

Influence of the Camera Viewpoint on Augmented Reality Interaction

G. M. N. Taira, A. C. Sementille and S. R. R. Sanches

Abstract— Many Augmented Reality (AR) applications run on conventional platforms such as desktops or laptops. Desktop applications usually display the 3D environment through a 2D screen while an external webcam attached to the computer performs the capture of the scene. In these applications, the camera viewpoint is an important factor since it can facilitate or hinder accuracy in tasks performed. The objective of this research is to find the optimal camera viewpoint for a desktop AR environment.

Keywords— Augmented Reality, User Perception, Camera Viewpoint.

I. INTRODUÇÃO

OS SISTEMAS de Realidade Aumentada (RA) vêm se tornando mais populares nos últimos anos. Sistemas desse tipo são caracterizados pela coexistência de três características essenciais: combinação de algo real e virtual; interação em tempo real; e alinhamento e sincronização precisos de objetos virtuais tridimensionais com o ambiente real, conhecido como problema do registro [1].

As aplicações de RA mais comuns utilizam marcadores fiduciais para alinhar os elementos virtuais no ambiente real [2]. Esses marcadores são normalmente cartões impressos que possuem uma moldura retangular e um desenho (ou padrão de textura) assimétrico em seu interior. Quando o sistema de RA identifica regiões na imagem que se assemelham com um marcador, essas regiões são comparadas com os modelos de marcadores previamente inseridos no sistema. Havendo sucesso na comparação, o elemento passa a ser tratado como um marcador.

O fato de as dimensões do marcador estarem armazenadas no sistema, assim como a orientação correta do padrão em seu interior, permite ao software calcular a posição e a orientação da câmera de vídeo em relação a esse marcador físico. Essas informações são utilizadas para a determinação das coordenadas relativas à projeção dos objetos virtuais na tela, a qual é feita, quadro a quadro, precisamente sobre o marcador real [4]. Desse modo, quando o marcador fiducial é movimentado, o movimento é reproduzido pelo elemento virtual a ele associado.

O desenvolvimento de sistemas de RA que podem ser executados em dispositivos móveis intensificou-se nos últimos anos [3]. No entanto, as aplicações desenvolvidas para plataformas como laptops e desktops [4] ainda são bastante comuns, uma vez que tais dispositivos são mais poderosos

computacionalmente, quando comparados aos dispositivos móveis. Além disso, a tela em que a cena da aplicação é exibida usualmente é maior quando laptops ou desktops são utilizados como plataforma, o que os torna mais adequados para determinadas aplicações.

Na maioria dos laptops, a câmera que captura os quadros de vídeo encontra-se embutida na parte superior da tela do próprio dispositivo e, por esse motivo, a posição e orientação dessa câmera não podem ser alterados, exceto quando a tampa da tela é reposicionada. Por outro lado, quando a aplicação de RA é executada em um desktop, a imagem real é normalmente capturada através de uma câmera externa, conectada ao dispositivo por meio de uma entrada USB (*Universal Serial Bus*). Nesse caso, tanto a posição quanto a orientação dessa câmera podem ser modificadas pelo usuário. Esse ajuste certamente influencia a qualidade da interação com o sistema, uma vez que o usuário deve planejar seus movimentos para que sejam reproduzidos adequadamente no ambiente da aplicação. O mau posicionamento da câmera pode dificultar o senso de orientação do usuário e, como consequência, prejudicar seu desempenho no ambiente da aplicação de RA.

Inserido neste contexto, o presente trabalho investiga como os efeitos de diferentes orientações da câmera influenciam na qualidade da interação do usuário com o ambiente de RA. Por meio de experimentos subjetivos, foi identificado o melhor ponto de vista para a câmera, considerando, inclusive, as limitações ergonômicas impostas pelos equipamentos.

II. TRABALHOS RELACIONADOS

Muitas pesquisas sobre fatores relacionados à percepção, que podem influenciar o desempenho de usuários em aplicações de RA, podem ser encontradas na literatura.

No trabalho de Knorlei et al. [5], por exemplo, os estudos tiveram como objetivo entender como o atraso na sintetização da cena pode degradar a qualidade da interação. Os autores realizaram experimentos subjetivos para identificar a tolerância dos usuários em relação a esses atrasos. Os resultados demonstraram que tais atrasos podem alterar a percepção quanto às características dos objetos virtuais dentro do ambiente de RA.

Em Dey et al. [6], os autores investigaram como o uso de diferentes tipos de dispositivos de visualização pode influenciar na percepção da profundidade correta dos elementos que compõem a cena de RA. Tablets e smartphones foram utilizados no estudo, que também se baseou em avaliações subjetivas para levantar o nível de satisfação de grupos de usuários associado ao uso de cada tipo de dispositivo.

Outro estudo voltado à percepção, mas quando a aplicação utiliza imagens estéreo, foi realizado no trabalho de

G. M. N. Taira, Universidade Tecnológica Federal do Paraná, Cornélio Procopio, Paraná, Brasil, taira@alunos.utfpr.edu.br

A. C. Sementille, Universidade Estadual Paulista Júlio de Mesquita Filho, Bauru, São Paulo, Brasil, semente@fc.unesp.br

S. R. R. Sanches, Universidade Tecnológica Federal do Paraná, Cornélio Procopio, Paraná, Brasil, silviosanches@utfpr.edu.br

Livingston et al. [7]. Na pesquisa, os autores mostraram que a precisão na percepção da localização dos elementos no mundo virtual depende, entre outros fatores, do tamanho da tela e do tipo de estímulo sob o qual o usuário é submetido pela aplicação.

A tolerância do usuário em relação ao atraso na apresentação dos elementos de uma cena de RA foi analisada no trabalho de Ganapathy et al. [8]. Os autores constataram que a tolerância para esse requisito pode variar, dependendo da dinâmica da movimentação e dos tipos de elementos que compõem a cena.

O atraso na geração dos elementos virtuais que compõem a cena de RA também foi objeto de estudo no trabalho apresentado por Lee et al. [9]. Os resultados mostraram que existe um nível de atraso aceitável a partir do qual o desempenho dos usuários na realização de tarefas pode ser prejudicado. Segundo os autores, apenas o registro visual não é suficiente para a percepção correta da localização dos elementos na cena, uma vez que o registro temporal (atraso) não pode ser desprezado.

No trabalho de Perritaz et al. [10], os pesquisadores analisaram vários aspectos da percepção do usuário em sistemas de RA. Fatores como o atraso na exibição da cena toda (não apenas dos elementos virtuais da cena), a taxa de quadros e a resolução da imagem (tela do dispositivo) mostraram-se impactantes na experiência com a aplicação. Os resultados do estudo foram utilizados na elaboração de um esquema que garante, por meio de ajustes em tempo real, valores aceitáveis para esses fatores, de forma a maximizar a qualidade da experiência com as aplicações de RA.

III. EXPERIMENTOS SUBJETIVOS

A tarefa fundamental realizada nesta pesquisa consistiu na definição e na aplicação de experimentos subjetivos que, por meio da simulação de diferentes orientações da câmera, permitiram identificar a mais adequada. Os detalhes do protótipo desenvolvido, os critérios usados no recrutamento de voluntários e a aplicação dos experimentos são descritos nesta seção.

A. Desenvolvimento do Protótipo

O protótipo desenvolvido para o experimento consistiu em uma aplicação que simulou um ambiente de RA baseado em marcadores fiduciais. Nesse caso, um único marcador foi configurado para ser reconhecido pelo sistema e uma esfera virtual texturizada foi associada a ele.

A cada quadro capturado pela câmera e processado pelo sistema, a esfera era gerada sobreposta ao marcador, de forma que, quando o usuário o movimentava, o mesmo movimento era reproduzido pela esfera virtual na cena aumentada.

Durante a execução do protótipo, são exibidos na tela cinco cubos virtuais da cor branca. Os cubos são inseridos no ambiente da aplicação de modo que pareçam estar posicionados sobre a mesa onde está sendo realizado o experimento. Esses cubos brancos não são associados com nenhum marcador, uma vez que permanecem na mesma posição durante todo tempo de execução da aplicação. A Fig. 1

mostra o ambiente do experimento, destacando o marcador fiducial utilizado, a esfera texturizada associada ao marcador e os cubos brancos inseridos na cena.

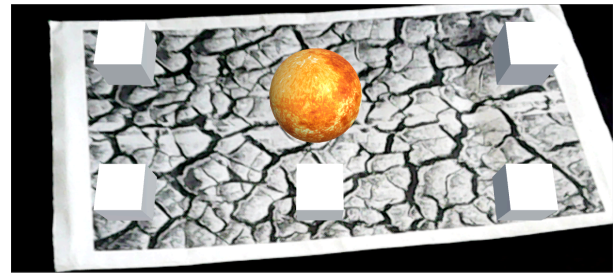


Figura 1. Ambiente da aplicação de RA: marcador fiducial e objetos virtuais utilizados nos experimentos.

O objetivo do usuário na aplicação consiste em fazer colidir a esfera com os cubos dispostos no ambiente, por meio da interação com o marcador. A mesma aplicação, que foi desenvolvida com o auxílio do motor de jogo Unity3D [11] e da biblioteca Vuforia [12], foi utilizada pelos participantes em todos os experimentos. O protótipo foi instalado em um computador equipado com uma câmera convencional. Um suporte na cor preta foi colocado sob o marcador, de forma a padronizar o ambiente visualizado na tela pelos usuários em todos os testes, além de facilitar a movimentação do marcador durante a realização do experimento.

B. Seleção dos Voluntários

O grupo de voluntários que participou do experimento foi composto de 40 pessoas, sendo 10 por bateria de testes. Uma característica importante do perfil de usuário escolhido é o fato de nenhum deles possuir afinidade com a tecnologia de RA. Esta escolha se justifica porque uma pessoa com tal conhecimento poderia utilizá-lo para melhorar seu desempenho na aplicação.

Além disso, foram selecionadas apenas pessoas destras para participar dos testes, uma vez que posicionar a câmera de maneira a facilitar a interação também de pessoas canhotas demandaria grande número de voluntários. Todos os experimentos foram aplicados na instituição de ensino em que o trabalho foi conduzido e os voluntários recrutados eram alunos, professores, funcionários e pessoas externas que se dispuseram a se locomover até a instituição.

C. Preparação e Aplicação dos Experimentos Subjetivos

Uma vez que a mesma aplicação seria utilizada por vários voluntários, um aspecto importante foi a manutenção das mesmas características do ambiente de testes em todas as avaliações. Detalhes como a luminosidade do ambiente, a luminosidade da tela e a distância do usuário em relação à tela seguiram as recomendações da *International Telecommunication Union (ITU)* [13], a qual fornece diretrizes para aplicação de experimentos subjetivos voltados para análise de qualidade de vídeo, inclusive em aplicações multimídia.

Diferentes configurações de ajuste de câmera foram

utilizadas nos testes para que fosse possível identificar o mais adequado. A câmera foi posicionada no centro da tela, presa na parte superior do monitor. Esse normalmente é o local em que tais dispositivos são habitualmente posicionados. Em todos os experimentos, a tela manteve-se perpendicular à mesa em que o computador se apoiou. Todas as configurações consideradas são exibidas na Tab. I.

TABELA I – CONFIGURAÇÕES DE ROTAÇÕES DA CÂMERA SOBRE OS ÂNGULOS X E Y UTILIZADAS NOS EXPERIMENTOS.

Baterias	Rotação X	Rotação Y	N° Participantes
1	45°	0°	10
2	35°	0°	10
3	45°	10°	10
4	35°	10°	10
Total			40

Como pode ser observado, foram definidas duas variações de posição da câmera (35° e 45°) no sentido do eixo de coordenadas X (movimento para cima e para baixo). Para cada uma dessas posições foram realizados testes com a câmera rotacionada 10° para direita (rotação sobre o eixo Y), totalizando 4 configurações (cada configuração representa uma bateria de teste).

A Fig. 2 mostra a o ponto de vista do usuário no ambiente do experimento, destacando o movimento de rotação da câmera sobre o eixo Y, o marcador fiducial e a área em que o marcador é movimentado pelo usuário sobre a mesa.

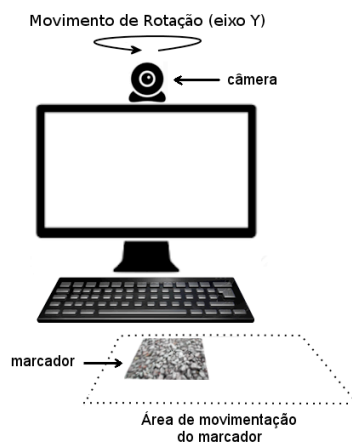


Figura 2. Ponto de vista do usuário no ambiente do experimento. Movimento de rotação da câmera sobre o eixo Y, marcador fiducial e área em que o marcador é movimentado pelo usuário sobre a mesa.

As medidas dos ângulos de rotação da câmera sobre o eixo Y foram definidas considerando questões ergonômicas, de forma a favorecer a tarefa que seria executada pelo usuário, nesse caso, mover o marcador fiducial sobre a mesa. Rotações sobre ângulos menores que 35° fariam com que a câmera apontasse para o teclado e rotações maiores que 45° apontariam a câmera para o próprio usuário, impossibilitando a interação da forma como foi proposta no protótipo.

Cada voluntário participou apenas de uma bateria de teste. Tal medida foi adotada para que um participante não utilizasse a experiência adquirida anteriormente para melhorar o desempenho. Em todos os experimentos, o usuário foi instruído a manter uma distância de aproximadamente 45 cm

da câmera, para que houvesse espaço suficiente na mesa para movimentação confortável do marcador fiducial.

A Fig. 2 mostra a postura do usuário no ambiente do experimento, destacando o movimento de inclinação da câmera (rotação sobre o eixo X) e distância aproximada do usuário em relação à câmera.

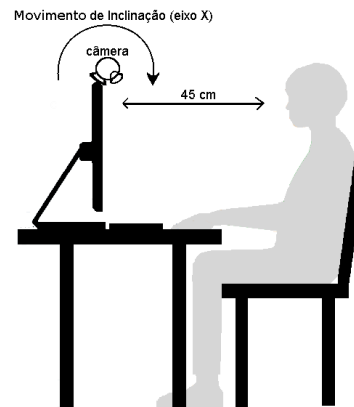


Figura 2. Ambiente do experimento: postura do usuário no ambiente do experimento, indicação do movimento de rotação sobre o eixo X e distância aproximada do usuário em relação à câmera.

Durante a aplicação dos experimentos subjetivos, o usuário necessitava provocar uma colisão da esfera com um dos cubos, o qual adquiria uma cor diferente. A Fig. 3 mostra a cena da aplicação quando a primeira colisão é provocada.

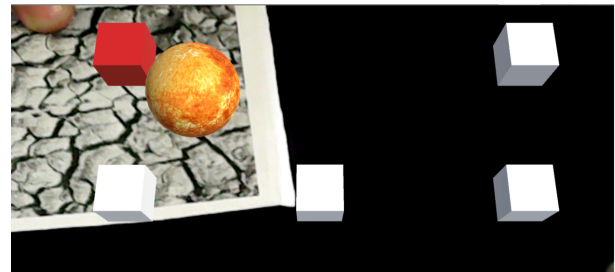


Figura 3. Colisão da esfera com um dos cubos faz com que a cor do cubo se altere.

A tarefa terminava quando o último cubo passasse a exibir uma cor diferente da inicial (Fig. 4), momento em que o tempo gasto para realização da tarefa era registrado.

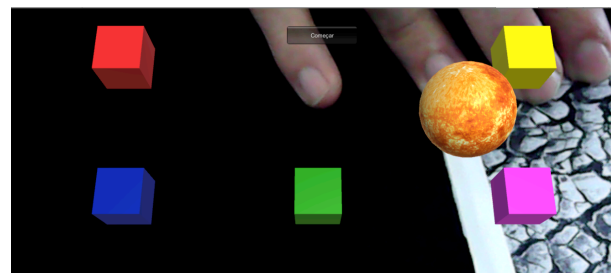


Figura 4. A esfera associada ao marcador colide com o último cubo, finalizando a tarefa do usuário.

Após a utilização do protótipo, cada usuário respondia um questionário com perguntas sobre fatores que prejudicaram seu desempenho na tarefa proposta. Era solicitado, inclusive, que fizesse um breve comentário geral sobre a experiência com a aplicação.

IV. ANÁLISE DOS RESULTADOS

A primeira análise realizada tratou dos dados obtidos das respostas dos voluntários às questões do formulário aplicado antes e após a realização de cada experimento. Nesta análise, inicialmente, procurou-se captar a impressão do usuário sobre a dificuldade da tarefa, antes de sua execução. Em outras palavras, o usuário atribuía um nível de dificuldade à tarefa logo depois de explicado a ele o que deveria ser realizado.

Na primeira bateria, por exemplo, o usuário visualizava o ambiente de teste e era questionado da seguinte forma: “Para a câmera na configuração 1 (rotação de 45° sobre o eixo X e nenhuma rotação sobre o eixo Y), qual o nível de dificuldade esperado para realizar a tarefa?”. Esse tipo de questionamento foi realizado em razão do posicionamento da câmera ser normalmente definido pelos usuários em uma etapa anterior à sua utilização. Desse modo, essa “impressão inicial”, obtida imediatamente após o voluntário posicionar-se no ambiente do experimento e estar ciente da tarefa, torna-se relevante.

Como cada voluntário participou de uma única bateria de testes, a opinião inicial dada referia-se àquela configuração específica. A Fig. 5 mostra os resultados dessa análise.

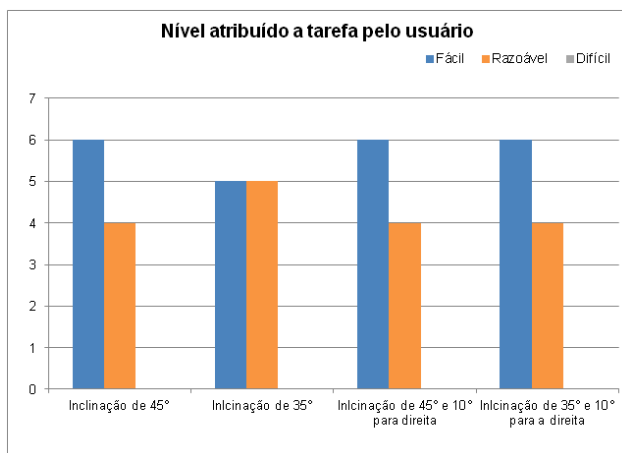


Figura 5. Impressão do usuário sobre a dificuldade de realizar a tarefa antes de executá-la.

A maior parte dos usuários respondeu que não encontraria dificuldade em realizar a tarefa nas configurações da bateria 1, 2 e 3. Por outro lado, a metade dos participantes considerou que o nível de dificuldade da tarefa seria moderado se a configuração da bateria 2 fosse escolhida. Nenhum participante respondeu que, em alguma das configurações, a tarefa seria difícil de ser executada.

A análise posterior procurou confrontar a opinião inicial do usuário com a real dificuldade encontrada por ele, após a realização da tarefa. Nesta análise, os resultados foram obtidos das respostas dadas pelos usuários no questionamento feito após a utilização do protótipo.

Para que não se sentissem tentados a emitirem a mesma opinião inicial, nessa fase, os participantes foram questionados simplesmente se a tarefa que lhe foi imposta era difícil de ser realizada. A questão elaborada para essa fase foi a seguinte: “Houve dificuldade em realizar a tarefa?”.

Novamente, vale ressaltar que cada voluntário participou

de uma única bateria de teste e, por esse motivo, a opinião emitida refere-se àquela única posição da câmera que o voluntário em questão utilizou no momento da execução do protótipo. Os resultados obtidos da análise das respostas referentes a essa etapa do experimento podem ser visualizados na Fig. 6.

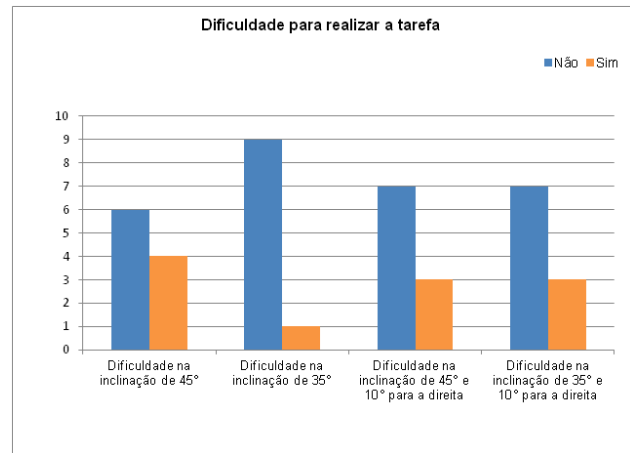


Figura 6. Dificuldade relatada pelos usuários após realizarem a tarefa de colidir os cubos na aplicação utilizada nos experimentos subjetivos.

Ainda que a opinião do usuário sobre a dificuldade em realizar a tarefa seja um dado importante, o tempo gasto pelos participantes para colidir a esfera com os cubos pode mostrar mais objetivamente a dificuldade associada com cada configuração de câmera. Em consequência disso, a análise considerou o cálculo desse tempo e sua associação com determinada configuração.

O tempo para realizar a tarefa no experimento, que é calculado automaticamente pela aplicação, iniciou-se no momento em que o participante clicou sobre um botão exibido na tela. A partir dessa ação, a interação passou a ser realizada apenas por meio da movimentação do marcador fiducial.

O cronômetro, que foi ocultado durante a execução da aplicação, só era interrompido quando a esfera associada ao marcador fiducial colidisse com o último cubo, fazendo com que nenhum cubo branco fosse visualizado na cena. Na Fig. 7 podem ser visualizadas as médias dos tempos gastos pelos voluntários em cada bateria de testes dos experimentos subjetivos.

Os resultados mostraram que quando a câmera encontrava-se sem qualquer rotação no eixo Y e inclinada 45° no eixo X (bateria 1), a média do tempo gasto para a realização da tarefa era menor que nas demais configurações.

Ainda que, para a maioria dos usuários, a rotação da câmera sobre os eixos X, Y (ou em ambos) intuitivamente parecesse não influenciar no desempenho da tarefa (Fig. 5), os dados do questionário aplicado após a experiência com o protótipo mostraram que essa rotação impactou o desempenho. A inclinação 15° maior sobre o eixo X (de 35° para 45°) causou a sensação de aumento da dificuldade na realização da tarefa, como mostram os dados do formulário aplicado após o experimento (Fig. 6).

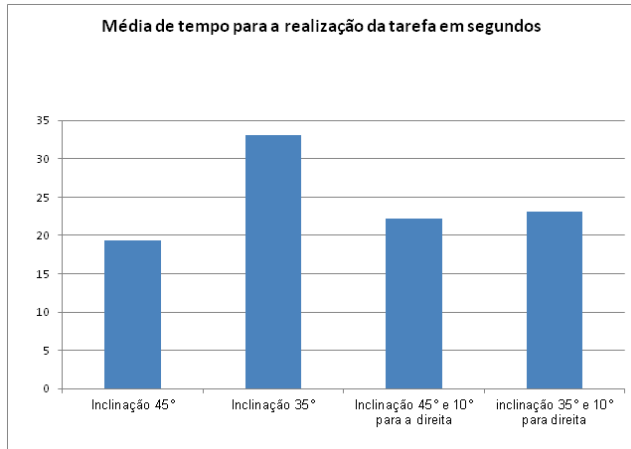


Figura 7. Tempo médio gasto pelos usuários para realizarem a tarefa de colidir os cubos na aplicação utilizada nos experimentos subjetivos.

O gráfico exibido na Fig. 7, relativo ao cálculo do tempo gasto na execução do protótipo, mostra que, uma vez encontrada a melhor inclinação no eixo X (45°), o desempenho é reduzido quando o ângulo é menor no sentido desse mesmo eixo. Da mesma forma, nota-se que rotações sobre o eixo Y podem afetar o desempenho.

O fato de o usuário intuitivamente considerar que as configurações que envolvem rotações sobre o eixo Y possam facilitar (ou pelo menos não comprometer) a interação (Fig. 5) pode estar relacionado com o posicionamento mais confortável do braço, o que facilita a movimentação do marcador. Na prática, ainda que seja mais adequada em termos ergonômicos, tal rotação pode prejudicar a percepção espacial em relação ao ambiente, que é visualizado na tela pelo usuário.

V. CONCLUSÕES

Muitas aplicações de RA, especialmente as baseadas em marcadores fiduciais, são executadas em plataformas desktops. Nessas plataformas, normalmente a câmera que captura a imagem do ambiente real não é embutida no dispositivo, como ocorre nos laptops. Essa câmera é conectada ao computador por meio de algum tipo de interface, por exemplo, a USB.

Como nas aplicações de RA baseadas em marcadores fiduciais normalmente o ambiente da aplicação é visualizado na tela do computador, a orientação da câmera, que é ajustada pelo próprio usuário, pode influenciar no desempenho das tarefas realizadas na aplicação.

Neste trabalho, por meio da aplicação de experimentos subjetivos, identificou-se a orientação adequada da câmera para que a percepção espacial do usuário em relação ambiente tridimensional fosse minimamente afetada. Como consequência, o desempenho na realização das tarefas é favorecido.

Os resultados mostraram que rotações sobre o eixo Y podem comprometer o desempenho de tarefas, ainda que ergonomicamente, esse posicionamento pareça facilitar a interação. Por outro lado, a rotação sobre o eixo X mostrou-se necessária por questões ergonômicas. No entanto, quando o ângulo de rotação não é suficientemente aumentado, verificou-

se que a dificuldade na realização da tarefa proposta na aplicação tornou-se maior.

REFERÊNCIAS

- [1] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 4, pp. 355-385, Ago. 1997.
- [2] E. F. Cardenas and M. S. Dutra, "An Augmented Reality Application to Assist Teleoperation of Underwater Manipulators," *IEEE Latin America Transactions*, vol. 14, no. 2, pp. 863-869, Feb. 2016.
- [3] T. Siu and V. Herskovic, "Mobile augmented reality and context-awareness for firefighters," *IEEE Latin America Transactions*, vol. 12, no. 1, pp. 42-47, Jan. 2014.
- [4] S. R. R. Sanches, D. M. Tokunaga, V. F. Silva, A. C. Sementille and R. Tori, "Mutual occlusion between real and virtual elements in Augmented Reality based on fiducial markers," *Applications of Computer Vision (WACV)*, 2012 IEEE Workshop on, Breckenridge, CO, 2012, pp. 49-54.
- [5] Knörlein, B., Di Luca, M., and Harders, M., "Influence of visual and haptic delays on stiffness perception in augmented reality," *In Science and Technology Proceedings - IEEE 2009 International Symposium on Mixed and Augmented Reality (ISMAR 2009)*, 2009, pp. 49-52.
- [6] A. Dey, G. Jarvis, C. Sandor and G. Reitmayr, "Tablet versus phone: Depth perception in handheld augmented reality," *Mixed and Augmented Reality (ISMAR)*, 2012 IEEE International Symposium on, Atlanta, GA, 2012, pp. 187-196.
- [7] M. A. Livingston, Z. Ai and J. W. Decker, "A user study towards understanding stereo perception in head-worn augmented reality displays," *Mixed and Augmented Reality*, 2009. ISMAR 2009. 8th IEEE International Symposium on, Orlando, FL, 2009, pp. 53-56.
- [8] S. Ganapathy, G. J. Anderson and I. V. Kozintsev, "Empirical evaluation of augmented information presentation on small form factors - navigation assistant scenario," *VR Innovation (ISVRI)*, 2011 IEEE International Symposium on, Singapore, 2011, pp. 75-80.
- [9] C. Lee, S. Bonebrake, D. A. Bowman and T. Höllerer, "The role of latency in the validity of AR simulation," *2010 IEEE Virtual Reality Conference (VR)*, Waltham, MA, 2010, pp. 11-18.
- [10] D. Perritaz, C. Salzmann and D. Gillet, "Quality of experience for adaptation in augmented reality," *Systems, Man and Cybernetics*, 2009. SMC 2009. IEEE International Conference on, San Antonio, TX, 2009, pp. 888-893.
- [11] UNITY TECHNOLOGIES, Unity, <https://unity3d.com>
- [12] PTC Inc, Vuforia, <https://www.vuforia.com>
- [13] ITU, BT.500: *Methodology for the subjective assessment of the quality of television pictures*, <https://www.itu.int/rec/R-REC-BT.500>



medidas de percepção do usuário.

Gabriel Mamoru Nakamura Taira é graduado (2016) em Tecnologia em Análise e Desenvolvimento de Sistemas pela Universidade Tecnológica Federal do Paraná (UTFPR), onde desenvolveu projeto de iniciação científica, entre os anos de 2015 e 2016, no Campus de Cornélio Procópio. Possui experiência na área de Realidade Aumentada com foco em aplicação de experimentos subjetivos para obtenção de



Antonio Carlos Sementille é graduado em Ciência da Computação pela Universidade Estadual Paulista Júlio de Mesquita Filho – UNESP (1988). Recebeu o título de mestre em Ciência da Computação pela Universidade Federal de São Paulo, São Carlos (1994), o título de doutor em Ciências (Física Computacional) pela Universidade de São Paulo (1999) e Livre Docência em Interfaces Avançadas pela UNESP (2010). Atualmente é professor adjunto na UNESP-Bauru. Suas áreas de interesse incluem Realidade Aumentada, Processamento de Imagens Digitais, Televisão Digital e Mídias Interativas.



Silvio Ricardo Rodrigues Sanches é bacharel (2003) e mestre (2007) em Ciência da Computação pelo Centro Universitário Euripides de Marília. Em 2013, recebeu o título de doutor em Engenharia Elétrica (Sistemas Digitais) pela Universidade de São Paulo. Atualmente é professor na Universidade Tecnológica Federal do Paraná (UTFPR), Campus de Cornélio Procópio. Tem experiência nas áreas de Realidade Aumentada e Visão Computacional.