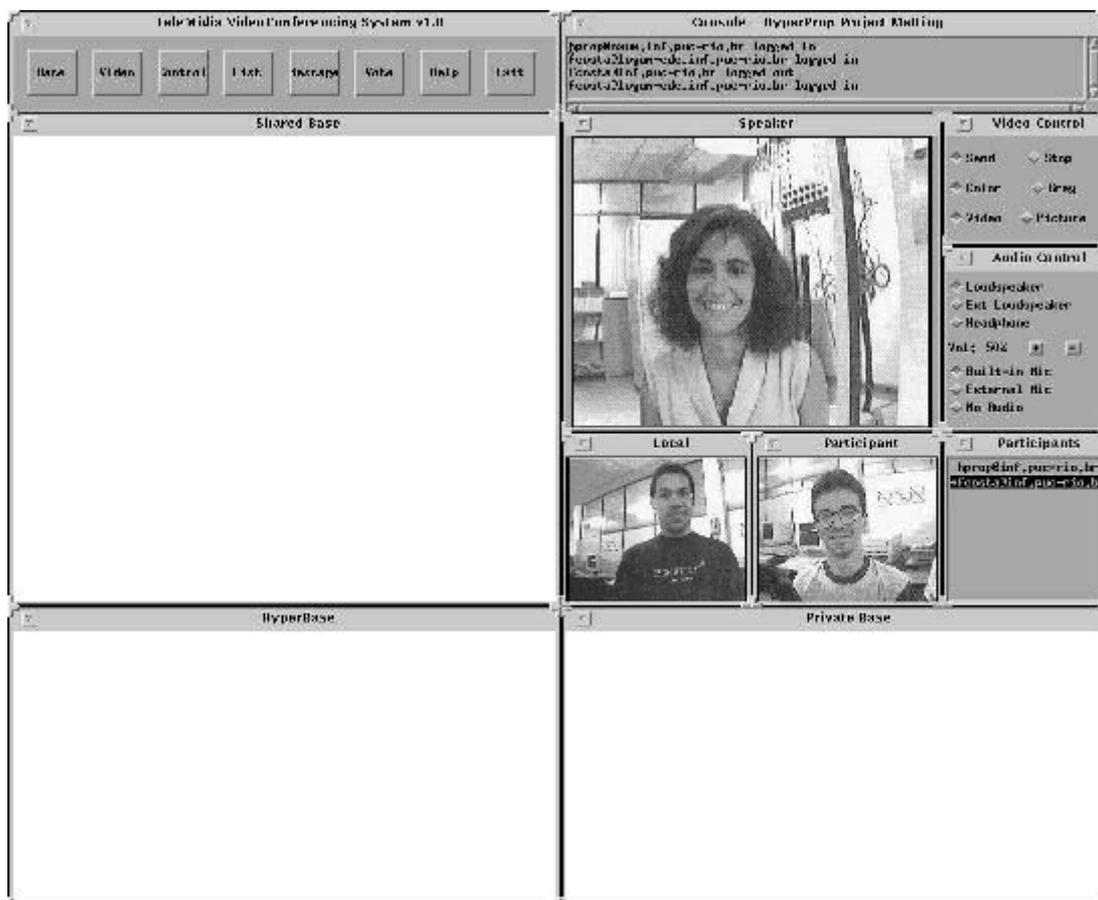


Figura 3.3: O Ambiente TVS

A Janela Principal consiste de um menu, com um ícone para cada grupo apresentado. Através de seleções, o usuário pode “ligar” ou “desligar” cada uma das demais janelas do sistema, excetuando, como mencionado, a própria janela principal e a console.

A Console é uma janela passiva, onde o usuário apenas recebe mensagens do sistema. É nesta janela também que o sistema apresenta ao participante as mensagens a ele endereçadas por outros participantes.

O vídeo do interlocutor é apresentado no formato CIF (352×288 pixels), por ser a imagem que deve ter maior atenção do usuário. Os demais vídeos utilizam o formato QCIF (176×144 pixels) de menor resolução. A utilidade da janela de vídeo de participante é questionável, embora endossada pela recomendação F.730 da ITU-T, uma vez que a facilidade da exibição do vídeo de um participante qualquer da conferência aumenta em muito o tráfego gerado na rede. Em geral o único vídeo difundido é o do interlocutor.

Na janela de controle de vídeo, o usuário faz seleções sobre o tipo de vídeo que ele próprio pretende enviar aos demais participantes da conferência. Esta janela apresenta botões para seleção de envio ou não de vídeo, envio de vídeo em cores ou em preto e branco e, finalmente, envio de vídeo ou de uma imagem estática (foto). Com estas opções o usuário pode desligar a transmissão do seu próprio vídeo se este for o seu desejo, pode enviar uma foto JPEG [Wall 91] se ele não dispuser de câmera, etc. Se a exibição do vídeo de um participante que tenha desligado sua transmissão for selecionada por outro participante, este último receberá na janela correspondente uma mensagem “Not Sending”.

Na janela de controle de áudio o usuário configura o envio e recepção de sinais de áudio [G.711] [G.728]. As seleções apresentadas são: saída de áudio desejada (Loudspeaker, External Loudspeaker ou Headphone), volume de saída de áudio e dispositivo de entrada de áudio desejada (Built-In Microphone, External Microphone ou No Audio). Os usuários que não enviam sinais de áudio não são considerados na disputa do controle de acesso, sendo portanto simples expectadores.

Na janela de participantes o usuário recebe uma série de informações sobre os demais participantes da conferência. Nessa janela aparecem os nomes dos vários participantes, com indicação gráfica dos usuários que o estão observando. Ainda nesta janela o usuário pode selecionar qual, dentre os participantes, se deseja o vídeo exibido na janela de vídeo de um participante. É também nesta janela que o usuário pode requisitar o envio de uma mensagem para um participante selecionado

A janela base privada consiste de um browser [MuSC 95] da Base Privada do usuário, dentro do conceito do Modelo de Contextos Aninhados [Hype 95]. Resumidamente, esta janela consiste de uma área privada de trabalho do usuário, onde é realizada a manipulação dos documentos multimídia/hipermídia, totalmente transparente a qualquer outro usuário.

A janela do repositório de documentos (HyperBase) consiste de um browser da Hyperbase Pública no conceito do MCA, que é o repositório dos documentos

multimídia/hipermídia que podem ser compartilhados por todos os participantes da conferência. Qualquer documento pode ser “levado” deste repositório (Hiperbase Pública) para a base privada do usuário (na realidade uma versão do documento é criada) e exibido em sua janela privada. Documentos também podem ser manipulados cooperativamente como descrito a seguir.

A janela Compartilhada (Shared Base) consiste de uma abstração criada para a manipulação cooperativa de documentos. Ela consiste também de uma Base Privada, no conceito do MCA, onde são realizadas manipulações de documentos multimídia/hipermídia pelos usuários, sobre o controle do sistema. Esta janela é exibida igualmente em todas as estações de usuários. Como na janela privada, versões de documentos da hiperbase pública podem ser manipulados nesta janela. Os documentos presentes na base compartilhada podem ser alterados, cooperativamente, por qualquer participante da conferência (desde que lhe tenha sido dado este direito). O acesso ordenado ao documento é realizado por uma regra simples. O interlocutor é quem detém o direito de alteração nos documentos da base compartilhada, podendo, opcionalmente, delegar este direito a um outro participante. Esta delegação de direito pode ser utilizada para criar um personagem similar a um secretário que, enquanto a discussão sobre um determinado tópico vai-se desenrolando, se encarrega de realizar as alterações solicitadas. Note que o interlocutor detém o controle da janela compartilhada por um espaço de tempo definido pelo mecanismo de controle de acesso. Trocando o interlocutor, troca também o participante com o direito à manipulação dos documentos compartilhados nesta janela. Conceitualmente, os documentos da janela compartilhada fazem parte de uma base privada cujo “dono” é o próprio sistema. Assim, as alterações realizadas são na realidade requisições de alteração que os usuários fazem ao sistema e este (dono da base) as realiza.

A base compartilhada pode ainda receber documentos provenientes da base privada do usuário, desde que tal documento esteja no estado “committed” do MCA [SoCR 95], uma vez que é gerada uma versão do documento na base compartilhada. Esta operação define um novo método para o MCA — a transferência de documentos entre bases privadas.

Na base compartilhada, apresentam-se conteúdos dos nós provenientes das bases privada e hiperbase pública. O interlocutor ou secretário podem manipular estes documentos durante a fatia de tempo alocada. Essa manipulação permite edições completas dos documentos — lembre que o documento hora apresentado é uma versão de um documento qualquer.

Finalmente, a janela de votação apresenta ao usuário as votações que estão ativas. O usuário pode criar uma nova votação, apagar uma votação existente (caso ele próprio a tenha criado) ou seleccionar uma das votações e efetivamente votar.

## 4. Arquitetura do Sistema

O TVS apresenta uma estrutura distribuída, apresentada na figura 4.1, utilizando a pilha de protocolos TCP/IP, mais precisamente o protocolo UDP, conforme mencionado anteriormente. Acima dessa camada, o TVS apresenta uma camada de comunicação, existindo uma subdivisão para comunicação de mensagens de controle e mensagens de transporte das mídias propriamente ditas. Acima desta camada, apresenta-se o MIU (Figura 4.1), que, através de interações com o usuário, possibilita a sua participação no ambiente da videoconferência.

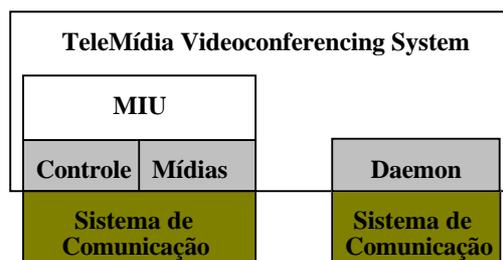


Figura 4.1: Camadas TVS

Ainda faz parte da arquitetura distribuída TVS, o daemon de controle de conexão, TVSD. Esse daemon é um processo, continuamente em execução, responsável pelo agendamento e controle das sessões de videoconferência TVS, através de troca de mensagens de controle com o MIU, via classe de comunicação de controle.

Um esquema de funcionamento de uma sessão de videoconferência pode ser visto na figura 4.2, que apresenta as trocas de informação entre os vários componentes do ambiente TVS. Verifica-se que as trocas de mensagens de controle são efetuadas via daemon, com exceção do envio das mensagens textuais — bilhetes — que se utilizam de transmissões de controle e são endereçadas ao destinatário, sem intermediários. Vê-se também que as transmissões das mídias são realizadas de maneira direta, sem passar pelo daemon, por motivos de melhoria de performance e de controle de congestionamento no sistema de comunicação.

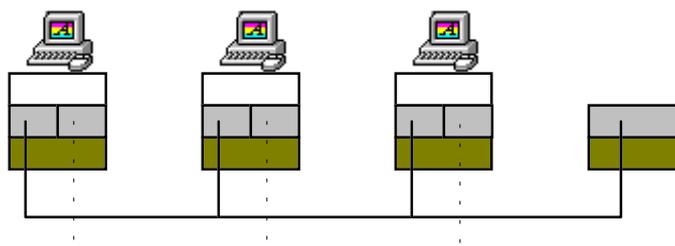


Figura 4.2: Sessão de Videoconferência — Interação do TVS com o Daemon (TVSD)

No que concerne à manipulação de documentos multimídia/hipermídia, o TVS efetua troca de mensagens com uma implementação de servidor da máquina HyperProp, em conformidade com o MCA, com o intuito de recuperar e exibir os documentos selecionados pelo usuário, permitindo o trabalho cooperativo nesses documentos. Assim, o TVS utiliza o conjunto de mensagens definidas pela máquina HyperProp para

apresentar na janela “HyperBase” e “Private Base” um browser [MuSC 95] dos diversos nós multimídia/hipermídia que podem ser acessados pelo participante. Posteriormente o sistema troca mensagens com a máquina para obter o conteúdo dos nós para exibição na base compartilhada, como descrito posteriormente.

## 5. Implementação

O TVS apresenta uma interface amigável através da utilização do sistema de interfaces IUP/LED [Levy 93] desenvolvido no Departamento de Informática da PUC-Rio. A versão utilizada é a do IUP-Motif. O código do sistema está escrito em C++.

A implementação atual do sistema cria uma abstração da interface apresentada ao usuário. Existe basicamente uma classe que dispara os métodos relacionados a cada elemento de interface (diálogo), uma classe que é responsável pela transmissão de sinais de controle e outra para transmissão das mídias (Figura 5.1). Todas as transferências de informações são realizadas através de envio de datagramas através de soquetes UDP.

O daemon possui um único soquete, associado à porta 5550, para envio e recepção de mensagens de controle. De acordo com a mensagem, um método adequado é disparado. Se o daemon passar um tempo maior que um valor convencional sem receber mensagens, ele envia mensagens “alive bit”, para verificar se as conexões estão ainda todas ativas. Caso alguma estação não responda à solicitação, o daemon envia uma mensagem de difusão avisando às demais estações, daquela conferência, da falha ocorrida. O daemon é um processo que usa poucos recursos de CPU, estando a maior parte do tempo bloqueado (select), aguardando a chegada de mensagens ou do timeout para envio da mensagem “alive bit”. Se nenhuma conferência está ativa, ele fica bloqueado até a chegada de alguma requisição.

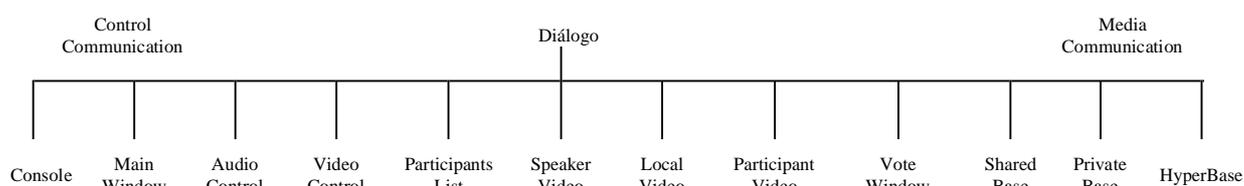


Figura 5.1: Hierarquia de Classes do Sistema.

Os dados sobre as conferências e votações são armazenadas em arquivos criptografados, lidos quando da ativação do daemon. Quando ativado, o daemon cria listas de conferências associando a cada entrada (conferência) um identificador, o assunto, o número de assentos disponíveis, uma estrutura homogênea para armazenar dados dos participantes, a data de realização da conferência, o número de votações ativas e uma lista com as informações das votações. De cada votação se armazena o título da votação, um identificador da votação, o identificador do participante que criou a votação, o número de opções e uma lista com as opções. De cada participante se armazena o nome, estação em que se encontra, domínio em que se encontra e dados de controle.

O TVS utiliza dois soquetes na sua camada de comunicação, um para envio de dados de controle e outro para envio das mídias, associados às portas 5551 e 5552 respectivamente.

A codificação de áudio e vídeo é feita, na versão corrente, por software. Os algoritmos utilizados são G.711 [G.711] e G.728 [G.728] para o áudio, H.261 para a codificação de vídeo e JPEG para transmissão de imagens estáticas. O empacotamento das mensagens é feito em conformidade com a recomendação H.221. O áudio é capturado diretamente do dispositivo `/dev/audio` do Solaris com o uso do dispositivo de controle `/dev/audiocntl`. O dispositivo de áudio do Solaris fornece as amostras PCM com a transformação logarítmica  $\mu$ -law (8 bits por amostra), através de recodificação do sinal G.711, obtém-se as amostras G.728, se este for o tipo de áudio escolhido pelo usuário. A captura de vídeo é feita através de funções da biblioteca Xil que acompanha a placa SUNVIDEO de captura de vídeo. Cada quadro é codificado de acordo com a recomendação ITU-T H.261, no formato CIF ou QCIF, conforme o participante seja o interlocutor ou não. Após a codificação, os sinais são acomodados em quadros H.221 e posteriormente enviados, pelo soquete reservado às mídias, para todos os participantes da conferência, caso o usuário seja o interlocutor, ou para os participantes com a marca indicativa de observação do usuário em questão.

Não é realizado qualquer controle de erro na transmissão das mídias de áudio e vídeo, uma vez que estas mídias toleram erros sem perda de qualidade perceptível ao ouvido e olho humano [STCN 92], respectivamente, se as taxas de erros nas redes em que transitam são baixas, hipótese assumida na implementação atual.

O mecanismo de votação é implementado de forma centralizada pelo daemon através de mensagens enviadas, pelo soquete de controle, ao daemon. O daemon valida o voto e contabiliza os resultados. Quando a maioria dos participantes cadastrados escolhe uma das opções, o daemon distribui o resultado aos participantes.

O TVS implementa a camada de comunicação através de duas classes: Transmissões de Controle e das Mídias. A classe de vídeo local é a responsável pela captura e codificação das mídias de áudio e vídeo, uma vez que esta classe é responsável pela apresentação da mídia vídeo na janela de vídeo local. Os métodos desta classe são ainda responsáveis pelo empacotamento das mídias em quadros H.221. Posteriormente os dados são enviados através da camada de comunicação de mídias para os participantes adequados. A transmissão das demais mídias — texto, imagem estática, etc. — é realizado quando da abertura de um nó da hiperbase pública ou da base privada de um participante. Posteriormente, apenas as atualizações são transmitidas, em tempo real.

O mecanismo de controle de acesso ao ambiente TVS é realizado através de detecção de silêncio, conforme estudo descrito em [Fari 92]. O mecanismo, esquematizado na figura 5.2, funciona através de um algoritmo bastante simples: Enquanto um interlocutor está expondo oralmente suas idéias, as demais estações ficam em modo de escuta. Cada estação executa um algoritmo de detecção de silêncio. Quando se detecta um intervalo de silêncio maior que um intervalo de tempo específico, as estações de usuários que demonstrarem interesse em tomar a palavra — basta que o participante comecem a falar — enviam uma mensagem de requisição de direito de acesso ao

daemon. O daemon, após receber todas as requisições, seleciona um dos participantes para ser o novo interlocutor. A decisão leva em conta uma lista de prioridades configurada pelo organizador e coordenador da conferência. A seguir, a decisão é enviada a todos os participantes da conferência, quando o novo interlocutor ganha o direito de acesso, iniciando sua locução, enquanto o secretário por ele indicado, ou ele próprio, ganha o direito de acesso à base compartilhada.

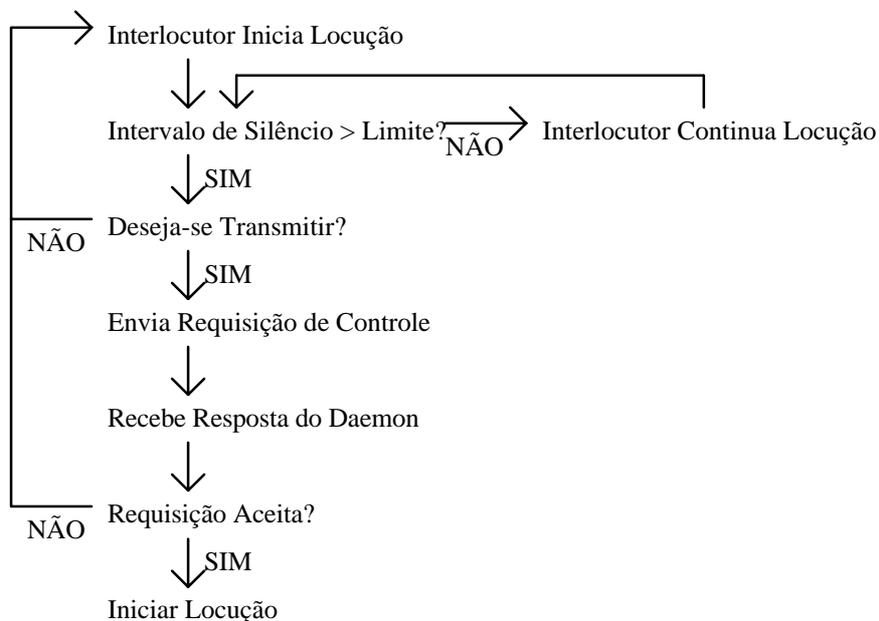


Figura 5.2: Estrutura Funcional do Mecanismo de Controle de Acesso

O interlocutor ou secretário por aquele indicado, tem o direito exclusivo de alteração nos documentos multimídia/hipermídia que estão na base compartilhada. Todas as operações de edição dos conteúdos dos nós são permitidas. Os demais participantes somente têm permissão de armazenar uma versão dos documentos da base compartilhada na sua base privada ou na hiperbase pública.

Em qualquer instante um usuário pode enviar mensagens textuais — bilhetes — a qualquer outro participante da conferência, desde que este esteja presente. A mensagem é enviada diretamente à estação destino através do canal de controle. No destino a mensagem é apresentada na console do MIU.

Todas as mensagens enviadas através da camada de comunicação de mídias utilizam a codificação H.221, já as mensagens enviadas pela camada de controle utilizam um formato de pacote genérico.

## 6. Conclusão

O TVS apresenta grande maleabilidade de configuração de interface, sendo adequado como suporte a vários tipos de reunião, desde teleseminários à reuniões de decisão. Algumas facilidades, em especial a manipulação cooperativa de documentos multimídia/hipermídia, não são encontradas nos protótipos e produtos comerciais hoje disponíveis, embora sejam previstas e solicitadas pelos padrões ITU-T [F.730].

Como trabalhos futuros, espera-se:

- Adaptar o sistema para uso em rede ATM;
- Incluir mecanismos de segurança, que devem levar em conta dois pontos de vista: transferência segura de informações confidenciais e autenticação de identidade de usuários. O primeiro problema pode ser resolvido através da utilização de um algoritmo eficiente de criptografia [Schn 95]. Já o segundo, é tema atual de pesquisas em vários centros de pesquisa e figura como problema ainda em aberto;
- Prover o cancelamento de eco. Nesta versão do TVS, o problema foi evitado pela existência de um único interlocutor em cada instante de tempo. A saída de áudio do interlocutor fica desativada. Este procedimento dificultaria a conversação quando o interlocutor estivesse utilizando fones de ouvido. Neste caso, o TVS não desabilita a saída de áudio na estação do interlocutor;
- Possibilitar a gravação da conferência para posterior assistência;
- Utilizar hardware específico para a codificação de vídeo, atualmente feita por software, tal mudança visa melhorar a performance do sistema.

## 7. Referências

[Come 93] Comer, D. - "Internetworking with TCP/IP - Vol I" - 2nd Edition - Prentice-Hall, 1993.

[Come XX] Comer, D. - "Internetworking with TCP/IP - Vol III BSD Socket Version" - 2nd Edition - Prentice-Hall.

[SzVe 93] Szyperski, C. Ventre, G. - "A Characterization of Multi-Party Interactive Multimedia Applications" - International Computer Science Institute TR-93-006, January 1993.

[Draft G.723] International Telecommunication Union, Telecommunication Standardization Sector - "Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 & 6.3 Kbit/s" - Draft ITU-T Recommendation G.723, September 1995.

[Draft H.263] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Video Coding for Low Bitrate Communications*" - Draft ITU-T Recommendation H.263, July 1995.

[Draft H.223] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Multiplexing Protocol for Low Bitrate Multimedia Communication*" - Final Draft ITU-T Recommendation H.223, to be published in next.

[Draft H.324] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Terminal for*

*Low Bitrate Multimedia Communication*” - Draft ITU-T Recommendation H.324, September 1995.

[Draft H.245] International Telecommunication Union, Telecommunication Standardization Sector - “Line Transmission of non-Telephone Signals - *Multimedia Control Protocol*” - Draft ITU-T Recommendation H.324, September 1995.

[Erik 94] Erikson, H. - “MBONE: The Multicast Backbone” - Communications of the ACM, August 1994, Vol 37, pp 54-60.

[F.701] Comité Consultatif International de Télégraphique et Téléphonie - “*Teleconference Service*” - CCITT Recommendation F.710, from CCITT Blue Book. ITU-T Recommendation F.701.

[F.710] International Telecommunication Union, Telecommunication Standardization Sector - “Telematic, Data Transmission and Teleconference Services: Operation and Quality of Service - *General Principles for Audiographic Conference Service*” - ITU-T Recommendation F.710, March 1991.

[F.711] International Telecommunication Union, Telecommunication Standardization Sector - “Operation and Quality of Service: Audiovisual Service - *Audiographic Conference Teleservice for ISDN*” - ITU-T Recommendation F.711, August 1993.

[F.720] International Telecommunication Union, Telecommunication Standardization Sector - “Telematic, Data Transmission, ISDN Broadband, UPT and Teleconference Services: Operations and Quality of Service - *Videotelephony Service - General*” - ITU-T Recommendation F.720, August 1992.

[F.721] International Telecommunication Union, Telecommunication Standardization Sector - “Telematic, Data Transmission, ISDN Broadband, Universal, Personal Telecommunications and Teleconference Services: Operations and Quality of Service - *Videotelephony Teleservice for ISDN*” - ITU-T Recommendation F.721, August 1992.

[F.730] International Telecommunication Union, Telecommunication Standardization Sector - “Telematic, Data Transmission, ISDN Broadband, Universal, Personal Communications and Teleconference Services: Operation and Quality of Service - *Videoconference Service - General*” - ITU-T Recommendation F.730, August 1992

[Fluc 95] Fluckiger, F. - “Understanding Networked Multimedia - Applications and Technology” - Prentice Hall, 1995

[Fred 94] Frederick, R. - “Experiences with real-time software video compression” - Xerox PARC, July 1994.

[G.711] Comité Consultatif International de Télégraphique et Téléphonie - “*Pulse Code Modulation (PCM) of Voice Frequencies*” - CCITT Recommendation G.711, from CCITT Blue Book.

[G.722] Comité Consultatif International de Télégraphique et Téléphonie - “*7kHz Audio-coding Within 64 kbits/s*” CCITT Recommendation G.722, from CCITT Blue Book.

[G.725] International Telecommunication Union, Telecommunication Standardization Sector - “General Aspects of Digital Transmission Systems; Terminal Equipment - *System Aspects for the Use of the 7 kHz Audio Codec Within 64 kbit/s*” - ITU-T Recommendation G.725, 1988.

[G.728] International Telecommunication Union, Telecommunication Standardization Sector - “General Aspects of Digital Transmission Systems; Terminal

Equipment - *Coding of Speech at 16 kbit/s Using Low-Delay Code Excited Linear Prediction* - ITU-T Recommendation G.728, September 1992.

[H.100] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Visual Telephone Systems*" - ITU-T Recommendation H.100, 1988.

[H.110] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Hypothetical Reference Connections for Videoconferencing Using Primary Digital Group Transmission*" - ITU-T Recommendation H.110, 1988.

[H.120] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Codecs for Videoconferencing Using Primary Digital Group Transmission*" - ITU-T Recommendation H.120, March 1993.

[H.130] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Frame Structure for Use in the International Interconnection of Digital Codecs for Videoconferencing or Visual Telephony*" - ITU-T Recommendation H.130, 1988.

[H.140] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Systems - *A Multipoint International Videoconferencing System*" - ITU-T Recommendation H.140, 1988.

[H.200] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Framework for Recommendations for Audiovisual Services*" - ITU-T Recommendation H.200, March 1993.

[H.221] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Frame Structure for a 64 to 1920 kbit/s Channel in Audiovisual Teleservices*" - ITU-T Recommendation H.221, March 1993.

[H.230] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Frame Synchronous Control and Indication Signals for Audiovisual Systems*" - ITU-T Recommendation H.230, March 1993.

[H.233] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Confidentiality System for Audiovisual Services*" - ITU-T Recommendation H.233, March 1993.

[H.242] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *System for Establishing Communication Between Audiovisual Terminals Using Digital Channels up to 2 Mbit/s*" - ITU-T Recommendation H.242, March 1993.

[H.261] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Video Codec for Audiovisual Services at  $p \times 64$  Kbit/s*" - ITU-T Recommendation H.261, March 1993.

[H.320] International Telecommunication Union, Telecommunication Standardization Sector - "Line Transmission of non-Telephone Signals - *Narrow-Band Visual Telephone Systems and Terminal Equipment*" - ITU-T Recommendation H.320, March 1993.

[Levy 93] Levy, C.H. - "IUP/LED: Uma Ferramenta Portátil de Interface com o Usuário" - Dissertação de Mestrado - Departamento de Informática / PUC-Rio, 1993.

[Liou 91] Liou, M. - "Overview of the p×64 kbit/s video coding standard" - Communications of the ACM, No. 4, April de 1991.

[McJa 95] McCane, S. Jacobson, V. - "vic: A Flexible Framework for Packet Video" - ACM Multimedia 95, San Francisco, CA, November 1995.

[MHEG 93] MHEG - Information Technology - "Coded Representation of Multimedia and Hypermedia Information Objects - Part 1: Base Notation", Committee Draft ISO/IEC CD 13522-1, 1993.

[MuSC 95] Muchaluat, D.; Soares, L.F.G; Casanova,M.A. - "Browsing in a Hypermedia System with Nested Composite Nodes" - Monografias em Ciência da Computação MCC23 - DI / PUC-Rio, 1995

[Park 95] Parker, T. - "Cornell Welcome Page - <http://cu-seeme.cornell.edu/>" - Cornell University, 1995.

[Schn 95] Schneier B. "Applied Cryptography" - 2nd Edition - John Willey, Novembro de 1995.

[RTP 94] Schulzrine, H.; Casner, S.; Jacobson, V.; Frederick, R. - "RTP: A Transport Protocol for Real-Time Applications" - Internet Draft ietf-avt-rtp-05, July 1994.

[SoMB 88] Soares, L.F.G.; Martins, S. de L.; Bastos, T.L.P. - "Lan Based Real Time Audio-Graphics Conferencing System, General Overview" - CCR066 Technical Report Rio Scientific Center-IBM Brasil, Novembro de 1988.

[SoCL 95] Soares, L.F.G.; Colcher, S.; Lemos, G. - "Redes de Computadores - Das LANS, WANS e MANS às Redes ATM" - Editora Campus, Janeiro de 1995.

[SoCR 95] Soares, L.F.G.; Casanova, M.A.; Rodriguez, N.L.R. - "Nested Composite Nodes and Version Control in an Open Hypermedia System" - Information Systems Vol 20, No. 6, pp. 501-519, 1995.

[Hype 95] Soares, L.F.G. at alii - "HyperProp - Uma visão geral" - I Workshop sobre Sistemas Multimídia Distribuídos, São Carlos, SP, Julho de 1995.

[Tane 88] Tanenbaum, A.S. - "Computer Networks" - Second Edition - Prentice-Hall International Editions, 1988.

[STCN 92] Soares, L.F.G.; Tucherman, L.; Casanova, M.A.; Nunes, P.R.R.L. - "Fundamentos de Sistemas Multimídia" - VIII Escola de Computação, Gramado, 1992.

[Turl 93] Turletti, T. - "A H.261 Software Codec for Videoconferencing over the Internet" - INRIA Research Report, No. 1834, January 1993.

[Turl 95] Turletti, T. - "Contrôle de Transmission pour Logiciel de Vidéoconférence sur l'Internet" - Thèse de Doctorat, L'Universite de Nice - Sophia Antipolis, Avril 1995.

[Wall 91] Wallace, G.K. - "The JPEG Still Picture Compression Standard" - Digital Equipment Corporation (Submetido para publicação em dezembro de 1991 em "IEEE Transactions on Consumer Eletronics"), 1991.