

Uma Plataforma de Teste para o Projeto Auditivo de Ambientes Virtuais 3D com Propósitos Educacionais

Marcelo da S. Hounsell, Vanessa Suzuki, Avanilde Kemczinski, Isabela Gasparini

Departamento de Ciência da Computação (DCC) - Universidade do Estado de Santa Catarina (UDESC) – Campus Universitário S/N – Bairro Bom Retiro – 89223-100 – Joinville – SC – Brazil

{marcelo,dcc6vs,avanilde,isabela}@joinville.udesc.br

***Abstract.** 3D Virtual Environments (VEs) have many advantages to be explored from the educational perspective. The fact that VEs can provide multisensorial experiences including the auditory one is a plus of the Virtual Reality technology. However, this sense has been barely used, not because of the lack of technological solutions but, because of the difficulty to answer “what”, “where”, “when” and “why” using such resource. This paper presents a testbed infrastructure that can be used to try out different setups towards learning results. Therefore, design experiences can be collected in the form of guidelines for future VEs.*

***Resumo.** Ambientes Virtuais 3D (AVs) têm muitas vantagens a serem exploradas sob o ponto de vista educacional. O fato de os AVs proporcionarem experiências multisensoriais, incluindo o sentido auditivo, é um diferencial da tecnologia da Realidade Virtual. Entretanto, o sentido auditivo tem sido pouco utilizado, não por falta de soluções tecnológicas, mas, pela dificuldade em responder “onde”, “qual”, “quando” e “porque” utilizá-lo. Este artigo apresentará uma infra-estrutura de testes que pode ser utilizada para experimentar diferentes respostas e avaliar o seu impacto em face dos resultados de aprendizagem. Assim, experiências de projeto poderão ser acumuladas como recomendações para AVs futuros.*

1. Introdução

O uso da Realidade Virtual (RV) no Processo Ensino-Aprendizagem (PEA) tem se concentrado em aplicações que valorizam muito a capacidade do aprendiz “ver” um determinado conteúdo sob estudo. Apesar de este ser um dos benefícios mais claros do uso desta nova tecnologia, há uma dificuldade na concepção e desenvolvimento de Ambientes Virtuais (AVs) pela própria juventude da tecnologia. Neste sentido, várias são as iniciativas que apontam para um estudo mais aprofundado do processo de projeto de AVs [Eastgate 2001][Deol, Sutcliffe e Maiden 1999].

O sentido auditivo tem sido amplamente sub-utilizado [Stuart 2001, pág. 35], apesar de que, algumas pessoas são mais sensíveis ao áudio do que às imagens e pode-se utilizar o áudio para minimizar problemas com imagens [Pinhanez e Intille 2004]. Deveria-se gastar tanto tempo no projeto do uso do áudio quanto da imagem. Nesta ênfase, já foi dito que “os ouvidos podem guiar os olhos” [Stuart 2001, pág. 35].

Várias soluções tecnológicas voltadas para a veiculação não trivial de áudios (como som estéreo, *surround* e 3D, [Pinho 2002]) têm sido aperfeiçoadas mas, as questões relativas a “quando”, “o que”, “onde” e “porque” usar o sentido auditivo nos AVs não parece terem recebido a mesma atenção. Estas e outras perguntas ainda não têm respostas definitivas e as implementações atuais estão passíveis de especulações quanto a estes aspectos havendo, portanto, dúvidas na sua efetividade.

O objetivo deste trabalho é contribuir nas respostas a questionamentos quanto a “o que”, “quando”, “onde” e “porque” usar o sentido auditivo em um AV. Para tal é apresentado um Modelo de Informações de Projeto (MIP) auditivo (que responde aos “o que” e “porque”) e aprofunda-se nas questões relativas ao gerenciamento da execução dos estímulos auditivos (que auxilia responder o “quando” e “onde”).

A próxima seção resgata a importância do sentido auditivo no PEA e a visão da literatura sobre ele. A seção 3 apresenta o MIP aplicado ao recurso auditivo e, em seguida, é apresentado um recurso, denominado “escalador de áudio” para satisfazer os requisitos do MIP, quanto ao gerenciamento do áudio no AV. A seção 5 apresenta como este recurso foi implementado num AV estudo de caso e depois, na seção 6, como este pode ser generalizado para uma plataforma de testes para inúmeras decisões de projeto específicos quanto ao áudio. A seção 7 conclui este texto.

2. O Sentido Auditivo em Ambientes Virtuais na Literatura

O uso do sentido auditivo é muito importante para a RV pois tem sido associado ao aumento da interatividade [Sawhney e Murphy 1996][Marshall e Nichols 2004], do senso de imersão [Apaydın 2002][Damasceno et al. 2004] e da navegabilidade [Eastgate 2001, págs. 44 e 58] de AVs. Estes autores ressaltam ainda a preocupação em relação a quantos e quais áudios podem ser utilizados para não ocorrer uma sobrecarga cognitiva (sobrecarga na memória de trabalho), pois o ser humano possui uma capacidade limitada de processamento da informação [Sawhney e Murphy 1996].

Há os que consideram [Paivio 1991][Mayer e Sims 1994] que se for utilizada a redundância auditiva (reprodução de informações apresentadas ao usuário concomitantemente nas formas textuais e auditivas), esta irá melhorar a retenção da informação pelo usuário, pois vários estudos confirmam que a memória humana está baseada na codificação separada da informação visual e verbal. Até mesmo quando o *feedback* de som é irrealista ele tem sido defendido como apropriado em alguns casos [Eastgate 2001, pág. 125].

Apesar da defesa do uso de dois canais de veiculação (visual e auditiva) ser melhor que um em qualquer mensagem instrutiva, há quem defenda [Kalyuga 2000] que não se deve utilizar estes dois canais integrados indiscriminadamente, pois pode-se ter efeito negativo no processo de aprendizagem: no caso onde são apresentadas explicações auditivas e visuais equivalentes concorrentemente pode-se aumentar o risco de sobrecarregar os canais sensoriais gerando uma carga cognitiva adicional desnecessária interferindo no aprendizado. Pode-se também ter um fator negativo quando um formato instrutivo não é equivalente à experiência do usuário, pois quando estes são mais experientes, as explicações auditivas também podem ser redundantes fazendo com que a aprendizagem seja inibida por causa da sobrecarga cognitiva [Kalyuga 2000]. Assim, pode-se constatar que depois que os usuários se tornam mais

experientes no domínio, a vantagem relativa inicial de usar o recurso auditivo desaparece enquanto que, usar textos aumenta. Ainda, quando uma das informações sensoriais (visual ou auditiva) possui um atraso ou se adianta da outra, o cérebro percebe e “acorda” o usuário tirando-o da imersão, gerando a necessidade estrita de sincronismo entre o visual e o auditivo [Apaydin 2002][Damasceno et al. 2004].

Certos autores [Penney 1989][Paivio 1991] sugerem que sejam separados os processos para informação auditivas e visuais, pois a quantidade de informações que podem ser processadas usando canais auditivos e visuais poderiam exceder a capacidade de processamento humano. Assim, através da limitação da carga de memória, a aprendizagem poderá ser mais fácil, pois usando um formato instrucional dual no qual fontes separadas de informação são apresentadas com texto na forma auditiva reduziria a carga cognitiva gerando vários benefícios [Mousavi, Low e Sweller 1995][Tindall-Ford, Chandler e Sweller 1997]. Estes efeitos benéficos acontecem quando dois ou mais componentes de uma apresentação puramente visual são incompreensíveis sozinhos e devem ser integrados mentalmente para que possam ser entendidos.

Portanto, observa-se que há a necessidade de ser reconhecida a importância de um bom projeto do recurso auditivo, não só quanto aos aspectos tecnológicos de latência que podem influenciar a questão cognitiva, mas quanto a seleção da forma e conteúdos a serem utilizados. Entretanto, a literatura ainda carece de investigações neste último ponto, ao qual este artigo traz uma contribuição.

3. Modelo de Informações para Projeto (MIP)

Com o objetivo de se identificar os elementos envolvidos no Projeto Auditivo de Ambientes Virtuais, foi proposto um Modelo de Informação para Projeto (MIP) do sentido auditivo [Suzuki e Hounsell 2006] que é um conjunto de informações estruturadas e inter-relacionadas que orientam as possibilidades de projeto em um determinado contexto, para ser aplicado principalmente na fase de concepção de uma solução (de *software*), e que representa principalmente (mas não exclusivamente) os conteúdos em questão e suas formas (do que as tecnologias requeridas).

Um MIP para o projeto auditivo tem 3 (três) constituintes (ver detalhamento em [Suzuki e Hounsell 2006]): as **restrições**, que são as preocupações gerais relacionadas ao projeto do uso do recurso auditivo, alguns já mencionados na seção anterior; a **taxonomia**, que classifica os “elementos” auditivos envolvidos, e; os **relacionamentos**, que indicam como os elementos da taxonomia são organizados entre si para funcionar conjuntamente, de acordo com as restrições.

Quanto a taxonomia, os recursos foram sub-classificados quanto a presença (que pode ser ausente; presente e típica ou; presente e atípica); aos tipos (locuções ou sons); ao objetivo do recurso auditivo (que pode ser para o realismo; para quebra de monotonia; para promover emoções ou; para apresentar informações); quanto a utilidade (para indicação; motivação; educação ou; *feedback*); aos tipos de fontes (fontes ambiente-distribuída- ou pontual); ao formato (caricato; real genérico ou; real específico) e; quanto a procedência do áudio (gravado ou sintetizado).

Quanto aos relacionamentos, a fim de estabelecer de que forma o áudio se relaciona com o AV, há de se considerar como os recursos de áudio são gerados (de

onde aparecem) e disparados e em que ordem estes são apresentados aos usuários. As formas de “ativação” (*triggering*) respondem pelas primeiras preocupações e as estratégias de “gerenciamento” (*scheduling*) da ativação, respondem pelas seguintes, conforme descritos a seguir. A ativação pode ser pelo usuário (iniciadas por proximidade; visibilidade; seleção, manipulação ou; habilitação) ou pelo sistema (e este pode ser guiado ou inteligente). Quanto ao gerenciamento, ou seja, a forma de composição do áudio antes de apresentá-lo ao usuário, deve-se considerar as prioridades, as características de cada áudio, e as restrições já indicadas anteriormente.

4. Escalonador de Áudios

A fim de proporcionar o gerenciamento necessário para obter-se todas as funcionalidades do MIP, é imperativo que se tenham infra-estruturas de *software* que possam tratar os recursos auditivos de maneira a priorizá-los conforme seus significados. Para tal, ficou evidente a necessidade de implementar um recurso “escalonador de áudios” pois este não é nativo nas linguagens e ambientes para RV. Em especial, este não é nativo do VRML (*Virtual Reality Modeling Language*), o qual está sendo usado neste trabalho.

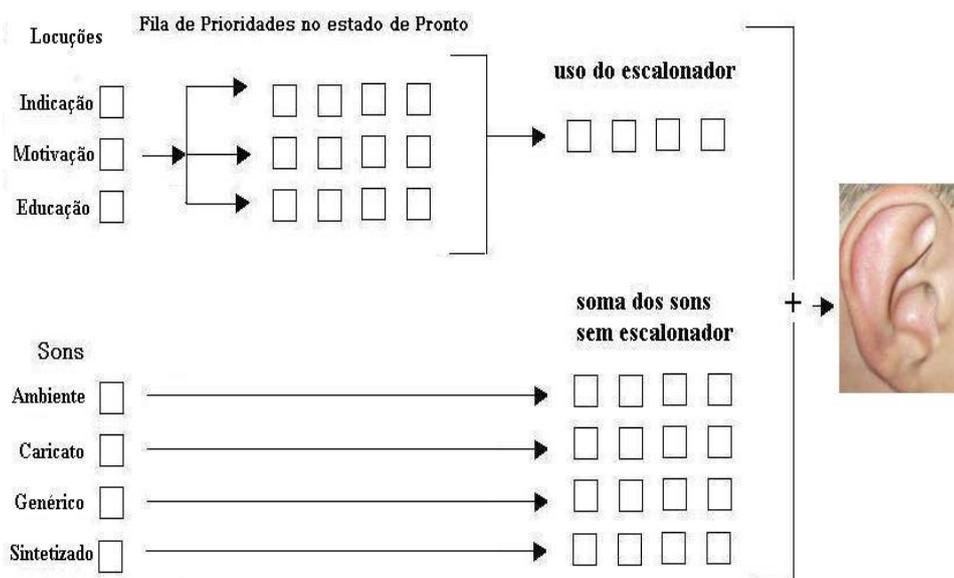


Figura 1: Esquema do Escalonador de Áudios

O objetivo então passa a ser a implementação de um escalonador com uma tecnologia que tenha as características do VRML, ou seja, código aberto, livre, gratuita, de fácil utilização e que se integre bem como o próprio VRML. Tentativas de utilizar-se a linguagem Java, integrada ao VRML, para gerenciamento da execução das mídias foi utilizada em trabalhos semelhantes, mas mostrou-se muito custosa em termos computacionais [Kniess e Souza 2004] e portanto foi descartada. Então, foram selecionados o PHP e o MySQL. O objetivo do escalonador é o de gerenciar a execução do áudio gerado pelo AV, seja ele locução ou sons, como mostra a Figura 1, sendo que as locuções necessitam de recursos especiais para serem veiculadas no ambiente e os sons não pois, são apenas adicionados ao ambiente de forma automática, considerando-

se a tecnologia VRML. A Figura 1 mostra que qualquer tipo de locução (gerada tanto pelo usuário quanto pelo sistema) é colocada na sua respectiva fila FIFO (*first in first out*, implementada em um Banco de dados MySQL). O “escalador” então consulta estas filas e gera uma outra fila, a de execução, que será “somada” (executada concomitantemente) aos sons não escalonados para finalmente ser apresentada ao usuário/aprendiz.

O escalonador proposto trata os tipos de áudio relacionados com a locução, pois esta possui alguns aspectos que devem ser atendidos para que a veiculação ocorra de maneira eficaz: a) as locuções não podem ser executadas paralelamente com outras locuções, pois podem ocasionar poluição sonora, e; b) existem algumas locuções que são mais importantes do que as outras, sendo que estas devem ser priorizadas.

4.1 Implementação

A Figura 2 mostra o esquema de implementação do escalonador. O PHP é responsável por realizar a conexão/integração entre o VRML e MySQL para que estas duas tecnologias possam trabalhar em conjunto. Para tanto, o PHP contém o código VRML e as chamadas para o Banco de Dados (BD), tratando o que deve ser inserido, executado e retirado do BD. O BD MySQL é utilizado para fazer o controle do escalonador de locuções através de uma tabela que contém uma fila de prioridades.

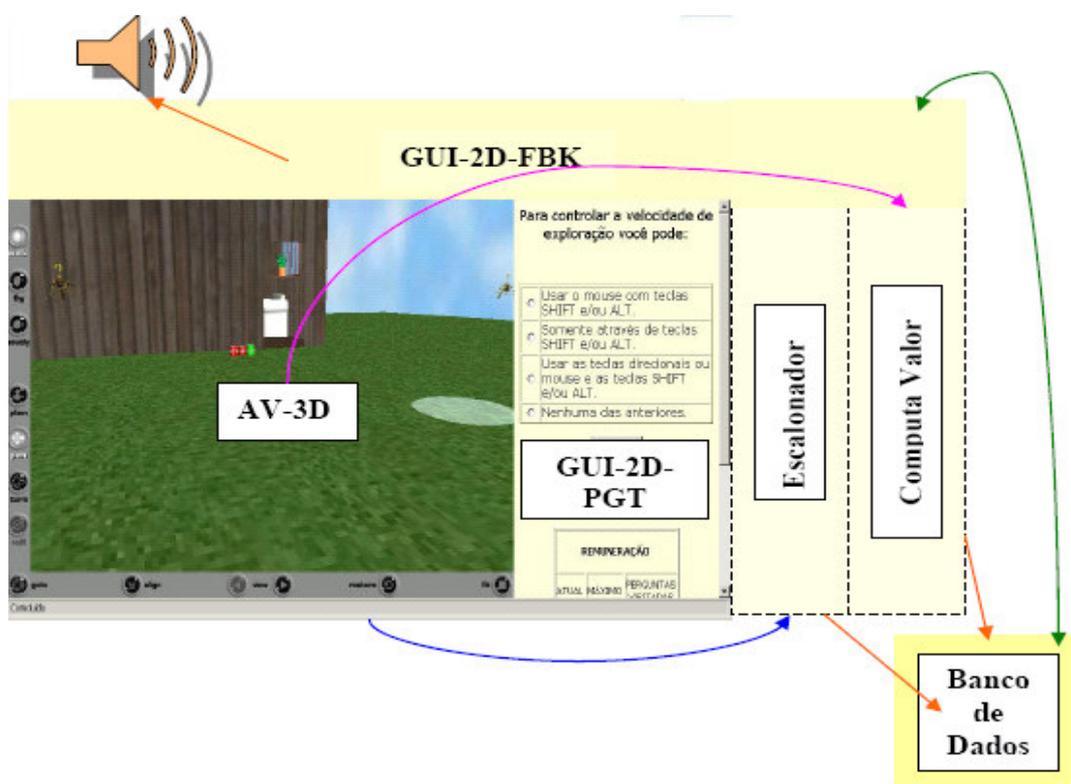


Figura 2: Implementação do Escalonador de Áudios em PHP e MySQL

O VRML é utilizado neste trabalho tanto para a modelagem do cenário virtual quanto para acionar os eventos auditivos do ambiente. Existem basicamente duas formas de interação realizadas: captura dos eventos do usuário ocorridos na aplicação 3D, e/ou eventos disparados pelo sistema. Essas interações ocorrem através do uso de

algumas técnicas disponíveis no próprio VRML, disparadas através do uso de *TouchSensors*, *ProximitySensors* e *TimeSensors*. Os efeitos sonoros são ativados através do nó *Sound* e *AudioClip*. Já as locuções são ativadas através de um código *javascript* contido dentro de um arquivo PHP.

Para a utilização do escalonador, foi necessário criar duas janelas auxiliares (“Escalonador” e “Computa Valor”), sendo que elas são ocultas (mostradas em pontilhado na Figura 2):

- as locuções geradas pelo sistema (de indicação/navegação, motivação, etc.) são tratadas pela janela oculta “Escalonador” e pelo escalonamento Banco de Dados Escalonador. A junção destes dois elementos forma a infra-estrutura do escalonador possibilitando organizar a reprodução auditiva no ambiente.
- as locuções de *feedbacks*, fatos, curiosidades e lembretes (geradas pelo usuário) são tratadas na janela oculta “Computa Valor” que é responsável por computar os impactos das interações do usuários (no caso, a pontuação atribuída às respostas). Primeiramente é verificado qual foi o estímulo obtido, depois é inserida a locução correspondente no BD (seta que apontam para o BD na Figura 2) para a futura execução juntamente com a exibição do texto (seta mais à direita na Figura 2).

A janela identificada como AV-3D envia os dados (a prioridade do som, a URL do som, o código do elemento chamador e, o tempo de duração) para a janela Escalonador. A janela primeiramente pesquisa alguns elementos referentes aos dados recebidos do AV-3D como: quanto tempo o evento possui, o caminho para a futura exibição do texto simultaneamente com o som (URL do som), entre outras coisas. De acordo com os dados passados, o código da janela irá inserir na tabela Escalonador do BD os atributos para a locução. Assim, a janela Escalonador é responsável por tratar o código recebido e obter os dados necessários para a inserção no BD.

Os sons (efeitos ou músicas) são executados diretamente pelo VRML (sem passar pelo escalonador, como indica a parte de baixo da Figura 1).

5. Estudo de Caso: Sherlock Dengue

Atualmente a aplicação é composta por três janelas (mas que não são necessárias em outro tipo de aplicação): AV-3D, GUI-2D-PGT e GUI-2D-FBK (os termos PGT e FBK são mnemônicos para “perguntas” e “feedback”). O MIP apresentado foi utilizado no projeto de áudio de um AV dedicado ao treinamento da inspeção de focos de dengue, denominado Sherlock Dengue [LARVA 2006]. Seguramente o MIP facilitou o processo decisório de qual e quando utilizar os estímulos auditivos no Sherlock Dengue, que é caracterizado por usar tecnologia Realidade Virtual Não Imersiva veiculada pela *web*, implementada basicamente com a linguagem VRML, uma vez que algumas possibilidades nem teriam sido reconhecidas se não fosse pelo MIP.

Considerando o MIP [Suzuki e Hounsell 2006], o Sherlock Dengue pode ser descrito como: projetado com objetos com áudio ausente se eles não são diretamente relacionados ao tema do AV (a dengue) e com áudio presente e típico para os objetos que são do tema. Tanto locuções quanto efeitos sonoros foram utilizados. As locuções foram objetivadas como fontes de informações do tipo educação e como *feedback*, todas

elas do tipo pontual, reais genéricas e gravadas (de textos existentes no próprio AV). As ativações foram selecionadas para serem iniciadas pelo usuário, tanto por proximidade quanto por seleção, sendo todos os estímulos priorizados e gerenciados por um “escalador de áudio”.

5.1. Interface do Sherlock Dengue

O Ambiente Virtual 3D, desenvolvido em VRML, é apresentado na janela AV-3D. Nas outras janelas da interface (GUI-2D) são expressas as informações formais textuais e auditivas referentes a estas informações textuais e/ou perguntas, conforme a Figura 2.

O AV-3D é um local onde existem diversos focos e falsos-focos da dengue. É o local onde o usuário, no papel de “inspetor virtual”, fará sua inspeção, em busca dos objetos “clacáveis”, com o intuito de ampliar seu conhecimento sobre a dengue tendo como motivação sua remuneração objetivando passar de fase. Este ambiente está povoado de recursos auditivos que auxiliarão na tarefa de inspeção (áudio por proximidade do foco, locuções dando dicas, locuções de curiosidades, dentre outras) ou de envolvimento do AV (emitindo efeitos sonoros mais próximos do real) (ver Figura 2). A cada fase alterna-se de um AV “apartamento” para um “barraco” para proceder as inspeções.

Para acionar o áudio do respectivo foco, primeiramente o sistema verifica se a pergunta respectiva já foi respondida, caso ela não tenha sido respondida, o sistema aciona o áudio de proximidade do foco (zumbido de um mosquito) através do nó *ProximitySensor* para indicar que o usuário está próximo do foco.

Os elementos na GUI-2D tanto na parte superior da tela (GUI-2D-FBK), como do lado direito (GUI-2D-PGT) possuem o objetivo de melhorar a apresentação gráfica 2D do sistema e obter uma maior interatividade 3D. A janela GUI-2D-FBK apresentará informações textuais e auditivas formais sobre o conteúdo, os quais foram priorizados da seguinte forma:

- **Lembretes** (informação do tipo Educação/Conteúdo): esta locução possui prioridade 1 (um), a maior, pois afeta o esquema de pontuação e aprendizagem. Se o usuário requisitar um lembrete é porque ele não tem certeza do conteúdo. Porém, ao utilizar este recurso, o usuário é penalizado através da perda de pontos, pois significa que não prestou atenção quando o conteúdo foi apresentado a ele. Além disso, este recurso é acionado pelo usuário necessitando assim de um *feedback* rápido da locução;
- **Fatos** (informação do tipo Educação/Conteúdo) e **Curiosidades** (do tipo “Motivação”): estas locuções possuem prioridade 2 (dois) pois, são acionadas pelo usuário (necessitam de um *feedback* rápido da locução), e são informações para o processo de aprendizagem. Estas informações podem ser resgatadas posteriormente e quantas vezes forem necessárias, não gerando punições ao usuário, e;
- **Curiosidades, reforços e dicas de navegação**: esta locução tem prioridade 3, pois é utilizada com o intuito de auxiliar o aprendiz na tarefa de navegação no AV.

Além das locuções ao usuário sobre o elemento selecionado, é executado um som referente a este elemento para proporcionar mais realismo ao ambiente.

6. Plataforma de Teste

Ficou aparente, depois de implementada a infra-estrutura do escalonador, de que havia-se construído, na verdade, uma plataforma de teste muito importante para avaliar os vários “parâmetros” e “estratégias” que podem ser adotados quando da inserção do áudio em um AV.

Os “parâmetros” representam valores numéricos que são associados a informações que conduzem a cadência da produção do áudio no AV; as “estratégias” representam decisões, não influenciadas pelos parâmetros, que governam as escolhas dos recursos auditivos a serem usados. Ou seja, uma vez tendo-se uma resposta tecnológica para o “como” gerenciar, volta-se a questão do “quando”, “onde” e “porque” usar determinados recursos, pois a maioria deles ainda não estão claramente estabelecidos na literatura e portanto, precisam ser devida e criteriosamente avaliados quanto a sua eficiência e/ou influência no PEA.

Os parâmetros que podem ser usados no gerenciamento do áudio incluem:

- Tempo entre as locuções: qual o tempo ideal entre as locuções a serem executadas?
- Tempo de cada locução: qual o tempo, a velocidade de narração que deveria ser executado para cada tipo de informação (no caso, educação, motivação e navegação)?
- Número de vozes diferentes: qual a influência de usar vários tipos de vozes na narração ou usar só uma?
- Tempo de resposta (*feedback*): qual o tempo de latência aceitável para não influenciar no sentimento de presença/imersão?

As estratégias que podem ser usados no gerenciamento do áudio incluem:

- Variação da voz conforme o tipo de locução: qual a influência de pausas, timbres e variações de volume para cada tipo de informação?
- Único áudio (locução ou som) para indicações de erro: deve-se usar o mesmo tipo de áudio para indicar erros (deve-se escolher entre quebra da monotonia ou uniformidade)?
- Quantidade de informações de cada tipo a ser reproduzido: qual a quantidade de informações (do conteúdo, navegação e controles) que devem ser reproduzidas e qual o número de repetições aceitáveis para cada tipo de informação?
- Linguagem da locução diferente da escrita: verificar se a informação transmitida na locução deve ser diferente da informação visual (complementar) ou igual (reforço). Ainda, vale a pena investigar o papel que a entonação (emoção) tem numa locução. Estes dois aspectos podem e devem ser bastante explorados pois, por si só, podem fazer grande diferença entre uma aplicação com, ou sem, áudio.

6.1. O Escalonador Experimentado

Abaixo, seguem as escolhas adotadas na implementação do AV Sherlock Dengue:

- Nenhuma locução possui duração maior que 60 segundos;
- Tempo entre as locuções é de 15 segundos;
- As locuções de dicas de navegação são transmitidas em um intervalo de 4 minutos;

- Uma única voz (feminina) foi usada para veicular as locuções;
- Um único tipo de som foi usado para a indicação de erros;
- Foram transmitidas as mesmas informações tanto na forma visual quanto na forma auditiva para os seguintes elementos: fato, lembrete e curiosidade e *feedbacks*;

As escolhas destes aspectos, identificados neste trabalho, podem ser posteriormente melhor investigadas quanto a sua influência no processo ensino-aprendizagem uma vez que agora se dispõe de uma infra-estrutura de *software* para sua implantação.

7. Conclusões

Este artigo mostrou que apesar dos sistemas de RV serem aplicações ditas multi-sensoriais, o sentido auditivo, considerado importante principalmente no contexto de aplicações educacionais, tem sido pouco utilizado. A revisão da literatura mostrou que existem dificuldades na utilização do sentido auditivo, principalmente sob o ponto de vista conceitual e foi até identificada certa divergência (sobre a veiculação simultânea ou não a outras fontes de informações – como textos, por exemplo). Esta dificuldade é expressa pela carência de pesquisas aprofundadas sobre o assunto e pela falta de um modelo de auxílio a projeto que seja maduro e amplamente aceito.

A solução apresentada aqui foi um Modelo de Informações para Projeto (MIP) auditivo que contempla restrições gerais, a identificação e tipificação das entidades envolvidas (áudios) sob vários pontos de vista e, as formas que podem ser usadas para efetivamente promover a inserção do áudio na dinâmica de um AV. O projeto exemplo suscitou a necessidade de implementação de uma infra-estrutura específica para gerenciar a reprodução dos áudios, que, por sua vez, fez aparecer novos detalhes de projeto (mais específicos e de baixo nível) que foram então devidamente explicitados e levou à implementação de uma infra-estrutura de apoio denominada “escalador de áudio”.

Este trabalho permitiu um melhor entendimento da questão da inserção do áudio em ambientes virtuais 3D interativos e suscitou outros questionamentos tanto do ponto de vista de sua eficácia, quanto de cognição, estética e prática que precisam ser futuramente melhor explorados.

Referências

- Apaydin, O. (2002) “Networked Humanoid Animation Driven By Human Voice Using Extensible 3D (X3D), H-ANIN and Java Speech Open Standards”. Disponível em: <www.stormingmedia.us/39/3971/A397104.html>. Acessado em: 07 abr. 2005.
- Damasceno, E. F. et al. (2004) “Implementação de Serviços de Voz em Ambientes Virtuais”. II Simpósio de Informática do CEFET-PI (INFOCEFET). Disponível em: <<http://www.cefetpi.br/eventos/infocefet/paginas/2004/arquivos/artigos/3.pdf>>. Acessado em: 7 abr. 2005.
- Deol, K.; Sutcliffe, A. e Maiden, N. (1999) “A design advice tool presenting usability guidance for virtual environments”. Workshop On User Centered Design And Implementation of Virtual Environments, UC DIVE’1999. Disponível em

- <http://www.cs.york.ac.uk/hci/kings_manor_workshops/UCDIVE/kaur.pdf>. Acessado em: 28/09/2006. University of York, England.
- Eastgate, R. (2001) "The Structured Development of Virtual Environments: Enhancing Functionality and Interactivity". Tese de Doutorado, University of Nottingham, UK.
- Kalyuga, S. (2000) "When using sound with a text or picture is not beneficial for learning". *Australian Journal of Educational Technology*. 16(2), 161-172. Disponível em: <<http://www.ascilite.org.au/ajet/ajet16/kalyuga.html>>. Acessado em: 29 set. 2004.
- Kniess, J. e Souza, J. L. R. (2004) "Apresentação e Implementação de Um Modelo Para Especificar Sincronização Em Realidade Virtual". IV Congresso Brasileiro de Computação CBCOMP, Itajaí – Santa Catarina.
- LARVA. (2006) "Laboratório de Realidade Virtual Aplicada". Projeto "Sherlock Dengue". UDESC – Universidade do Estado de Santa Catarina. Disponível em: <www.joinville.udesc.br/larva>. Acessado em: 30 Jun. 2006.
- Marshall, E. e Nichols, S. (2004) "Interaction with a desktop virtual environment: a 2D view into a 3D world". Springer-verlag London Ltd. ISSN: 1434-9957. Vol 8. pp. 17 – 25.
- Mayer, R. E. e Sims, V. K. (1994) "For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning". *Journal of Educational Psychology*, 86(3), 389-401.
- Mousavi, S.; Low, R. e Sweller, J. (1995) "Reducing cognitive load by mixing auditory and visual presentation modes". *Journal of Educational Psychology*, 87(2).
- Paivio, A. (1991) "Dual coding theory: Retrospect and current status". *Canadian Journal of Psychology*, 45(3), pp. 255-287.
- Penney, C. G. (1989) "Modality effects and the structure of short-term memory". *Memory & Cognition*, 17(4), 398-422.
- Pinhanez, C. e Intille, S. (2004) "Building Interactive Spaces". Mini-curso. VII Symposium on Virtual Reality. ISBN: 85-904873-1-8.
- Pinho, M. S. (2002) "Som Tridimensional". Disponível em: <<http://www.inf.pucrs.br/~pinho/TCG/Docs/Aula5-Som.doc.pdf>>. Acessado em: 10 ago. 2004.
- Sawhney, N. e Murphy, A. (1996) "Designing Audio Environments - Not Audio Interfaces". ACM/SIGCAPH Conference on Assistive Technologies (ASSETS'96). Disponível em: <<http://citeseer.ist.psu.edu/7742.html>>. Acessado em: 28 abr. 2005.
- Stuart, R. (2001) "The Design of Virtual Environments". ISBN 1-56980-207-6. Barricade Books, Canadá. 274 páginas.
- Suzuki, V. e Hounsell, M. S. (2006) "Um Modelo de Informações para o Projeto Auditivo de Ambientes Virtuais". In: VIII Symposium on Virtual Reality, Belém, PA. SVR 2006. V. 1. p. 1-4.
- Tindall-Ford, S.; Chandler, P. e Sweller, J. (1997) "When Two Sensory Modes are Better than One". *Journal of Experimental Psychology: Applied*, 3(4), 257-287.