

Analyzing the benefits of the combined interaction of head and eye tracking in 3D visualization information

Alexandre A. Freitas, Tiago D. O. Araújo, Paulo R. S. C. Junior

Laboratório de Visualização,
Interação e Sistemas Inteligentes
Universidade Federal do Pará
Belém, Brasil
{alexandre.freytas, tiagodavi70,
paulorobertojr100}@gmail.com

Brunelli P. Miranda, Bruno A. F. Murakami, Carlos G. R. Santos

Laboratório de Visualização,
Interação e Sistemas Inteligentes
Universidade Federal do Pará
Belém, Brasil
{brunelli.miranda,
murakami.bruno,
gustavo.cbcc}@gmail.com

Marcelle P. Mota, Bianchi S. Meiguins

Laboratório de
Visualização, Interação e
Sistemas Inteligentes
Universidade Federal do
Pará Belém, Brasil
{mpmota,bianchi}@ufpa.br

ABSTRACT

This work presents an evaluation of the joint interaction of eye tracking and head tracking in a 3D information visualization environment. In this context, it was conducted a task-based evaluation of the interactions, in a prototype using 3D scatter plots, such as navigation, selection, filters, and other typical interactions of an information visualization tool. The tasks mentioned were performed through interactions using head tracking for navigation and eye tracking for selection, and they were evaluated according to quantitative metrics (time and response of a questionnaire) and qualitative (extracted using the Think-Aloud Protocol). The results show that the "click by blinking" configuration was unstable, but the head tracking as a form of navigation showed a greater accuracy in the interaction.

Author Keywords

Interaction; Head Tracking; Eye Tracking; 3D Visualization Information.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Experimentation; Human Factors; Measurement.

INTRODUÇÃO

Interfaces em três dimensões (3D) estão se tornando cada vez mais comuns no cotidiano dos usuários, sendo usadas em dispositivos móveis, computadores domésticos, displays de larga escala e outras tecnologias que podem oferecer a disponibilidade de interação com este tipo de interface.

O uso de uma nova interface implica também no estudo de novas interações. Algumas características se destacam no uso de interfaces com conteúdo 3D, dentre elas: o grau de

liberdade do usuário e a exploração da interface, permitindo que metáforas acerca de interação possam ser elaboradas, como: "flutuar" e "zoom" [4].

Bowman et al. [3] elaboram boas práticas para construção de interfaces 3D para interações e avaliações na perspectiva do usuário, demonstrando a necessidade de estudos sobre interações para essas interfaces. Seu estudo, incorpora sugestões sobre localização e configuração dos elementos virtuais, para melhor adaptação do usuário à interface 3D.

O uso de interfaces 3D em vários contextos apresenta uma oportunidade para o surgimento de novas tecnologias de interação, já que as tecnologias de interação para ambientes 2D podem apresentar limitações no ambiente 3D. Uma proposta é combinar diferentes tipos de interações, aproveitando o melhor que cada uma pode oferecer. Desta forma, interações multimodais oferecem múltiplos modos de comunicação entre homem-máquina, promovendo uma interação mais comum, intuitiva e próxima ao seu cotidiano [5].

Estas interações não convencionais podem compor tarefas em diferentes interfaces. A voz é uma interação não convencional que tem ganhado influência em vários ambientes. Furtado et al. [8] analisam em seu estudo a voz como interação em uma ferramenta de visualização da informação, e identifica a dificuldade de usar a voz para selecionar um atributo específico.

Os gestos se tornam outra interação não convencional que apresenta dificuldade para tarefas específicas de visualização da informação. Miranda et al. [16] apresentam a dificuldade de selecionar mais de um item na interface. É necessário especificar que a interface usada por Miranda et al. se assemelha ao protótipo utilizado neste trabalho.

A partir desta análise, o rastreamento de cabeça e o rastreamento de olhos são interações que tem ganhado muito destaque em pesquisas atuais. Estas interações não convencionais proporcionam ao usuário a possibilidade de interagir com sistemas em diferentes locais, como *displays* de larga escala [6], onde as interações com *mouse* e teclado

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

são muitas vezes difíceis de integrar, e estes são ambientes comuns para interação com cenários de visualização de dados.

A Visualização da Informação (InfoVis), por sua vez, possui duas grandes etapas: a representação visual de dados abstratos e a interação com esses dados, recursos estes, necessários para conhecer melhor os dados e o relacionamento entre eles, descobrindo nesse processo padrões, *outliers*, tendências, entre outras informações intrínsecas aos dados [21].

Este trabalho objetiva avaliar a experiência do usuário utilizando interações não convencionais em um ambiente de InfoVis 3D. Um conjunto de tarefas foi apresentado para que um grupo de usuários a realizassem em um gráfico de dispersão 3D utilizando a combinação de rastreamento de cabeça e olhos. A experiência mínima esperada dos usuários estava na utilização de pelo menos uma ferramenta de visualização da informação.

No aspecto da análise quantitativa, serão visualizados os tempos totais das tarefas de todos os participantes e as respostas de um questionário em escala *Likert*. Uma análise qualitativa será realizada através do protocolo *Think-aloud*, de forma que os usuários possam falar sobre a experiência da utilização do rastreamento de cabeça e olhos, podendo apresentar informações relevantes sobre a interação com ferramenta utilizada.

Este artigo está organizado da seguinte forma: Trabalhos Relacionados, que mostram a utilização de diferentes técnicas de interações não convencionais em interfaces 3D imersivo ou não imersivo; Contextualização, sobre os conceitos acerca de rastreamento de cabeça, rastreamento de olhos e características do protótipo usado; Metodologia aplicada no trabalho; e Análise dos resultados nos aspectos quantitativos e qualitativos da experiência dos usuários.

TRABALHOS RELACIONADOS

Nesta seção serão apresentadas algumas pesquisas que auxiliaram no desenvolvimento deste trabalho, tanto em conceitos, quanto na construção da metodologia aplicada, demonstrando como as interações não convencionais se destacam quando usadas em um ambiente 3D tanto imersivo quanto não imersivo.

Punpongsanon et al. [18] apresentam uma extensão do trabalho de LazyNav [10], que é um modelo de controle e navegação terrestre *mid-air*, apresentando um sistema que pode ser usado para aplicações em *head-mounted display* (HMD) e *Large Displays* (LD). No contexto de LD o sensor está localizado abaixo do display público. E se tratando de HMD, o sensor está abaixo de uma câmera que realiza o rastreamento da cabeça. No presente trabalho, o contexto de *displays* de larga escala é uma motivação do estudo, pois, dispositivos de interação não convencional possibilitam ao usuário um maior grau de liberdade na interação, destacando a importância de estudos voltados para interfaces que usem um grande espaço para interagir.

Yoo et al. [22] propuseram uma interface para usuário baseado em pontos de fixação e gestos para interação com *displays* de larga escala. O trabalho utiliza uma interação na parede através do display para navegação. O objetivo é combinar interações por pontos de fixação e gestos provindos da mão, de forma que seja possível manipular o conteúdo 3D disponível no display. Foi observado que os usuários aderiram a interface proposta, no entanto é necessário procurar formas para diminuir a fadiga para esse tipo de interação. No estudo, a fadiga é apresentada como uma métrica importante para analisar a experiência do usuário, uma vez que é refletida diretamente na aceitação da interação. Dessa forma, utiliza-se neste trabalho a fadiga como métrica de avaliação do estudo.

Lim e Kim [14] propuseram uma interface voltada para a interação de rastreamento de olhos com conteúdo 3D. Os autores desenvolveram seu próprio algoritmo de reconhecimento dos olhos, que utilizava câmeras RGB conectadas à *Smart TVs* e marcadores em óculos 3D como dispositivos para captura das informações. Para realização das interações, a tela da aplicação era dividida em uma matriz 3x3. No presente trabalho, a abordagem de divisão da tela em 9 áreas, apresentado por Lim e Kim, serviu de inspiração para o design da interface em algumas funcionalidades do protótipo.

Florian et al. [2] analisam as interações com aplicações 3D utilizando rastreamento de olhos, através de experimentos comparando dois métodos para cálculo de pontos de fixação e comparação entre interação por olhar no ambiente 2D e 3D. Neste trabalho, as interações são realizadas em um ambiente 3D, porém tais interações também são comuns ao ambiente 2D, então o estudo de Florian et al. demonstram que é possível que as interações de um ambiente sejam portadas para outro, neste caso do 2D para o 3D.

CONTEXTUALIZAÇÃO

Esta seção apresenta os principais conceitos necessários para a compreensão do trabalho. Abordando o conhecimento de ambas as tecnologias usadas e demonstrando o protótipo de visualização da informação 3D usada na pesquisa.

Rastreamento de olhos

O rastreamento de olhos se caracteriza pelo reconhecimento e monitoramento das atividades oculares, analisando como eles se comportam, o que os atrai ou como se adaptam a algum estímulo aplicado [13].

Muitos estudos abordam novas formas de realizar o rastreamento dos olhos, desde utilizar técnicas que usam extração das características físicas dos olhos, como posição da *glint* (brilho dos olhos) e tamanho da pupila, até análises psicológicas sobre a atenção visual, percepção, fixação do olhar, entre outros [7].

Nesta pesquisa, o rastreamento dos olhos é efetuado com o uso do dispositivo chamado The Eye Tribe [20]. Este dispositivo visa habilitar o rastreamento dos olhos para uso em aplicações, como meio de análise, navegação ou

interação no ambiente computacional. O Eye Tribe possui duas câmeras no centro do dispositivo, e em suas extremidades canhões infravermelhos, usados para obter características específicas dos olhos do usuário no momento da interação (Figura 1).

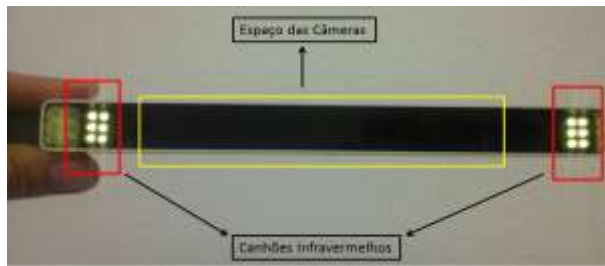


Figura 1. Eye Tribe. Fonte: adaptado de [20]

As câmeras constroem um espaço tridimensional em torno da cabeça do usuário (Figura 2.A), o que permite a movimentação da cabeça enquanto o mesmo fixa o olhar para uma região de interesse. Essa movimentação é importante, pois não é necessário ficar com a cabeça parada para fazer o rastreamento dos olhos.



Figura 2. (A) Espaço tridimensional criado pelo Eye Tribe. Fonte: adaptada de [20]

O Eye Tribe fornece duas interações advindas de seu software de instalação, que são: movimento do ponteiro, e, calibração. A primeira funcionalidade é pouco precisa para interação que necessite estabilidade, justificando assim o uso de outro modo de interação não convencional para cobrir esta falha (rastreamento de cabeça).



Figura 3. (A) Pontos de interesse da calibração. (B) Resultado da calibração. Fonte: adaptada de [20]

A segunda funcionalidade fornece ao usuário uma calibração acerca da observação de pontos específicos, que é feita através do treinamento das características dos olhos relacionada à uma região da tela (Figura 3.A). Essa

calibração é medida através de “estrelas” (Figura 3.B), onde uma estrela refere-se ao pior tipo de calibração, e cinco estrelas é a melhor calibração para uso do dispositivo.

Rastreamento de cabeça

Diferentes áreas fazem uso do movimento da cabeça para diversas funcionalidades, seja no suporte a outras tecnologias de rastreamento ou mesmo como forma de interação em muitos sistemas. Como interação, é possível destacar a naturalidade com que o usuário realiza o movimento da cabeça.

Os movimentos da cabeça se destacam por serem naturais, simples e efetivos para apontar objetos, interagir com sistemas e realizar uma determinada comunicação [1]. Assim, navegar em um ambiente virtual utilizando uma abordagem que detecta o movimento do rosto e/ou da cabeça torna a interação mais simples, intuitiva e discreta [12].

Para o rastreamento de cabeça foi usado o *Enable Viacam* (eViaCam) [15], por ser um *software* que oferece uma versão gratuita e *open-source* (código aberto), sendo necessário uma *webcam* para fazer o rastreamento do movimento da cabeça, além de possuir versões para computadores e dispositivos móveis. Em sua tela inicial (Figura 4), o software apresenta a janela do sistema onde mostra a imagem recebida pela *webcam*, dando a demarcação do rosto do usuário através de um algoritmo de detecção facial, demonstrando as características encontradas dentro da demarcação.

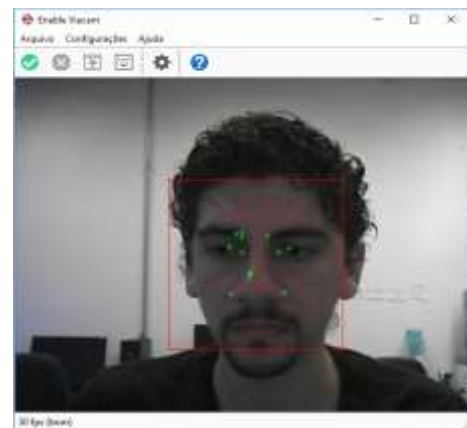


Figura 4. Reconhecimento da face pelo eViaCam

Protótipo de Visualização da Informação

Neste trabalho foi utilizado um protótipo de visualização da informação que apresenta um Gráfico de Dispersão 3D. O protótipo foi modificado a partir do trabalho de Miranda et al. [16] para que o usuário possa interagir através do rastreamento de olhos e cabeça.

Uma base de dados contendo informações simplificadas sobre carros produzidos na década de 80 foi pré-carregada e três atributos distintos da base foram selecionados para configurar os eixos da visualização, sendo: *potência* (eixo x), *rpm* (eixo y) e *nrcilindros* (eixo z). Na ferramenta, são carregados 150 atributos para facilitar a visualização dos dados. A Figura 5 demonstra a tela inicial da aplicação após

a base ser carregada, os respectivos pontos demonstram os dados. As funcionalidades para as tarefas básicas de uma ferramenta de visualização de informação: *Overview*, zoom, filtro e detalhe-sob-demanda [19], são oferecidas pelo protótipo, assim como a navegação, rotação e movimentação da visualização (Figura 5). Segurar e Resetar são interações adicionadas ao protótipo, sendo respectivamente usadas para manter a visualização parada sem interação direta, e para retornar à visualização ao estado original (*Overview*).

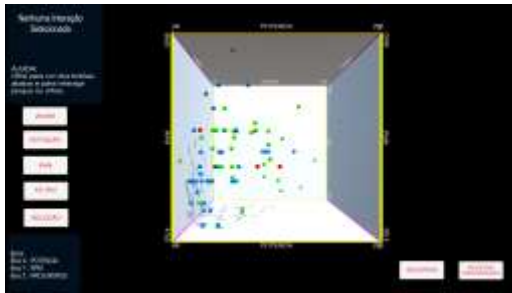


Figura 5. Overview da visualização da informação carregada no protótipo

No momento em que as funcionalidades de ZOOM, PAN ou ROTAÇÃO estão selecionadas, o usuário pode interagir com a visualização movimentando o ponteiro até uma das setas que surgem na frente do cubo, sendo feita assim a ação correspondente à direção desejada. A Figura 6 exemplifica o uso dessa interação com a rotação.

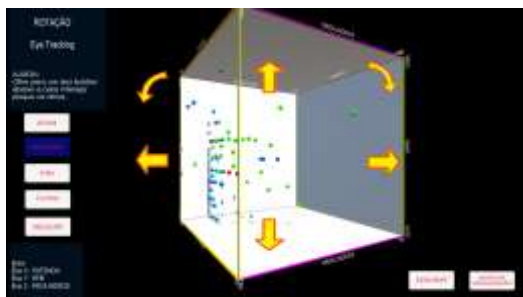


Figura 6. Funcionalidade de Rotação selecionada

Os filtros da aplicação são apresentados através de uma lista de atributos. Um atributo categórico quando selecionado mostra a lista de categorias desse atributo, como na Figura 7, em que se apresenta a seleção de diferentes modelos dos carros da base de dados.



Figura 7. Filtro da aplicação, com o atributo categórico selecionado

O filtro contínuo é ativado com a seleção de um atributo contínuo, e apresenta dois controles deslizantes, como demonstra a Figura 8, além da possibilidade de movimentar a barra para manter a mesma proporção selecionada. Ao selecionar uma das barras laterais a cor rosa é acionada ao usuário dando o *feedback* visual da respectiva ação desejada.

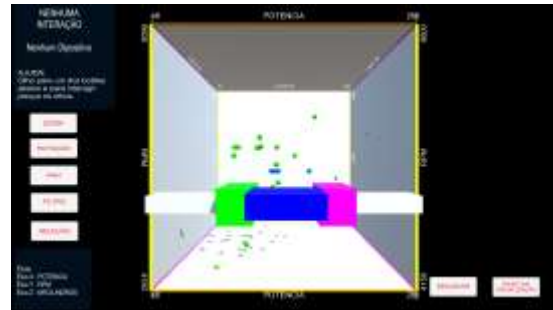


Figura 8. Filtro da aplicação, com o atributo contínuo selecionado

METODOLOGIA

Nesta seção será descrita a metodologia abordada para o desenvolvimento do trabalho, apresentado o cenário de teste, o perfil dos participantes e as métricas usadas na avaliação tanto no aspecto quantitativo quanto no aspecto qualitativo.

No trabalho foram usadas as tecnologias de rastreamento de olhos combinada com rastreamento de cabeça. O movimento da cabeça é capturado pelo software *eViaCam* para efetuar a ação de navegar na interface da visualização, enquanto que o rastreamento dos olhos, através da ação de piscar, fornece a ação de selecionar na aplicação.

No aspecto de implementação, o *eViaCam* fornece em sua interface, o controle de navegação do ponteiro através do movimento da cabeça ao respectivo lado que é desejado movê-lo. Apesar do rastreamento dos olhos efetuar a interação de clique através do piscar de ambos os olhos, o Eye Tribe não fornece esse tipo de interação em sua interface [20]. Por conseguinte, o evento de clique foi implementado a partir da *SDK* fornecida pelos desenvolvedores da ferramenta.

Para os testes foi utilizado um computador *HP EliteDesk* com monitor *EliteDisplay E231* de resolução 1920x1080, com sistema operacional Windows 10 e processador Intel(R) Core i5-4570, com memória RAM de 8GB. Para o dispositivo de rastreamento de olhos apresentado anteriormente, utilizou-se a versão mais atual de seu *SDK*, que é a versão 0.9.77.0. Quanto ao rastreamento de cabeça, a versão usada neste trabalho consiste na 2.1.0.

Procedimento dos testes

Como o objetivo do trabalho é a avaliar quais benefícios da interação por rastreamento de olhos e cabeça combinados no ambiente 3D, uma tarefa foi definida de forma que os usuários possam completá-la através de etapas, e estas

servirão para avaliar a interação e obter o *feedback* dos usuários.

Inicialmente o participante do experimento foi orientado sobre o funcionamento dos dois dispositivos, esclarecendo a interação que cada tecnologia oferece, sendo o rastreamento de cabeça para navegação na interface, e o rastreamento dos olhos para interação de seleção (clique). Posteriormente, a participação do usuário na pesquisa foi condicionada à assinatura do termo de consentimento livre e esclarecido, onde continham informações sobre o supervisor do teste, objetivos e aspectos éticos do estudo. Ao assinar o termo, o participante estava ciente que sua voz seria gravada, transcrita, categorizada e usada para posterior análise.

Após a orientação e tendo o termo de consentimento assinado, o participante deve realizar uma calibração que é oferecida pelo software do Eye Tribe. Esta calibração é necessária, pois os participantes eram orientados a ficar em uma distância confortável do dispositivo de rastreamento de olhos. Dessa forma, o dispositivo deve se ajustar a distância que o novo usuário de encontra, para um rastreamento mais preciso. Não é apenas a distância em relação ao dispositivo que muda entre cada usuário, mas também características importantes para o reconhecimento dos olhos, como a distância entre as pálpebras, o tamanho das pupilas e a distância entre os cantos dos olhos, que são exemplos de atributos que influenciam no reconhecimento do dispositivo.

O usuário teve cerca de três minutos para se habituar ao protótipo de visualização da informação 3D e aos dispositivos utilizados para interação com a mesma. A tarefa do usuário consistia em algumas etapas, que eram:

- Executar um filtro categórico pelo atributo TIPO e selecionar os carros do tipo HATCH;
- Com o subconjunto formado, executar outro filtro, porém agora sendo contínuo, através do atributo PESO;
- Com o novo subconjunto apresentado, identificar qual carro possui maior NÚMEROS DE CILINDROS, usando técnicas de ZOOM, PAN ou ROTAÇÃO;
- Através da técnica de seleção, encontrar o ano de fabricação do carro identificado;

Ao fim do teste, o participante deveria responder um questionário adaptado da ISO 9241-400:2007 [11], presente em um de seus anexos. A ISO citada se refere à avaliação de dispositivos apontadores de forma geral. Este questionário apresenta-se em escala *Likert*, contendo treze perguntas acerca da fadiga do usuário ao usar as tecnologias e da usabilidade de modo geral. A adaptação ocorreu na retirada de três perguntas que não correspondiam ao movimento que seria usado para as interações, estes movimentos eram respectivos aos membros superiores, tendo mesma resposta entre todos os participantes, dessa forma não influenciaram para análise dentro dos resultados obtidos.

Perfil dos Participantes

O estudo foi realizado com 10 participantes, estudantes de graduação e pós-graduação da Universidade Federal do Pará (UFPA) de ambos os sexos. Para participar do estudo apresentado neste artigo, os participantes deveriam ter experiência mínima de uso de uma ferramenta de visualização da informação.

Avaliação

Esta seção é dedicada a apresentar a metodologia relacionada à avaliação da interação que os participantes tiveram ao utilizar os dispositivos de interação não convencional em visualização da informação 3D.

A avaliação dos resultados é feita a partir da análise nos aspectos quantitativos e qualitativos, observando o tempo de conclusão da tarefa e das falas categorizadas através dos principais pontos que os participantes destacaram tanto na usabilidade das interações não-convencionais nas respectivas etapas da tarefa, quanto na interface que a ferramenta oferece ao participante.

Tempo

Esta métrica analisa a relação do tempo em que o usuário realiza a interação dentro do ambiente, explorando a eficiência e os estímulos em que o usuário está envolvido [17]. Nesta pesquisa analisa-se os tempos totais que os usuários obtiveram para concluir a tarefa.

Protocolo Think-Aloud

No aspecto qualitativo, esta pesquisa utilizou o protocolo *Think-Aloud*. Seu estudo se aplica em análises de usabilidade, sendo usado mesmo quando a amostra é obtida por uma quantidade pequena de participantes. Como esclarecido anteriormente, o protocolo consiste no participante falar o que pensa enquanto efetua determinada tarefa [9]. Para este trabalho, a utilização deste protocolo é consistente, por representar a experiência e interação na utilização das interações do rastreamento de cabeça e rastreamento de olhos.

O áudio dos participantes deste trabalho foi gravado, transcrito e analisado posteriormente, extraindo categorias que identificam as qualidades de ambas as tecnologias e juntamente, os problemas de interação. O protótipo também recebeu comentários acerca da organização dos componentes gráficos, gerando assim boas práticas para usar as tecnologias de interação não convencional juntas a uma interface 3D.

Para uma análise mais específica sobre a experiência dos usuários foi gerada uma *word cloud*. Esta visualização consiste em montar uma “nuvem” das palavras que foram transcritas através dos áudios dos participantes. Para montar a visualização, primeiro foram excluídos os conectivos da língua portuguesa e os jargões da fala. Após esse pré-

processamento, através de um site¹ específico para construir este tipo de visualização, foi carregado o arquivo com as palavras transcritas, e gerada a *word cloud* para análise dos resultados.

RESULTADOS

Esta seção demonstra os resultados obtidos através do uso das tecnologias pelos participantes. Como referido anteriormente, os resultados serão apresentados em uma análise quantitativa e qualitativa, apoiando as conclusões e discursões acerca da experiência do usuário.

Resultados Quantitativos

Como especificado anteriormente os resultados quantitativos se concentram na análise do tempo total de duração da tarefa que os participantes obtiveram. A Figura 9 apresenta o tempo que cada participante levou para concluir toda a tarefa do experimento. Destaca-se a variação do tempo entre os participantes, sendo obtida por alguns fatores, como a introdução de uma nova interação com um sistema, diferenciando-se dos métodos tradicionais de interação *mouse* e teclado.

A mesma tarefa foi realizada por alguns dos autores deste trabalho, com intuito de obter a média do tempo total para interação com o *mouse* obteve-se uma média de tempo de 02:00 minutos. Comparando com a média de tempo total da interação da combinação do rastreamento de cabeça e olhos que foi de 06:12 minutos, percebe-se uma diferença proporcional, ocorrida pela configuração do clique e o novo meio de interação proposto.

Os participantes 2 e 4 apresentaram destaque no tempo total da tarefa. O participante 2 apresentou dificuldade com as interações não-convencionais, onde especificou em seu áudio, problemas na ferramenta de visualização e na configuração de clicar com o piscar dos dois olhos.

O participante 4, não se adaptou à interação e ao protótipo apresentado. A dificuldade se concentrou em haver poucos movimentos do corpo. Em seu relato o participante 4 comenta sobre a possibilidade de ficar de pé diante da ferramenta forneceria a ele uma maior liberdade. Este usuário destacou-se por não completar a tarefa, foi um dos participantes que apresentou dificuldades com a interação de selecionar, uma justificativa deste é a calibração mal executada, sendo o único participante que não conseguiu usar o rastreador de olhos sem óculos.

O movimento diante dos rastreadores deve ser limitado, pois ao mover a cabeça demais para ambos os lados o dispositivo de rastreamento de olhos poderia sofrer com a perda de referência do olhar.

A Figura 10 apresenta a qualidade das calibrações de todos os participantes. É importante ressaltar que a calibração é afetada se o participante estiver usando óculos, pois as lentes dos óculos causam refração nos raios infravermelhos do

hardware, dois dos participantes não conseguiram usar o dispositivo sem retirar os óculos, tendo suas calibrações definidas como ruins. Enquanto que os demais participantes da pesquisa tiveram resultados bons de calibração para o uso dos dispositivos propostos.

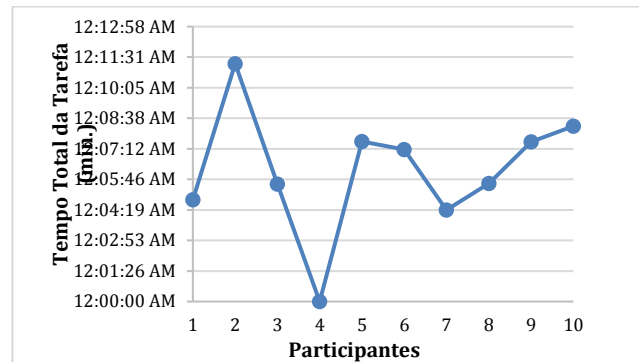


Figura 9. Representação do tempo total da tarefa de cada usuário

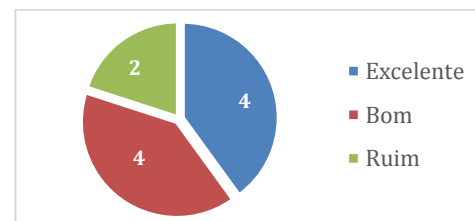


Figura 10. Qualidade das calibrações dos participantes

Na Figura 11 é demonstrada a comparação entre os usuários que usam óculos durante o dia e os que não usam, juntamente com quantos usuários conseguiram usar os dispositivos de rastreamento durante a tarefa com e sem os óculos. Analisando os dados quanto ao uso de óculos, conclui-se que grande parte dos participantes conseguiram utilizar os *hardwares* sem o uso dos óculos (8 participantes), dois participantes não conseguiram realizar a tarefa sem os óculos. Foi questionado ao usuário o grau referente aos óculos em uso, identificou-se que o grau dos óculos e a medida de refração são diretamente proporcionais. A calibragem é prejudicada com o uso dos óculos, sendo difícil a interação de selecionar causando assim cliques indesejáveis. A frustração nos participantes ao utilizar o sistema proposto foi muito evidente, sendo comentado que o clique possui grandes dificuldades na interação.

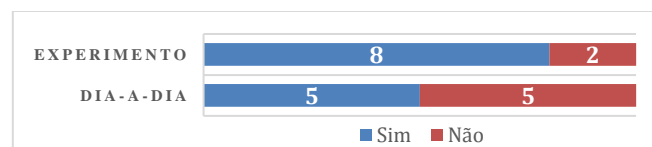


Figura 11. Gráfico quanto ao uso de óculos

O questionário da escala *Likert*, que todos os participantes responderam após a experiência com as tecnologias, é

¹ <https://www.jasondavies.com/wordcloud/>

A pergunta sobre a fadiga nos ombros apresentou escala bastante ruim, sendo colocada como 1, no geral. Os participantes tinham movimento livre no ambiente de teste, e a movimentação dos ombros foi muito elevada, gerando alta fadiga. Apontar com o rastreamento de cabeça se mostrou efetivo, pois os participantes informaram que a interação com o rastreamento de cabeça é precisa e adequada para a ação de navegar, porém a configuração de clique com olhos precisa ser melhor adaptada, sendo necessário mais pesquisas acerca de interações com o rastreamento de olhos. O conforto foi considerado aceitável de acordo com as respostas dos usuários, tendo em média a escala 4 como resposta. Esse conforto pode ser refletido no cenário de testes, onde o usuário possui uma cadeira confortável, uma tela de grande proporção, e movimentação livre.

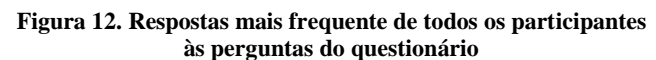


Gráfico de pizza mostrando a distribuição das respostas para a pergunta "Qual a melhor escala?". O gráfico é dividido em quatro segmentos coloridos, cada um rotulado com um número. A legenda indica que os números correspondem às escalas:

- 1 (vermelho): Escala 1
- 2 (laranja): Escala 2
- 3 (amarelo): Escala 3
- 4 (verde): Escala 4

Os segmentos representam as seguintes proporções:

Escala	Proporção
Escala 1	1
Escala 2	2
Escala 3	3
Escala 4	4

Figura 13. Proporção das respostas da pergunta 10 de todos os participantes

A análise no aspecto qualitativo apresenta a categorização das falas de todos os participantes, tanto em relação à interação com os dispositivos não convencionais, quanto à interface do protótipo de visualização da informação 3D.

Figura 14. Word Cloud das palavras mais presentes da transcrição dos áudios

Para análise, também foram consideradas as três palavras com sentido negativo e maior frequência: “travar”, “difícil” e “sensível”. A palavra “travar” está relacionado ao reconhecimento do rastreamento de cabeça, principalmente quando o participante gira a cabeça e a face sai do campo visual de reconhecimento, algumas frases estão dentro desta

classificação, como: “*Ta vendo, ta dando uma travada no head tracker*”. A palavra “difícil” se associa à disposição dos componentes na tela, sendo representada por frases, como: “*Interação no canto da tela está um pouco difícil*”. A palavra “sensível”, está relacionada à ação de clicar, envolvendo o piscar dos olhos, sendo representada por frases do tipo: “*Interação no canto da tela está um pouco difícil*”.

A palavra “óculos” deve ser destacada por apresentar diferentes sentidos na transcrição. Esta é relacionada à interação com os componentes da interface, e principalmente na falha ao reconhecer os olhos, de modo que esta palavra está presente em frases como: “*Clicar com os óculos é difícil*” e “*Não, com os óculos é pior. Ele não reconhece direito*”.

Para melhores entendimentos acerca do objetivo desta análise, foram definidas três categorias (Interação, Interface e Experiência) que definem como os comentários se organizaram dentro de cada etapa definida na tarefa.

Desta forma, a análise será realizada mostrando as subcategorias classificadas dentro de cada categoria que estão relacionadas em cada etapa da tarefa. A Figura 15 apresenta uma linha de tempo que demonstra como estão organizadas cada categoria. Neste ponto, é preciso esclarecer que esta foi uma forma de organização observada pelos autores para representar os melhores *feedbacks* dos participantes sobre a interação, interface e experiência.

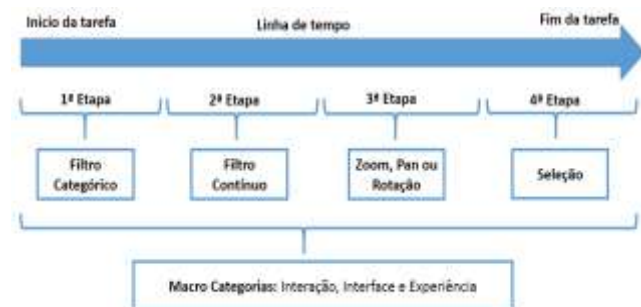


Figura 15. Linha do tempo com as etapas e cada categoria apresentada

Após a definição das categorias, a análise dos áudios foi realizada e extraída subcategorias para cada etapa. Estas análises das categorias fornecem conclusões sobre melhorias e identificação das dificuldades entre os modos de interação usados neste trabalho. A Figura 16 apresenta a hierarquia da primeira etapa, as categorias definidas e as subcategorias extraídas dos áudios dos participantes.

A categoria de INTERAÇÃO obteve mais comentários negativos dos participantes, sendo mais evidentes quanto à categoria “Reconhecimento do Piscar Involuntário”, a interação deveria ser feita no filtro categórico, porém a ação de selecionar (clicar) encontrou problemas na execução chegando a obter dois cliques seguidos ou cliques quando o usuário não havia piscado. Um exemplo de frase que levou à esta categorização é: “*Olha, ele fez uma seleção que eu não cliquei*”.

Quanto à INTERFACE, os comentários foram sobre erros que o protótipo apresentou, tanto de execução quanto de disposição dos componentes, uma das frase que levou à esta classificação de “Erro no Protótipo” foi: “*Quando tu entra num nível, ele não volta!*”. A EXPERIÊNCIA do participante foi um ponto analisado, pois o *feedback* sobre o conforto do participante ao usar as tecnologias em uma interface 3D é fundamental para gerar sugestões sobre a usabilidade.



Figura 16. Categorias da primeira etapa da tarefa

A Figura 17 mostra as categorias da segunda etapa da tarefa, que consistia no participante realizar um filtro contínuo no protótipo. Mesmo sendo a mesma interação da primeira etapa, a categoria que mais se destacou foi sobre a INTERFACE, a principal classificação dos comentários negativos foi “Erro no protótipo”, uma das frases que foram integradas nessa categoria é: “*Ta vendo o componente? Eu não tô vendo o componente!*”. Quanto à INTERAÇÃO, o principal comentário ainda foi em relação ao “Piscar Involuntário”, sendo mais citados erros na interação com os componentes do protótipo. Na EXPERIÊNCIA, surge uma nova subcategoria chamada de “Dúvida”, neste ponto o participante apresentou um tipo de dúvida sobre o que efetuou na interação, e essa dúvida refletiu diretamente em sua experiência gerando um sentimento de frustração, por exemplo. Uma frase usada para mostrar a classificação é: “*Eu não tô entendendo como isso tá acontecendo!*”. A configuração de piscar para selecionar (clicar) mostrou-se instável, havendo cliques consecutivos e cliques involuntários. A navegação com o rastreamento de cabeça obteve boas opiniões dos participantes, sendo preciso com os componentes existentes na interface.

Para amenizar os problemas encontrados nas interações e interface do filtro categórico, é preciso implementar e oferecer sugestões de como interagir com a tecnologia. Neste trabalho o piscar mostrou-se sensível para selecionar, do mesmo modo, a interface demonstrou falhas afetando a experiência do participante. Assim a configuração tanto da

interface quanto da interação deve ser oferecida ao usuário, para que o mesmo se adapte ao sistema proposto.



Figura 17. Categorias da segunda etapa da tarefa

A Figura 18 mostra as categorias referentes à terceira etapa. A categoria de INTERFACE foi o destaque sobre os comentários dos participantes, onde as subcategorias de "Erro do Protótipo", "Forma do Componente" e "Localização do Componente" foram mais relevantes nos resultados, dentre estas o "Erro do protótipo" sendo destaque por não apresentar os componentes na interação, e um exemplo de frase usada para esta classificação é: "Não mudou nada na verdade, não mudou nada".



Figura 18. Categorias da terceira tarefa da tarefa

O modo de interação com rotação, *pan* e *zoom*, mostrou bons resultados, pois manter o ponteiro sobre um componente não afetou na experiência do participante, frase como: "Movimentar está agradável" mostram que o rastreamento de cabeça foi preciso em estacionar sobre um componente.

A Figura 19 destaca a categoria de INTERAÇÃO na quarta etapa. Na subcategoria de "Tipo de interação", a classificação foi obtida por frases como: "Pra isso eu vou usar o que?". Uma explicação para isso se deve ao fato do participante não observar as informações sobre as legendas

na interface. Quanto à INTERFACE, apenas a subcategoria de "Localização de Componente" foi observada. Isto foi obtido por frases como: "Tem alguém isolado?". Esta categoria está presente nos resultados pelos participantes entenderem a mudança na interface de seleção e buscar ler cada característica do atributo apresentado.



Figura 19. Categorias da quarta etapa da tarefa

CONCLUSÃO

Neste trabalho, foi realizada uma análise sobre o uso de interação de rastreamento de cabeça e olhos para um ambiente de visualização da informação 3D. Esta análise trouxe discussões acerca de resultados quantitativos e qualitativos.

A análise tem características exploratórias, onde podem ser elaboradas hipóteses e sugestões para novos estudos. Buscando elaborar novas formas de interagir e organizar a respectiva interface. Os testes foram feitos através de uma tarefa específica disposta aos participantes, esta tarefa foi dividida em etapas que refletiam na respectiva tarefa de visualização da informação.

Os resultados qualitativos foram coletados através do protocolo *Think-Aloud*, foram categorizados de acordo com as etapas da tarefa apresentadas no teste. As categorias apresentadas mostram as opiniões dos participantes acerca da interação, interface e experiência.

Os próximos passos deste trabalho concentram-se em: avaliar as interações em outras interfaces 3D, sendo de visualização de informação ou não; identificar novos paradigmas de interfaces para interações não convencionais e gerar táticas de configuração da interação, como planejar modos de clicar e integrar novas interações como clicar e arrastar.

AGRADECIMENTOS

Os autores agradecem à UFPA por sediar o espaço para o desenvolvimento da pesquisa e à CAPES pelo financiamento das bolsas para os alunos envolvidos. Também agradecemos ao LABVIS-UFPA pelo apoio ao desenvolvimento deste projeto, e a GOL SOFTWARE por sediar o espaço para pesquisa.

REFERÊNCIAS

1. Amer Al-Rahayfeh, Miad Faezipour. 2013. Eye tracking and head movement detection: A state-of-art

- survey. In *IEEE journal of translational engineering in health and medicine*, 1: 2100212-2100212.
2. Florian Alt, Stefan Schneegass, Jonas Auda, Rufat Rzayev and Nora Broy. 2014. Using eye-tracking to support interaction with layered 3D interfaces on stereoscopic displays. In *Proceedings of the 19th international conference on Intelligent User Interfaces*, 267-272. https://doi.org/10.1007/978-3-642-40477-1_13
3. Doug A. Bowman, Ernst Kruijff, Joseph J. LaViola Jr., Ivan Poupyrev. 2004. *3D user interfaces: theory and practice*. Addison Wesley Longman Publishing Co.
4. Doug A. Bowman, Jian Chen, Chadwick A. Wingrave, John Lucas, Andrew Ray, Nicholas F. Polys, Qing Li, Yonca Haciahetoglu, Ji-Sun Kim, Seonho Kim, Robert Boehringer and Tao Ni. 2006. New directions in 3D user interfaces. *IJVR*, 5, 2: 3-14.
5. Bruno Dumas, Denis Lalanne, Sharon Oviatt. 2009. Multimodal Interfaces: A Survey of Principles, Models and Frameworks. In *Human Machine Interaction*, 5440: 3- 26.
6. Giannis Drossis, Dimitris Grammenos, Ilia Adami, Constantine Stephanidis. 2013. 3D visualization and multimodal interaction with temporal information using timelines. In *IFIP Conference on Human-Computer Interaction*, 214-231.
7. Andrew T. Duchowski. 2017. *Eye Tracking Methodology: theory and practice*. Springer International Publishing.
8. Lennon Furtado, Anderson Marques, Nelson Neto, Marcelle Motta, Bianchi Meiguins. 2016. IVOrpheus 2.0-A Proposal for Interaction by Voice Command-Control in Three Dimensional Environments of Information Visualization. In *International Conference on Human Interface and the Management of Information*, 347-360. https://doi.org/10.1007/978-3-319-40349-6_33
9. Amy M. Gill, Blair Nonnecke. 2012. Think Aloud: Effects, Validity. In: *Proceedings of the 30th ACM International Conference on Design of Communication*. New York, NY, USA: ACM. (SIGDOC '12), 31-36. ISBN 978-1-4503-1497-8. <http://doi.acm.org/10.1145/2379057.2379065>
10. Emilie Guy, Parinya Punpongsanon, Daisuke Iwai, Kosuke Sato. 2015. Tamy Boubekur. LazyNav: 3D ground navigation with non-critical body parts. In *3D User Interfaces (3DUI)*, 43-50.
11. International Organization for Standardization. 2007. ISO 9241-400:2007: Ergonomics of human — system interaction — part 400: Principles and requirements for physical input devices. Retrieved May 22, 2017 from <https://www.iso.org/standard/38896.html>
12. Bing S. Kang. 1999. Hands Free Interface to a virtual reality environment using head tracking. U.S. Patent 6,009,210, Filed March 5, 1997, issued December 28, 1999.
13. LC Technology Inc. 2017. What is Eye tracking? Retrieved May 22, 2017 from <http://www.eyegaze.com/what-is-eye-tracking/>
14. Chang J. Lim, Donghan Kim. 2012. Development of gaze tracking interface for controlling 3D contents. *Sensors and Actuators: A. Physical*. 185: 151-159.
15. Cesar M. Loba. 2008. Enable Viacam. Retrieved May 22, 2017 from <http://eviacam.sourceforge.net>
16. Brunelli Miranda, Carlos Santos, Nikolas Carneiro, Tiago Araújo, Anderson Marques, Marcelle Mota, Nelson Neto, Bianchi Meiguins. 2016. Evaluation of Information Visualization Interaction Techniques Using Gestures and Widgets in 3D Environments. In *International Conference on Virtual, Augmented and Mixed Reality*. Springer International Publishing, 9740: 255-265. https://doi.org/10.1007/978-3-319-39907-2_24
17. Kumiyo Nakakoji, Atsuko Takashima, Yasuhiro Yamamoto. 2001. Cognitive Effects of Animated Visualization in Exploratory Visual Data Analysis. In *Proceedings of Fifth International Conference on Information Visualization*, 77-84. <https://doi.org/10.1109/IV.2001.942042>
18. Parinya Punpongsanon, Emilie Guy, Daisuke Iwai, Kosuke Sato, Tamy Boubekur. 2016. Extended LazyNav: Virtual 3D Ground Navigation for Large Displays and Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics*, 23, 8: 1952-1963.
19. Ben Shneiderman. 1996. The eyes have it: a task by data type taxonomy for informations visualizations. In *Proceedings of IEEE Symposium on Visual Languages*. 336-343. <https://doi.org/10.1109/VL.1996.545307>
20. The Eye Tribe. 2017. Our big mission. Retrieved May 22, 2017 from <https://s3.eu-central-1.amazonaws.com/theeyetribe.com/theeyetribe.com/index.html>
21. Ji S. Yi, Youn ah Kang, John Stasko, Julie Jacko (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13, 6: 1224-1231.
22. ByungIn Yoo, Jae-Joon Han, Changkyu Choi, Kwonju Yi, Sungjoo Suh, Dusik Park, Changyeong Kim. 2010. 3D user interface combining gaze and hand gestures for large-scale display. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, 3709-3714. <https://doi.org/10.1145/1753846.1754043>