NetGiz - Geração Automática de Conteúdo Cognitivo Adaptável

Renata Falcão Corrêa, Luiz Fernando Rust, Luci Pirmez

Núcleo de Computação Eletrônica – (UFRJ)
Caixa Postal 2324 – 20001-970 – Rio de Janeiro – RJ – Brazil

re.falcao@br.inter.net, rustlf@utrc.utc.com, luci@nce.ufrj.br

Abstract. *This paper describes a strategy to overcome professor's low technological knowledge, by separating the authoring process into two phases: the first automatically build the multimedia document by capturing some classroom elements; the second aggregates the document adaptability creating a control archive. This archive is an element from the ServiMídia [1] adaptation mechanism, project developed at NCE/UFRJ. The strategy allows various approaches' specifications. In order to compare and validate those approaches a perceptual quality metric is used.*

Resumo. *Este artigo descreve uma estratégia para suprir a deficiência de conhecimento tecnológico dos professores, dividindo o processo de autoria em duas fases: a primeira consiste na elaboração do documento multimídia, automatizada pela captura de elementos da aula presencial; a segunda agrega a adaptabilidade gerando um arquivo de controle. Este arquivo é um elemento do mecanismo de adaptação proposto pelo ServiMídia [1], projeto desenvolvido pelo NCE/UFRJ. A estratégia adotada permite a especificação de vários métodos de adaptação. Para a análise destes métodos foram utilizadas métricas específicas de avaliação perceptiva da qualidade (QoP).*

1. Introdução

A demanda por aprimoramento profissional e o crescimento tecnológico geram desafios para o sistema educacional, que procura se adequar à nova necessidade do mercado. A inovação ocorrida tem como um de seus representantes a "Educação à Distância". Nesta nova forma de aprendizado, o contato entre o professor e o aluno ocorre de forma indireta. Logo, o material utilizado deve estimular o aprendizado sem a presença física do professor. As ferramentas de geração de conteúdo devem viabilizar esse processo e fornecer os meios necessários para que o professor estruture sua aula. O uso de recursos multimídia torna-se inevitável para o aprimoramento do material didático, porém aumenta a complexidade para os professores.

Visto que muitos professores não possuem conhecimentos em informática, o processo de geração do material didático fica comprometido. A maioria dos métodos de autoria de documentos multimídia, como TIEMPO [2], MAEstro [3] e o CMIF [4], não proporcionam interfaces amigáveis e flexíveis aos autores, prejudicando o desenvolvimento desses documentos. A solução utilizada pelo Classroom 2000 [5], embora mais automatizada, não prevê a adaptação dos documentos na rede.

O objetivo deste artigo é apresentar uma estratégia que simplifica a autoria do documento multimídia cognitivo mantendo a flexibilidade que os sistemas distribuídos demandam. Para isso, a estratégia proposta é dividida em duas fases: a primeira automatiza o processo de autoria por meio da captura de elementos da aula presencial do professor, construindo o documento multimídia; a segunda gera o arquivo de controle proposto pelo ServiMídia [1], no qual estão as informações de adaptação do documento original.

Na primeira fase, a estratégia é capturar individualmente os elementos da aula presencial, flexibilizando o processo de transmissão e adaptação do documento multimídia pela rede. Os elementos são capturados por diferentes dispositivos e sincronizados internamente através de marcas de tempo. Para a elaboração do documento multimídia é utilizada a linguagem SMIL (*Synchronized Multimedia Integration Language*) [6].

Na segunda fase, o documento SMIL é interpretado e suas mídias são avaliadas, de acordo com parâmetros pré-fixados, para determinar qual delas precisa ser adaptada. Com os resultados, são especificados os parâmetros de QoS (Qualidade de Serviço), as mídias alternativas e as condições para que as mesmas sejam ativadas durante a adaptação. As informações obtidas são armazenadas no arquivo de controle, descrito através da linguagem SCL (*SMIL Control Language*) [1].

Para a implementação das funcionalidades de captura e adaptação é proposta a utilização do conceito de fábricas. O objetivo principal da fábrica é a formalização dos processos, com a definição das etapas e tarefas. Neste artigo, são definidas duas fábricas: a Fábrica de Geração de Conteúdo e a Fábrica de Agregação da Adaptabilidade. Embora possuam funcionalidades diferentes, seus objetivos focam a automação do processo de elaboração de um documento multimídia adaptável.

Com as especializações das máquinas e a definição das etapas do processo, é possível utilizar diferentes abordagens para solucionar um único problema. Na especificação da Fábrica de Agregação da Adaptabilidade serão apresentados quatro métodos para gerar o arquivo de controle. Tais métodos são baseados na automação do processo de edição e diferenciam-se pelo grau de interação com o usuário e pelo modo de manipulação das mídias. Para a avaliação das propostas, as fábricas foram implementadas em uma única ferramenta, denominada NetGiz.

Para o material didático, é imprescindível que a apresentação adaptada não altere nem prejudique os aspectos cognitivos da apresentação. Portanto, foi desenvolvida uma metodologia para validação perceptiva das apresentações original e adaptada. Essa metodologia mapeia os requisitos de QoP (Qualidade Perceptiva) em requisitos de QoS e valida a apresentação utilizando critérios perceptivos. A avaliação permite também comparar os métodos de geração dos documentos adaptados.

2. Projeto ServiMídia

Para suprir os desafios enfrentados pelos documentos multimídia na transmissão em meios distribuídos, o projeto ServiMídia [1] propõe o desenvolvimento de aplicações multimídia adaptativas dentro do contexto de ensino à distância. Esse tipo de aplicação requer a absorção dos aspectos cognitivos da apresentação pelos usuários finais/alunos,

apesar do ambiente de execução (Internet ou Intranet) não os favorecerem. Para que diante da necessidade de adaptação da apresentação original o conteúdo não seja perdido, é preciso manter um certo nível de qualidade entre a apresentação adaptada e a original.

Basicamente, a estratégia especifica arquivos de controle que são associados a documentos multimídia convencionais. Os arquivos de controle contêm informações necessárias para a realização da adaptação de um documento multimídia. Por meio dessa metodologia, o autor define relações semânticas entre as mídias e, com isso, mantém o aspecto cognitivo das apresentações.

Dentro dessa proposta, há a definição de dois arquivos distintos. O primeiro utiliza a linguagem SMIL 1.0 para especificar a apresentação original, que seria executada em um sistema com todos os recursos necessários disponíveis. Para o segundo foi especificada a linguagem SCL, baseada em XML (*Extended Markup Language*). No arquivo SCL são descritas as variantes de QoS para os elementos multimídia que compõem a apresentação original, as mídias alternativas e as condições para que as mesmas se tornem ativas. A linguagem SCL é flexível e permite ao autor especificar inconsistências no arquivo de controle. Essas inconsistências foram apontadas e solucionadas em [7] por meio da validação da estrutura de adaptação.

Este artigo está voltado para o ambiente adaptativo proposto pelo ServiMídia. Por isso, o documento original e o arquivo de controle, elaborados pela ferramenta NetGiz, serão descritos, respectivamente, pelas linguagens SMIL e SCL.

3. Estratégia de Geração Automática de Conteúdo

O estudo das metodologias existentes e o objetivo de simplificar o processo de geração de conteúdo, produzindo um material que estimule o aluno, resultou na busca pela automação do processo de autoria. Para facilitar a geração de conteúdo, optou-se por trabalhar sobre a aula presencial do professor. Pois, muitas vezes, o professor possui sua aula estruturada no papel, mas não sabe como utilizar o computador para transmiti-la pela rede.

Embora a captura da aula por câmeras de vídeo seja o processo mais simples, ela não é suficiente. Os arquivos produzidos são muito grandes e pouco interativos. O ambiente distribuído impõe que o arquivo a ser transmitido consuma a menor largura de banda possível. Portanto, a geração do material didático deve considerar os limites da rede. De acordo com todas as necessidades levantadas, optou-se por dividir a aula em diversos objetos. Esta separação permite o manuseio independente e interações direcionadas. Os objetos podem ser transmitidos de acordo com suas características e individualmente tratados para que não onerem a rede.

A estratégia proposta neste artigo é voltada para o professor e baseada no conceito de fábricas. A principal característica das fábricas é a formalização dos processos e produtos, trabalhando em linha de produção, com etapas e tarefas perfeitamente definidas. Com a alta especialização das máquinas, que compõem a fábrica, cada uma garante a produtividade da etapa de produção em que está engajada, trazendo uma série de benefícios. A fábrica definida tem o objetivo de gerar um documento multimídia de forma automática. Dentro dessa definição, as máquinas são

especializadas na captura de mídias que formam o ambiente de sala de aula virtual. Para a especificação das máquinas foram selecionados os elementos principais da aula presencial: a fala do professor; a escrita no quadro; a sala de aula; e o personagem do professor. Denominou-se Fábrica de Geração de Conteúdo a fábrica responsável pela captura dos elementos da aula presencial.

3.1. Fábrica de Geração de Conteúdo

A Fábrica de Geração de Conteúdo é composta por cinco máquinas: (i) Máquina Captura Quadro, (ii) Máquina Captura Mídias, (iii) Máquina Captura Voz, (iv) Máquina Captura Professor, e (v) Máquina SMIL. Além das máquinas, há o Centro de Controle, responsável por gerenciá-las. A fábrica é ilustrada na Figura 1.

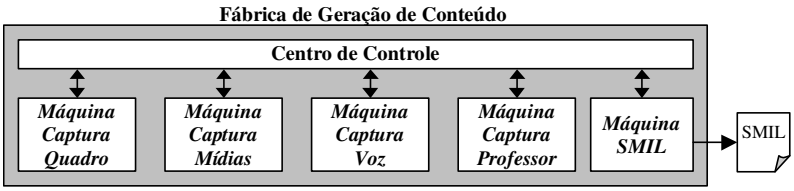

Figura 1 -Fábrica de Geração de Conteúdo

As quatro máquinas de captura agem de forma independente, gerando fluxos separados para cada tipo de captura. Esses fluxos são armazenados com marcas de tempo e organizados pelo Centro de Controle. A definição da funcionalidade de cada módulo será apresentada a seguir:

Centro de Controle

O Centro de Controle possui duas funções principais. A primeira é a interface com o usuário, que apresenta-se de forma amigável e simples para que o professor acione a captura de sua aula.

A segunda função do Centro de Controle é comandar as máquinas. O Centro é responsável por ativá-las e controlar as ações resultantes da interface com o usuário. Nesse comando, ele é capaz de requisitar as mídias capturadas e organizá-las estruturando o documento final.

Máquina Captura Quadro

O quadro negro representa o componente principal da sala de aula. É nele que o professor escreve os principais tópicos que serão abordados e complementa suas explicações orais. A função da Máquina Captura Quadro é explorar os recursos do ambiente de desenvolvimento para capturar as anotações. A mídia capturada é armazenada com estampas de tempo para manter o sincronismo com os produtos das demais máquinas. Essa máquina implementa as funcionalidades de gravação, pausa, finalização, retrocesso e avanço.

O processo de captura do quadro pode ser feito com o auxilio do quadro eletrônico branco [8], ou de dispositivos como o Mimio [9], especializados na digitalização dos dados escritos sobre o quadro. A utilização do quadro branco eletrônico pelo professor pode ser observada na metodologia de edição Classroom 2000 [5], na qual o professor escreve no quadro eletrônico e suas anotações aparecem nos computadores dos alunos.

Os dispositivos de captura digital dos quadros se conectam a um computador ou a um projetor LCD (*Liquid Cristal Display*) para criar uma grande tela interativa. O usuário pode controlar o ambiente do computador, bem como fazer anotações na área ativa de trabalho em tempo real. Todas as funções de programa e anotação são transferidas para o computador através de canetas eletrônicas.

Máquina Captura Mídias

A Máquina Captura Mídias surgiu como alternativa para os professores que estão habituados a utilizarem mídias (como transparências, vídeos, imagens e vídeos) em sua aula. A função desta máquina é permitir a inclusão e exibição de mídias conforme determinado pelo usuário. A exibição é monitorada de forma que o instante de apresentação e a permanência em tela de cada mídia sejam armazenados com estampas de tempo. Essa máquina implementa as funcionalidades de gravação, pausa, finalização, retrocesso e avanço.

Máquina Captura Voz

A explicação oral do professor permite compreender melhor o conteúdo cognitivo. A Máquina Captura Voz, comunica-se com o dispositivo de áudio e grava o fluxo oriundo do microfone. O áudio, assim como os demais fluxos, também são marcados com o instante de tempo inicial e final. Essa marcação é necessária para casos nos quais o professor gere uma pausa na gravação. Assim como as demais, ela implementa as funcionalidades de gravação, pausa, finalização, retrocesso e avanço.

Máquina Captura Professor

A figura do professor é essencial em um documento que visa a recriar o ambiente de sala de aula. A função da Máquina Captura Professor é capturar a imagem do professor. As mídias resultantes, assim como as demais, são marcadas com estampas de tempo. Ela também implementa as funcionalidades de gravação, pausa, finalização, retrocesso e avanço.

A problemática envolvida na captura do professor está relacionada ao ambiente distribuído. A simples filmagem da aula não apenas cria um ambiente paralelo, no qual o vídeo e os demais fluxos são apresentados em regiões diferentes, como também sobrecarrega o documento multimídia ao requisitar mais recursos. A solução é capturar apenas a figura do professor, separando-o do resto da cena.

Existem diversos métodos para a realização dessa tarefa como a técnica do fundo azul, a câmera tridimensional [10] e até algoritmos complexos [11][12][13] que, além de extraírem a imagem do professor, geram um objeto 3D com as informações capturadas.

A técnica do fundo azul foi imediatamente descartada, pois era preciso que o professor realizasse sua aula sem o quadro. Os algoritmos estudados utilizam diversas visões para construir um objeto 3D. Cada visão é representada por uma imagem contendo um perfil do professor. As imagens são posteriormente combinadas por meio de cálculos vetoriais gerando o personagem 3D. Mas o fato de separarem o processo de captura do professor em dois passos prejudica a sincronização com os demais fluxos que são capturados.

A câmera tridimensional, desenvolvida pela 3DV Systems, é capaz de captar a distância relativa de cada pixel de um quadro. Sua tecnologia permite que o usuário

capture um objeto 3D, como o professor, e manipule seus dados. A essência dessa tecnologia é a captura simultânea de cor (RGB) e profundidade (Z) para cada pixel da imagem, em vez da leitura pixel a pixel. Seu funcionamento pode ser resumido da seguinte forma: quando a luz, projetada da câmera, atinge a superfície de um objeto na cena, algumas componentes da luz retornam ao detetor. A intensidade delas indica a reflexividade, ou seja, a cor (RGB) da atual região iluminada do objeto alvo, o tempo entre a transmissão e o retorno dessas componentes é utilizado para calcular a distância até o ponto iluminado, retornando a profundidade (Z). A posse da informação de profundidade permite a separação automática de objetos da imagem em tempo real.

Máquina SMIL

Para compor o documento final, aplicou-se a linguagem descritiva SMIL 1.0 [6]. A linguagem SMIL 1.0 possui diversos elementos que permitem definir o instante de apresentação de cada objeto, a região na qual este objeto será exibido e o sincronismo dos fluxos de mídia. Além disso, é a linguagem utilizada pelo projeto ServiMídia. A Máquina SMIL é responsável por descrever a aula capturada utilizando o DTD (*Document Type Definition*) da linguagem SMIL. Ela recebe as informações sobre os fluxos de mídias capturados e os transforma em um documento representando a apresentação.

3.2. Execução da Proposta

O ambiente de desenvolvimento utilizado para a elaboração do protótipo é composto por um computador pessoal com 64Mb RAM, placas de som e de vídeo, microfone profissional WM-308, WebCam da Creative Labs e sistema operacional Windows. Como linguagem de programação, foi utilizado o Borland Delphi versão 5.0.

Toda a programação utilizou a técnica de orientação a objetos, que permite a reutilização de código. O conceito de componentes visuais existente no Borland Delphi foi empregado também com o propósito de reutilizar os códigos de forma fácil. O Borland Delphi permite a criação de componentes visuais e a adição dos mesmos em sua barra de objetos. Esta funcionalidade facilita a inserção dos componentes nos formulários. Os objetos transformados em componentes visuais são aqueles que lidam com os dispositivos de captura. Os componentes são: TMicrophone (Voz), TMediaShow (Mídias), TCamera (Professor) e TWhiteBoard (Quadro).

Para a implementação das máquinas algumas estratégias foram utilizadas e encontram-se descritas a seguir:

Máquina Captura Quadro

As informações enviadas para o computador podem ser interpretadas de duas maneiras: por meio de imagens que representam o quadro escrito ou de coordenadas que indicam a posição da caneta em determinado instante da apresentação. A utilização de imagens onera os arquivos multimídia, pois requisitam certos recursos da rede. Por isso, a leitura das coordenadas da caneta apresenta-se como a solução mais viável para a geração de documentos sem que se sobrecarregue a rede.

Implementada pela classe TWhiteBoard a captura do quadro foi simulada através do armazenamento das coordenadas do mouse na tela. Em um ambiente gráfico, definiu-se um editor gráfico através do qual o usuário pode realizar suas anotações. Para manter

o sincronismo com os demais fluxos, a leitura das coordenadas é marcada com instantes de tempo. O resultado é o conjunto de pontos, representados pelo trio (x, y, t_i), no qual x e y são as coordenadas da caneta no quadro e t_i o instante de sua captura. Esse trio é adaptado para a região de espaço e tempo em que ocorre a apresentação do documento.

Os pontos são salvos em um arquivo com extensão *wbp* e identificados como animação no arquivo SMIL final. Os arquivos *wbp* são caracterizados por serem pequenos e não necessitarem de muita banda para sua transmissão.

Máquina Captura Voz

A captura da voz foi implementada pela classe TMicrophone, que usa a MCI (*Media Control Interface*) para a interface com os dispositivos de áudio. O áudio gravado é marcado com o instante inicial e final da gravação e devidamente inserido no documento multimídia. A desvantagem na utilização da MCI está na codificação do arquivo final, ele utiliza a codificação PCM (*Pulse Code Modulation*) e 8 bits de armazenamento das amostras, o que prejudica a qualidade da mídia gerada.

Máquina Captura Professor

A captura do professor utilizou o pior caso, no qual todo o ambiente é filmado e transformado em um vídeo. Para os testes de captura do professor foi utilizada uma WebCam. Todavia, os resultados não foram satisfatórios, pois a qualidade do vídeo obtida era ruim e, embora tenha sido criado um componente visual para a captura da câmera, ele não foi utilizado durante a fase de testes da ferramenta. Esse fato não impossibilitou a captura, apenas alterou a estrutura do documento final. Ao invés de ser apresentado um ambiente de sala de aula virtual, foi criado um ambiente alternativo, no qual o personagem do professor não é figura essencial. O componente gerado, TCamera, é apresentado para fins de compreensão da metodologia de implementação.

Máquina Captura Mídias

Essa máquina foi implementada pela classe TMediaShow. Ela permite a inserção e ordenação das mídias de acordo com a apresentação. Cada mídia tem o formato (nome, t_i, t_f), no qual t_i é o instante de exibição inicial e t_f o instante final da mídia na tela. De acordo com o tipo da mídia um dispositivo de exibição é acionado para apresentá-la.

Máquina SMIL

A Máquina SMIL foi implementada pela classe TParserSMIL, o algoritmo utiliza a estrutura montada pelo Centro de Controle para transcrever os requisitos para a estrutura SMIL. A construção do documento de controle é baseada no DTD do SMIL 1.0.

4. Agregação da Adaptabilidade

A estratégia de agregação dos requisitos de adaptabilidade segue o mesmo conceito de fábrica desenvolvido para a geração automática de conteúdo. Para que o produto final possua qualidade, é preciso definir máquinas com objetivos específicos. Nesta fábrica, diferente da Fábrica de Geração de Conteúdo, o processo de produção não é paralelo, gerando a necessidade do detalhamento das fases para que cada executor tenha perfeito conhecimento do que o antecede, assim como dos eventos subseqüentes. Ao tornar o

processo modular e serial, cada fase pode ser explorada ao máximo para que a ferramenta gere, com qualidade, o arquivo de controle.

4.1. Fábrica de Agregação da Adaptabilidade

O objetivo da Fábrica de Agregação da Adaptabilidade é a geração do arquivo de controle de acordo com o documento multimídia gerado pela Fábrica de Geração de Conteúdo. Para compor esse processo a fábrica é composta por seis módulos: (i) Centro de Controle; (ii) Máquina SMIL; (iii) Máquina de Homologação; (iv) Máquina de Adaptação; (v) Máquina SCL e (vi) Banco de Parâmetros. A arquitetura proposta pode ser observada pela Figura 2 e as máquinas serão descritas a seguir.

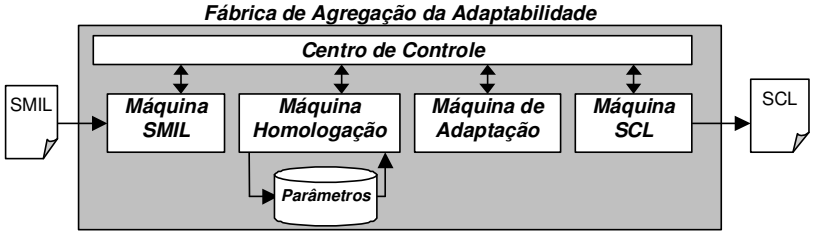

Figura 2 – Principais Componentes da Edição Automática

Máquina SMIL

A Máquina SMIL abre o documento SMIL 1.0 e gera o diagrama de estados da apresentação. No diagrama, cada estado possui seu instante inicial t_i, seu instante final t_f e as mídias ativas AM (*Active Medias*). O diagrama é construído da seguinte forma: partindo do estado inicial S1 ($t_i = 0$), no qual nenhuma mídia está ativa, cada estado S_i subseqüente possui o grupo de mídias ativas AM (*Active Medias*) em seu intervalo de tempo (t_i a t_f). As transições têm origem no estado S_i e destino S_{i+1}, sendo elas formadas pelo início (t_i) ou término (t_f) da apresentação de uma mídia. Portanto, o intervalo de tempo de um estado s_i é determinado pelo instante de início ou término de mídias. As Figura 3.a e b ilustram o diagrama originado dos arquivos SMIL apresentados na mesma figura. A construção do diagrama baseia-se nas informações contidas nos elementos <par> e <seq> do documento SMIL.

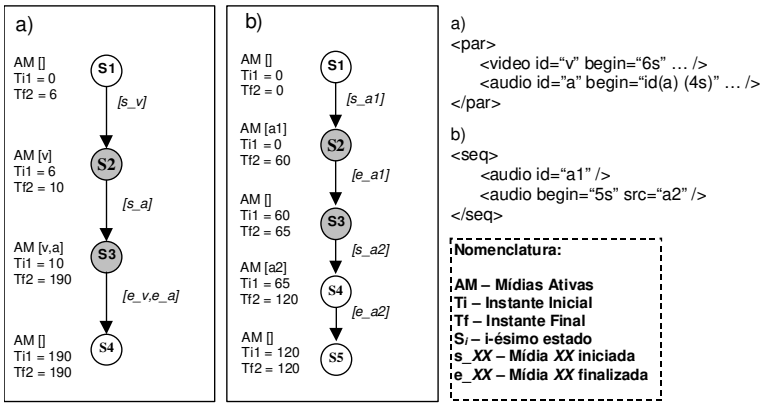

Figura 3 – Diagrama de estados do arquivo SMIL

Máquina de Homologação

Essa máquina analisa os estados do diagrama e aponta as mídias que precisam ser adaptadas. Cada estado é analisado levando em consideração o ambiente de rede e a importância das mídias no contexto da apresentação. Este fato permite flexibilizar a geração do documento adaptado, tornando-o mais próximo do original e de acordo com os requisitos da rede em que será transmitido.

A informação das mídias que precisam ser adaptadas é preenchida levando em consideração as dependências entre elas e a possibilidade de inconsistência no documento de controle. O esquema adotado pelo ServiMídia, de definição de links condicionais entre pares (ou grupos) de mídias, permite que ocorra um encadeamento indesejável de links provocando algum tipo de inconsistência semântica nas informações de adaptação. As inconsistências podem causar um comportamento indesejável nos mecanismos de adaptação.

Como abordado em [7], esse procedimento é essencial para que o arquivo final esteja consistente. As inconsistências mais comuns são: (i) links condicionais redundantes; (ii) expressões condicionais de links, nas quais os objetos são mutuamente exclusivos, impedindo que o mesmo seja disparado; e (iii) links dependentes do tempo de execução das mídias.

Banco de Parâmetros

O Banco de Parâmetros é modelado de forma a parametrizar as informações que serão utilizadas pela Máquina de Homologação. Desta forma, ele contém os requisitos da rede, da prioridade de adaptação das mídias e das possíveis dependências entre elas. A prioridade das mídias é classificada utilizando a escala de 1 a 10. Sendo o valor 10 correspondente a alta necessidade de adaptação, 5 a necessidade regular e 1 a necessidade baixa de adaptação.

Máquina de Adaptação

A Máquina de Adaptação recebe as mídias selecionadas pela Máquina de Homologação e gera mídias alternativas. O produto dessa máquina é a estrutura de adaptação do arquivo SMIL. Como será apresentado mais adiante, foram sugeridas duas modalidades de edição para a implementação dessa máquina.

Máquina SCL

A Máquina SCL recebe a estrutura adaptada e gera o arquivo de controle SCL baseado na linguagem definida pelo DTD do SCL. A linguagem SCL é formada pelas informações de controle relativas às mídias presentes no arquivo SMIL original.

Centro de Controle

O Centro de Controle é a unidade responsável pela comunicação entre o usuário e as máquinas. Seu principal objetivo é fornecer interface clara e amigável ao professor. A idéia é ajudar ao professor a construir um documento adaptado para sua apresentação, utilizando todos os recursos possíveis para minimizar a falta de conhecimento tecnológico.

A segunda função desempenhada pelo Centro de Controle é o comando das máquinas por meio da coordenação das etapas de geração do arquivo SCL. O Centro de Controle é capaz de ativar e desativar uma máquina de acordo com a etapa do processo.

4.2. Métodos de Adaptação

Neste artigo serão focados quatro métodos que permitirão analisar a melhor estratégia de adaptação. Esses métodos variam quanto a forma de adaptar as mídias e quanto a escolha de quais mídias serão adaptadas. A seleção das mídias segue as prioridades definidas no Banco de Parâmetros. Cada mídia recebe um valor de 1 a 10, sendo 10 correspondente a alta necessidade de adaptação, 5 a necessidade regular e 1 necessidade baixa de adaptação. A Máquina de Homologação utiliza dois limites para a seleção das mídias prioritárias para adaptação, o limite 5 (cinco) quando há uma necessidade alta de adaptação e o limite 7 (sete) quando há uma necessidade baixa de adaptação, i.e., as mídias acima destes limites são selecionadas para serem adaptadas. Todavia, caso nenhuma mídia esteja acima do limite é selecionada a mídia mais prioritária. Os métodos descritos a seguir, utilizam as prioridades das mídias e os limites para indicar quais devem ser adaptadas.

Método Redução

O objetivo deste método é gerar rapidamente o documento adaptado sem a ajuda do professor. Para manter a semântica e diminuir a banda passante, a estratégia utilizada é a redução da qualidade da mídia por meio de algoritmos de compressão. Mídias como vídeo e áudio são consideradas críticas para a rede e portanto, são priorizadas para serem adaptadas. Para os testes, a codificação das mídias foi realizada por outras ferramentas de edição.

Método Cortes Banda Máxima

A estratégia deste método é realizar cortes sobre determinadas mídias sem degradar a qualidade das demais. O áudio, segundo Steinmetz [14], é essencial para a compreensão do contexto da apresentação e no contexto da educação, a explicação oral do professor é extremamente importante para a perfeita compreensão do conteúdo cognitivo. Logo, o áudio é considerado mídia essencial do documento. Este tipo de mídia necessita de razoável largura de banda para ser transmitida. Por isso, este método é denominado Corte Banda Máxima, i.e., o áudio é a última mídia que deverá sofrer alguma alteração. Todas as demais mídias que se apresentarem paralelamente poderão ser retiradas da apresentação após a adaptação.

Método Cortes Banda Mínima

O método Cortes Banda Mínima propõe o corte de mídias sem considerar sua importância no contexto do documento. As mídias são eliminadas de acordo com os requisitos da rede.

Método Caminho Alternativo

O Caminho Alternativo tem como objetivo substituir as mídias mais críticas por outras mais simples, como a utilização de textos para a substituição do áudio e imagens para a substituição do vídeo. Para este método foram definidas duas modalidades de edição que facilitam a elaboração de mídias alternativas. Tais modalidades se diferenciam pelo grau de interação com o professor e são descritas a seguir:

- **Edição Inteligente**

A Edição Inteligente tem como princípio não usar a interação do professor. Por isso, são definidos processos que tentam automatizar a geração de mídias alternativas.

- VÍDEO: Para a adaptação do vídeo é utilizada a eliminação de quadros redundantes, originando uma seqüência de imagens que substitui o fluxo sem onerar a rede. Esse processo pode ser realizado automaticamente por meio de algoritmos de detecção de borda como o de Sethi e Patel [15].
- ÁUDIO: A opção para substituição do áudio é torná-lo textual. Os textos requisitam menos recursos da rede do que o áudio e conseguem passar a mesma mensagem. O processo de textualização da fala pode ser realizado por meio de dispositivos de reconhecimento de voz [16]. Estes são dispositivos de software que convertem o sinal acústico para o sinal digital e enviam a fala em forma de texto para o aplicativo. As ferramentas existentes utilizam dois métodos de reconhecimento de voz: (i) Ditado; e (ii) Comando e controle.

- **Edição Conduzida**

A Edição Conduzida, como o próprio nome já a define, tem o objetivo de criar novas mídias por meio da interação conduzida com o professor. Todos os passos percorridos pelo professor, para a elaboração de novas mídias, são guiados por mensagens explicativas. A vantagem desse método é que o novo documento gerado é elaborado sob a supervisão do professor, garantindo, portanto, a qualidade cognitiva do documento adaptado.

- VÍDEO: Para a adaptação do vídeo o professor deve selecionar imagens do próprio vídeo para substituí-lo. Para tal, os quadros do vídeo são dispostos lado a lado para que o professor selecione quais ele deseja manter na apresentação. Cada quadro selecionado é salvo com o padrão JPEG e inserido no documento adaptado de forma a preservar o tempo de execução do vídeo.
- ÁUDIO: Utilizando a mesma metodologia do método inteligente, o usuário é induzido a criar textos que substituam suas falas. A seqüência sonora é quebrada em pequenos intervalos, aos quais o professor deve associar um texto explicativo. Como o tempo total da apresentação deve ser preservado, o texto deve permanecer na tela de acordo com o intervalo da fala a que se refere.

4.3. Execução da Proposta

O ambiente de desenvolvimento se manteve o mesmo da Fábrica de Geração de Conteúdo. Como linguagem de programação, foi utilizado também o Delphi 5.0. Para o banco de dados foi utilizado o *Advatage DataBase Server*, um servidor de banco de dados gratuito e de fácil instalação. Para o Método Inteligente testou-se a tecnologia para o reconhecimento de voz utilizando a API, criada pela Microsoft, denominada SAPI (*Speech Application Programming Interface*) [17].

Os testes realizados sobre o processo de reconhecimento de voz não resultaram em mídias adaptadas coerentes com o áudio. Para que o reconhecimento produza algum resultado, é preciso que o dispositivo seja adequado à fala do usuário. No contexto do trabalho, significa que o professor precisaria treinar o dispositivo para que este reconhecesse sua fala. Além de impor essa atividade ao professor, os dispositivos de reconhecimento de voz existentes apresentam muitas falhas, reconhecendo uma

porcentagem baixa para as expectativas da ferramenta. Portanto, durante os testes, o modo de Edição Inteligente não foi utilizado.

5. Avaliação Perceptiva Pós-Adaptação

Ghinea, em [18], descreve um método para mapear os dados de QoP em parâmetros de QoS. De acordo com [18], cada apresentação multimídia pode ser caracterizada pelo seu dinamismo *D* e pela importância relativa de seus componentes (vídeo *V*, áudio *A* e texto *T*) como fontes de informação. Ghinea utiliza cinco parâmetros de QoS: taxa de erro (*BER*), perda de segmento (*SL*), ordem de segmento (*SO*), atraso (*DEL*) e variação do atraso (*JIT*). A Tabela 1 exibe a matriz de conversão entre QoP e QoS, que classifica como: baixo, médio e alto; a relevância dos parâmetros de QoS para cada mídia.

Mapeamento QoP para QoS		QoP			Fórmula
		Vídeo	*Áudio*	*Texto*	
QOS	*Taxa de Erro (bit)*	Baixo	Baixo	Baixo	$QoP = \frac{\Sigma (N_M * C_{iM} * P_{CM})}{\Sigma (N_M)}$
	Perda de Segmento	Baixo	Alto	Alto	
	Ordem de Segmento	Alto	Médio	Médio	
	Atraso	Médio	Médio	Baixo	
	Variação do atraso	Médio	Baixo	Médio	

Tabela 1 – Mapeamento de QoP para QoS

A relação obtida, Tabela 1 coluna "Fórmula", relaciona os parâmetros de QoS "C", com o número "N" de mídias "M" do documento (V, A e T) e o peso "P" de cada parâmetro de QoS. Estes pesos estão relacionado à rede em que a aplicação é executada. Logo, o valor final da QoP é calculado de acordo com o contexto da aplicação.

5.1. Critérios de Avaliação

A fim de analisar os métodos propostos para a agregação dos requisitos de adaptabilidade foram adotados sete critérios. Estes critérios serão descritos a seguir:

– **C1..C5: QoS das Mídias**

Cada apresentação adaptada será comparada à original de acordo com o somatório dos parâmetros de QoS. O mapeamento obtido por Ghinea é transformado em números na Tabela 2. Para as animações serão utilizadas as mesmas pontuações do texto, pois, na aplicação que foi desenvolvida, a animação corresponde a uma seqüência de pontos que representa o quadro branco. A imagem pode ser interpretada como sendo uma seqüência de pixel que, unidos constroem a informação final. A QoP da imagem se baseou também na pontuação do texto.

De acordo com a Tabela 2, cinco critérios são definidos: taxa de erro (C1), perda de segmento (C2), ordem do segmento (C3), atraso (C4) e variação do atraso (C5). Esses critérios são avaliados utilizando a fórmula de Ghinea [18], na qual os pesos dos parâmetros de QoS possuem o valor um. Como será visto mais adiante alguns critérios

Mapeamento QoP para QoS		QoP				
		Vídeo	*Áudio*	*Texto*	*Animação*	*Imagem*
QOS	*Taxa de Erro (bit)*	1	1	1	1	1
	Perda de Segmento	1	3	3	3	3
	Ordem de Segmento	3	2	2	2	2
	Atraso	2	2	1	1	1
	Variação do atraso	2	1	2	2	2

Tabela 2 – Mapeamento dos requisitos de QoP em QoS

– **C6: Carência de Contexto**

Os métodos propostos não alteram o contexto, entretanto, podem ocasionar perda de informação. Quando se trata de um documento educacional o conteúdo cognitivo deve ser assegurado para não prejudicar o aluno. Por isso, este critério possui os seguintes valores: 3 – Conteúdo deficiente; 2 – Houve perda de informação mas não prejudicou o conteúdo; 1 – Não alterou o conteúdo.

– **C7: Necessidade de Readaptação**

O documento adaptado não deve apresentar a necessidade de se readaptar, uma vez que sua qualidade perceptiva é pior que a do documento original. Conforme definido pela ferramenta, a necessidade de adaptação é verificada comparando-se a largura de banda da apresentação com os requisitos mínimo e máximo da rede. Se esta se apresentar acima do requisito máximo a necessidade de adaptação é alta, caso esteja acima do requisito mínimo a necessidade é baixa mas, é adaptável. Somente se a largura de banda estiver abaixo do mínimo é que a apresentação pode não sofrer adaptação. Com base nesses dados, esse critério tem os seguintes valores: 3 – Necessidade de adaptação; 2 – Necessidade baixa de adaptação; 1 – Sem Necessidade de adaptação.

5.2. Testes e Resultados Obtidos

Nesta seção serão apresentadas os testes e os resultados obtidos com a utilização das ferramentas de geração de conteúdo e agregação da adaptabilidade. Foram estipulados que o limite mínimo da rede é de 64 Kbps e o máximo 128Kbps.

Testes Propostos

Na descrição dos testes, as mídias serão definidas no formato: (nome, tipo, tamanho), i.e., uma mídia de nome T1, do tipo TXT e de tamanho 6Kb será definida como (T1,TXT,6Kb). As seqüências de apresentação das mídias em cada teste podem ser visualizadas pelos gráficos da Figura 4. Cada teste foi elaborado sobre os arquivos gerados pela Fábrica de Geração de Conteúdo, algumas mídias porém, foram usadas em mais de um teste. Quatro testes são propostos: Teste A – formado por dois vídeos, (V1, AVI, 571Kb) e (V2, AVI, 4807Kb), e um áudio, (A1, WAV, 4273Kb); Teste B – formado por um áudio (A1, WAV, 4273Kb) e uma animação (An1, WBP, 56Kb); Teste C – formado por três áudios, (A3, WAV, 130Kb), (A4, WAV, 596Kb) e (A5, WAV, 367Kb), e três imagens, (I1, BMP, 150Kb), (I2, BMP, 150Kb) e (I3, BMP, 150Kb); Teste D – formado por um vídeo, (V1, AVI, 571Kb), um áudio, (A1, WAV, 4273Kb), uma animação, (An2, WBP, 45Kb), e três imagens, (I4, BMP, 11Kb), (I5, BMP, 41Kb) e (I6, BMP, 4Kb).

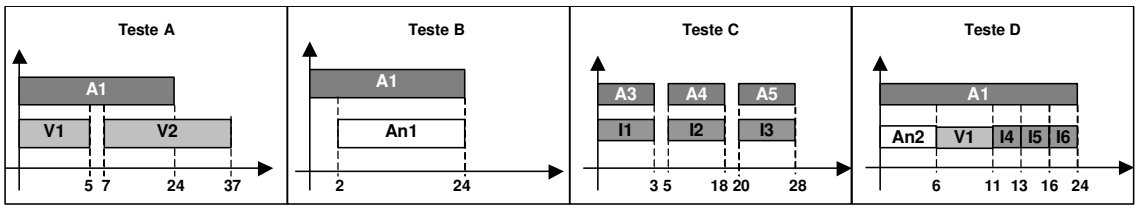

Figura 4 - Seqüência de apresentação dos testes

Resultado da Ferramenta de Agregação da Adaptabilidade

Para o cálculo da QoP dos critérios de 1 a 5 são importantes o número e os tipos de mídias existentes no documento multimídia. A Figura 5 representa o gráfico originado a partir das informações referente a avaliação de QoP realizada sobre os documentos adaptados.

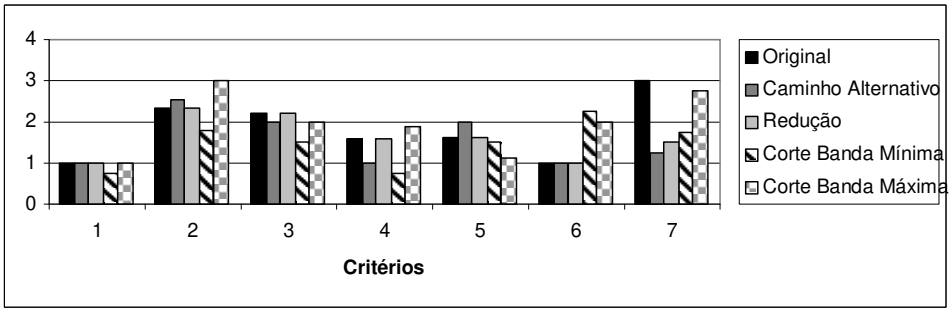

Figura 5 – Gráfico referente a avaliação de QoP

A média final dos métodos foi calculada adicionando peso dois aos critérios 4 e 7 e peso três ao critério 6. O critério 4 refere-se a variação de atraso, prejudicial por representar o fator de inquietude no usuário final. O critério 7 representa a necessidade de readaptação do documento, inviabilizando a alternativa sugerida. O critério 6 relaciona-se ao conteúdo da aplicação, caso haja alguma carência, o documento adaptado pode não estar passando o conhecimento de forma correta.

De acordo com os resultados obtidos, nos quais a menor média é o melhor resultado, os métodos Caminho Alternativo (média 1.46) e Redução (média 1.48) ficaram próximos. Porém, pela análise individual dos testes [19] pode-se perceber que o de Redução depende do método de compressão utilizado, i.e., se a compressão for alta e ainda mantiver a qualidade das mídias, este método possuirá uma média melhor que a do Caminho Alternativo. Contudo, se a compressão não resultar em uma boa qualidade, o documento adaptado sofrerá com a carência de contexto. Em compensação, o Caminho Alternativo por ser caracterizado pela diminuição da banda passante sem a alteração ou perda de contexto, mantém a qualidade do documento final em qualquer situação. Sua vulnerabilidade é também seu ponto forte, a elaboração de mídias alternativas com a ajuda do professor assegura a integridade de contexto mas, permite a criação de mídias que podem gerar a necessidade de readaptação.

A redução do fluxo de mídias realizado pelo Corte de Banda Mínima (média 1.64) favoreceu o resultado da avaliação sobre os critérios referentes à QoS. Entretanto, a perda de contexto, evidente no teste A que teve todas as suas mídias cortadas, caracterizou a média perceptiva elevada em relação aos dois primeiros métodos. O Corte Banda Máxima (média 1.96) provocou mudanças de contexto mantendo a necessidade de readaptação, fatores essenciais para a avaliação perceptiva.

6. Conclusão

A proposta deste trabalho focou o setor educacional e a problemática na elaboração de conteúdo voltado para o ambiente distribuído. O objetivo principal foi definir uma

metodologia que permitisse a autoria de documentos multimídia sem que fosse necessário o professor conhecer as tecnologias.

A utilização da estratégia de fábricas implicou as seguintes vantagens: (i) utilização de dispositivos específicos para a captura de cada ponto da aula; (ii) não necessidade de desenvolvimento integral da ferramenta para alteração da funcionalidade de uma determinada máquina; e (iii) possibilidade de especificação de diversos métodos de adaptação.

No processo de implementação da proposta verificou-se que alguns dispositivos produziam soluções de baixa qualidade como a WebCam, para a captura do professor, e a tecnologia de reconhecimento de voz, para a modalidade de Edição Inteligente. Pode-se perceber que ambos os dispositivos ainda não podem ser utilizados para ajudar no desenvolvimento de documentos multimídia.

Os resultados obtidos pelos testes mostraram que a adaptação pelo Caminho Alternativo foi a melhor opção. Além de reduzir a largura de banda da apresentação, ela mantém a semântica do documento pois, utiliza o modo de Edição Conduzida para a elaboração das mídias alternativas. O método de Redução apresentou também, resultados satisfatórios, porém, dependentes do tipo de compressão utilizado. Os demais métodos, como já era esperado, causaram a perda de conteúdo e a necessidade de readaptação, colocando-os acima da média.

A ferramenta implementada mostrou-se bastante funcional no sentido de automatizar o processo de geração de conteúdo, garantindo que este seja adaptável de acordo com a rede. Como trabalho futuro é sugerida a criação da ferramenta do aluno, na qual é simulada o ambiente de sala de aula virtual com iteração entre alunos.

Referências

[1] GOMES R. – "Autoria e Apresentação de Documentos Multimídia Adaptativos em Redes" – Dissertação de Mestrado UFRJ/IM/NCE, Rio de Janeiro, 85p., 2001.

[2] WAHL, T., WIRAG, S., ROTHERNMEL, K. – "TIEMPO: Temporal Modeling and Authoring of Interactive Multimedia". – IEEE International Conference on Multimedia Computing and Systems, p.274–277, Washington D.C., USA Fev. 1995.

[3] DRAPEAU G., GREENFIELD H. – "MAEstro — A Distributed Multimedia Authoring Environment" – In: Proc. Summer 1991 USENIX Conference, Nashville, TN, pp. 315–328. 1991

[4] ROSSUM G., JANSEN J., MULLENDER K.S., BULTERMAN D. – "CMIFed: A Presentation Environment for Portable Hypermedia Documents" – In: Proceedings ACM Multimedia'93, Anaheim CA, pp. 183–188, Agosto 1993.

[5] ABOWD G., ATKESON C., BROTHERTON J., BHALODIA J. – "Classroom 2000: A System for Capturing and Accessing Multimedia Classroom Experiences." – In: CHI'98 Demonstration Paper, Maio 1998.

[6] W3C. – "Synchronized Multimedia Integration Language (SMIL) 1.0 Specification". – W3C Recommendation, Junho 1998.

[7] RUST L.F., PIRMEZ L., GOMES R., CORREIA R., CORREA R. – "Validação da Estratégia de Adaptação Dinâmica de QoS do Projeto ServiMídia" – Simpósio Brasileiro de Redes de Computadores SBRC, Buzios – Rio de Janeiro, maio 2002.

[8] INTERACTIVE WHITEBOARD: Visual Collaboration Tools. Disponível na INTERNET via www.url: http:\\www.picturetel.com. Arquivo Consultado em 2000.

[9] MIMIO: Think it. Share it. Disponível da INTERNET via www.url: http:\\www.mimio.com. Arquivo Consultado em 2001.

[10] ZCAM: Real-Time 3D studio camera. Disponível na INTERNET via www.url: http:\\www.3dvsystems.com. Arquivo consultado em 2001.

[11] DYER C., DAVIS L., KLUWER – "Volumetric Scene Reconstruction from Multiple Views" In: Foundations of Image Understanding, 469–489, Boston, 2001.

[12] SAITO H., KANADE T. – "Shape Reconstruction in Projective Grid Space from Large Number of Images." – IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99), Junho 1999.

[13] SEALES B., WELCH G., JAYNES C. – "Real-Time Depth Warping for 3-D Scene Reconstruction" – IEEE Aerospace Conference 1999, Aspen, CO, Mar. 1999.

[14] STEINMETZ R. – "Human Perception of Jitter and Media Synchronisation" – IEEE Journal on Selected Areas in Communications, 14(1), pp. 61-72, 1996.

[15] SETHI I., PATEL N. – "A statistical approach to scene change detection," – In: IS&T SPIE Proceedings: Storage and Retrieval for Image and Video Databases III, vol. 2420, pp. 329– 339, San Jose , Fevereiro 1995.

[16] HOW SPEECH WORKS – MICROSOFT. Disponível na INTERNET via www.url:http://www.microsoft.com/speech/news/how.asp. Consultado em 2000.

[17] SAPI: SPEECH API 5.0 OVERVIEW. Disponível na INTERNET via www.url: http://www.microsoft.com/speech/technical/sapioverview.asp. Consultado em 2000.

[18] GHINEA G., THOMAS J.P. – "An Approach towards Mapping Quality of Perception to QoS in Multimedia Communications" – Proceedings of IEEE Multimedia Signal Processing Workshop, Copenhagen, Denmark, 1999.

[19] CORREA R. – "Projeto NetGiz – Geração Automática de Conteúdo Adaptável para Educação à Distância" – Dissertação de Mestrado UFRJ/IM/NCE, Rio de Janeiro, 95p., dezembro 2002.