

# Considerações de Usabilidade para Interação por Voz em Aplicações Móveis de Realidade Aumentada

## Usability Considerations for Voice Interaction in a Mobile Augmented Reality Application

Tiago Araújo\*, Brunelli Miranda†, Carlos Santos\*, Nikolas Carneiro\*, Marcelle Mota\*†, Nelson Neto\*†, Bianchi Serique\*†

\*Programa de Pós-Graduação em Ciência da Computação - PPGCC

†Faculdade de Computação – FACOMP

Universidade Federal do Pará, Belém, Pará, 66075-110

**Resumo** – O aumento da popularidade de aplicativos de realidade aumentada móvel tornou possível cobrir novos modos de interação com o seu conteúdo. O uso de linguagem natural torna possível interagir dentro de um aplicativo de uma forma mais intuitiva. O uso da voz, embora não popular entre os pesquisadores está emergindo. Este artigo discute a inclusão de um serviço de reconhecimento de voz e as implicações deste em um aplicativo de realidade aumentada móvel.

**Abstract**—Increased popularity of mobile augmented reality applications made possible to cover new modes of interaction with the its content. The use of natural language makes possible interact within an application in a more intuitive way. The usage of the Voice although not popular among researchers is emerging. This paper discusses the inclusion of a voice recognition service and the implications of this in a mobile augmented reality application.

**Keywords**—realidade aumentada, voz, usabilidade, aplicação móvel, interação, Android, reconhecimento de voz

### I. INTRODUÇÃO

A Realidade Aumentada Móvel (RAM) ganhou um importante destaque em múltiplas áreas de aplicação, dentre elas: engenharias, saúde, educação turismo, automação e outras. Este destaque está relacionado a vários fatores, como: o desenvolvimento tecnológico dos dispositivos móveis, aumento da popularidade e comercialização de *smartphones* e *tablets*, e a inovação e facilidades da Realidade Aumentada (RA).

Apesar do desenvolvimento tecnológico e da crescente popularidade da RAM serem motivações para sua ampla utilização, são os estudos na usabilidade das ferramentas RAM que podem manter uma boa experiência de uso, consolidado assim a longevidade da utilização dessas ferramentas nas mais diversas áreas.

Martínez *et al.* [1] apresenta alguns desafios para o desenvolvimento de aplicações de RAM. Dentre os desafios apresentados, pode-se destacar a escassez de padrões de

desenvolvimento e o pouco espaço para a apresentação de informações.

A escassez de padrões de desenvolvimento pode ser considerada tanto escassez de padrões para codificação dessas aplicações de RAM, assim como para a organização e *design* das interfaces gráficas. E o pouco espaço para apresentação das informações se dá pelo fato de que a RA necessita mostrar ao usuário o ambiente real acrescido de informações virtuais em sincronia [5], e isto, em uma tela de dispositivos móvel como um *smartphone*, que pode variar de 3 a 8 polegadas aproximadamente.

Ao dispor de um espaço limitado os projetistas de aplicações de RAM devem trabalhar arduamente para mostrar ao usuário a cena misturada (ambiente real acrescido de informações virtuais) e ainda apresentar-lhe as possíveis interações e *feedbacks* do sistema.

Inevitavelmente, o projetista terá que optar por deixar algumas informações escondidas (geralmente em menus que podem apresentar várias opções) infringindo alguns princípios de usabilidade, como a primeira heurística de Nielsen [2], em troca de uma melhor visibilidade do conteúdo principal e deixando a tela do usuário com menos elementos gráficos desnecessários. Entretanto, esconder funcionalidades em menus pode acarretar em um maior esforço cognitivo para encontrar e utilizar tais funcionalidades.

Um importante recurso que pode contribuir com a organização da interface gráfica no pouco espaço de tela é a utilização de *feedbacks* e comandos voz na qual o usuário pode interagir e receber informações utilizando a comunicação oral.

Outro ponto importante relacionado a utilização de aplicações móveis é o uso das mãos na aplicação [3]. Aplicações de RAM que necessitam que o usuário aponte sua câmera para uma direção específica, como é o caso das aplicações de localização de pontos de interesse por GPS, podem demandar que o usuário segure seu dispositivo com uma mão fixando o dispositivo em uma direção e tenha que interagir com a outra.

A situação de segurar o dispositivo com uma mão e interagir com a outra pode ser incômoda a medida que o dispositivo é maior e mais pesado, como os *tablets*. As interações por comando de voz podem inibir esse problema, fazendo com que o usuário consiga interagir com a aplicação mesmo que segurando o dispositivo com as duas mãos.

A utilização de comandos de voz em aplicações de RAM ainda é rara e há uma notável carência de estudos da utilização de voz em aplicações de RAM em dispositivos como *smartphones* e *tablets*.

Em uma pesquisa realizada nos principais indexadores de trabalhos acadêmicos da computação (IEEE *Xplore*, *Science Direct* e *ACM Digital Library*) foi constatada uma carência de estudos que contém RAM e comandos e *feedbacks* de voz.

Ao pesquisar com as seguintes palavras chaves: “*Mobile*”, “*Augmented*”, “*Reality*”, “*Voice*” e “*Speech*”, (sendo que *voice* e *speech* foram pesquisados separadamente por serem utilizadas como sinônimos) foram encontrados um total de 42 artigos, sendo 29 na IEEE, 13 na ACM e nenhum na *Science Direct*. Após a leitura dos resumos, remoção de artigos repetidos e remoção de artigos que não estão relacionados com assunto proposto, resultou um total de 12 artigos, sendo 9 na IEEE e 3 na ACM. A Figura 1 (a) mostra o resultado desta busca.

Após uma leitura mais aprofundada dos artigos selecionados, verificou-se que os artigos são sobre aplicações que utilizam comandos de voz, mas não há um estudo mais detalhado sobre os impactos de sua utilização.

A Figura 1 (b) mostra um gráfico dos artigos selecionados por área em que a voz e a RAM foram utilizados. Embora o número de aplicações com voz seja restrito existe uma ampla variedade de aplicações que vai do ensino e aprendizagem até as engenharias e robótica. A Tabela 1 apresenta as citações dos artigos selecionados na pesquisa por área.

Rede, Comunicação	2	[4], [5]
Localização Outdoor/Indoor	2	[6], [7]
Aeroespacial	1	[8]
Automação e Robótica	3	[9], [10], [11]
Aprendizado	3	[12], [13], [14]
Engenharia	1	[15]

TABELA 1. ARTIGOS ENCONTRADOS

Ao realizar a mesma pesquisa, porém sem as palavras *voice* e *speech* foram retornados um total de 1927 resultados, que se comparado aos 42 resultados que incluíram a voz representam bem a ausência de estudos de RAM em conjunto com a voz.

Este trabalho apresenta algumas considerações da utilização de voz em uma aplicação de RAM, e as implicações arquiteturas e interface gráfica provenientes da utilização de voz em uma aplicação de RAM. A aplicação utilizada como estudo de caso chama-se ARGuide e trata-se de uma aplicação de localização e exploração de pontos de interesse por Navegador de RA, Mapa geográfico e leitura de códigos QR Code.

## II. REALIDADE AUMENTADA MÓVEL

A RA trata da inserção de elementos virtuais no cenário real, visando aumentar a percepção do usuário sobre os elementos que o cercam [16]. Com a inovação de dispositivos móveis, que agora possuem processadores e componentes mais avançados, capazes de suportar melhor a renderização de objetos 3D e o processamento de informações, surgiu uma boa oportunidade para o desenvolvimento da RAM.

A RAM tem despertado interesse tanto no desenvolvimento de aplicações na área, quanto para a utilização dessas aplicações. Como mostrado em [17], já existem algumas bibliotecas e Kits de Desenvolvimento de Software (SDK), que oferecem funcionalidades apropriadas à RA e facilitam o desenvolvimento de aplicações RAM.

Na pesquisa realizada em [17], usuários foram convidados a utilizar várias aplicações RAM em diversos contextos, para citar alguns, ambientes *indoor*, *outdoor*, com vários pontos de interesse e em grandes áreas urbanas. Resultados mostraram que há interesse por parte dos usuários em aplicações RAM, mostrando uma aceitação positiva, porém, ainda são necessários estudos em como incluir essa tecnologia em serviços relevantes para o dia a dia do usuário.

## III. INTERAÇÃO POR VOZ

A voz é um dos meios de comunicação mais utilizados pelos seres humanos, estando entre as três formas de comunicação (voz, gestos e expressões faciais) mais utilizadas no dia-a-dia de uma pessoa [18]. A voz, quando utilizada em uma aplicação, está inserida no conceito de linguagem natural.

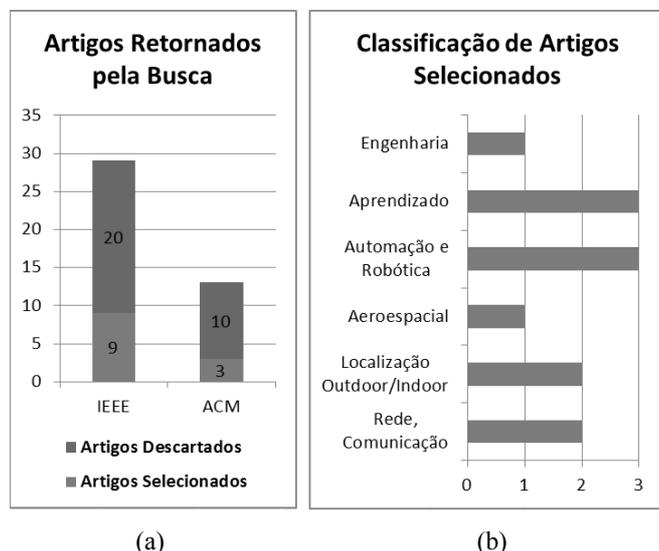


Fig. 1. Informações sobre os artigos pesquisados.

A interação por linguagem natural pode ser definida como a comunicação entre humano e máquina utilizando uma linguagem familiar ao humano [19]. A área de Interação Humano-Computador (IHC) atribui papel importante para esse tipo de interação, devido aos benefícios proporcionados ao usuário, entre eles, o fato das ações serem mais intuitivas, portanto minimizando o esforço cognitivo e permitindo que o usuário se concentre na tarefa e não na interação em si [20].

O uso de voz em *smartphones* e *tablets* também apresenta essas vantagens para o usuário. É cada vez mais comum que as aplicações móveis ofereçam mais funcionalidades ao usuário. Porém, esse aumento de funcionalidades exige que as aplicações usem menus mais complexos, que podem acabar escondendo algumas funcionalidades do usuário ou demandar um tempo considerável até que a funcionalidade desejada seja encontrada. Nesse cenário, a interação por voz pode facilitar a busca do usuário por uma funcionalidade, ou até mesmo servir como atalho para a execução da funcionalidade em questão [21].

Apesar da utilização de comandos por voz poder variar de acordo com o tipo de aplicação, de maneira geral, o reconhecimento de voz nas aplicações segue um modelo padrão, que pode ser observado na Figura 2.

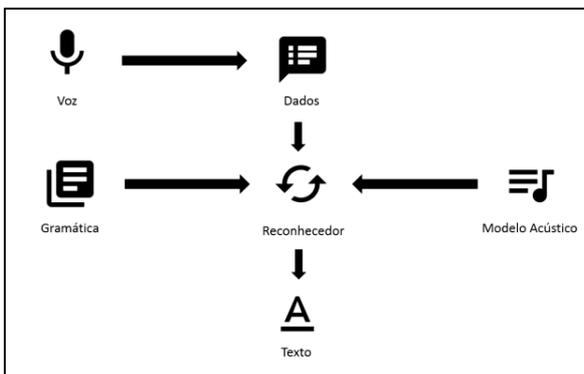


Fig 2. Funcionamento padrão de um modelo de reconhecimento de voz

No modelo apresentado, o usuário entra com a informação em forma de comando de voz, que é repassada para o reconhecedor, onde ela é analisada juntamente com a gramática e modelo acústico empregados na aplicação. O resultado desse processo é então retornado ao usuário em forma textual.

A interação por voz apresenta uso prático em diversos cenários, porém ainda há poucos estudos que a utilizam juntamente com a RAM, portanto, apresenta necessidade de estudos a seu respeito nesse área.

#### IV. ARGUIDE

O ARGuide é uma aplicação de RAM que utiliza conteúdos georeferenciados como base para a navegação do usuário [22]. A aplicação permite que o usuário possa explorar pontos de interesse (*Points of Interest* - POI) alternando entre várias visões coordenadas para acessar o conteúdo disponível. O conteúdo associado a cada POI pode ser multimídia.

As visões podem ser entendidas como diferentes formas de ver os POIs, sendo elas coordenadas significa que as

alterações realizadas em uma visão refletem em todas as outras, dando essa sincronia ao usuário faz com que ele perceba a mesma informação por diferentes perspectivas. As visões disponíveis na aplicação são as seguintes:

- 1) **Mapa:** permite a visualização em um mapa geográfico dos POI da aplicação;
- 2) **Navegador RA:** utiliza a posição do POI para definir marcadores na câmera do aplicativo correspondentes a posição no mundo real;
- 3) **Lista de POI:** Uma visão que organiza os POI em forma de lista;
- 4) **Visualização de POI:** mostra o conteúdo do POI em uma janela, organizando o texto e as mídias disponíveis.

Neste trabalho, o ARGuide recebeu um serviço de reconhecimento por voz para ajudar na navegação e visualização de conteúdo.

#### V. CONSIDERAÇÕES SOBRE A USABILIDADE DA UTILIZAÇÃO DE VOZ NO ARGUIDE

Incluir uma nova forma de interação em uma aplicação de RAM pode acarretar em mudanças no modo em que essa aplicação exibe o conteúdo para o usuário. Lee e Grice [23] identificaram os problemas relacionados à integração de comando de voz em aplicações móveis, e sugerem que esses problemas devem ser tratados desde o início da integração da aplicação com um serviço de reconhecimento de voz.

Para definir quais mudanças devem ser feitas para inserir um serviço de reconhecimento de voz no ARGuide, foi definido um processo para a obtenção de das mudanças que devem ser feitas.

Juntamente com o estudo da literatura, para dar apoio na inserção de interações por voz nessa aplicação, a ajuda de uma especialista em IHC foi necessária. Entrevistas foram realizadas para gerar mudanças no desenvolvimento de forma em que as funções de interações por voz estivessem sendo adicionadas. Foram realizadas duas entrevistas, e entre cada entrevista uma nova versão do ARGuide foi gerada para passar por nova avaliação. A Figura 3 mostra o processo de obtenção das modificações a serem feitas no ARGuide para melhorar a usabilidade quando forem introduzidas interações por voz.



Fig. 3. Processo de inserção de interação por voz no ARGuide

##### A. 1ª Entrevista

Sem componentes funcionais de reconhecimento de voz, essa versão recebeu avaliação baseada em mudanças de interface simples que facilitariam o uso do reconhecedor de voz. A primeira entrevista foi conduzida com explicação prévia da ferramenta e observação da especialista enquanto utilizava a aplicação, mostrando as funcionalidades no decorrer da entrevista.

As considerações feitas na primeira entrevista foram com elementos de interface. A especialista notou que vários

componentes de interface estavam sem indicação de que poderiam ser acionados por um comando de voz. Elementos como os marcadores do mapa ou ícones de navegação tiveram que ser transformados para elementos textuais.

Um exemplo de retorno obtido através da entrevista está na Figura 4. Essa figura mostra dois aspectos na interface que sofreram alteração em função da primeira entrevista. A primeira observação é a troca dos ícones para texto. Os ícones que representavam ações da aplicação foram substituídos por palavras que podem ser usados como comandos de voz. Os marcadores utilizados para identificar os POI foram trocados rótulos de texto.

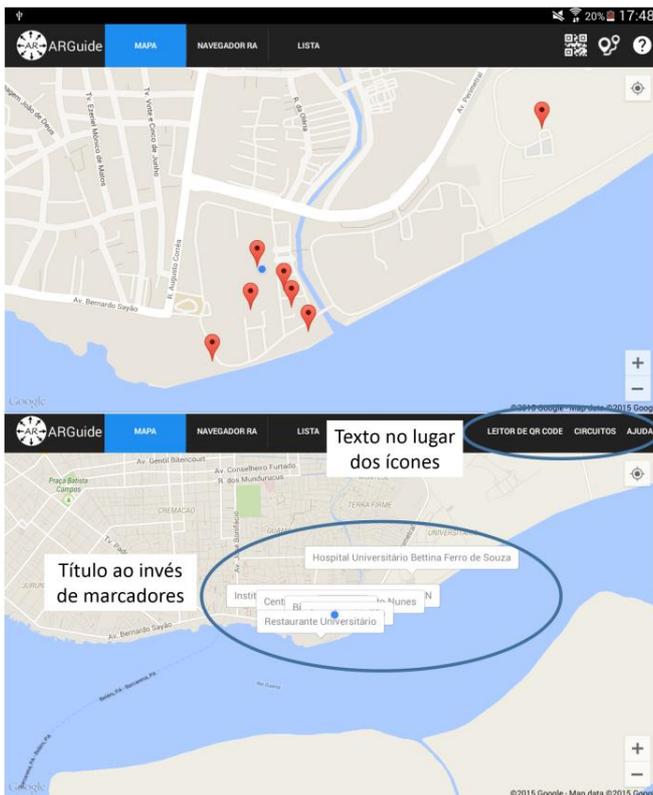


Fig. 4. Alterações realizadas após a primeira entrevista.

### B. 2ª Entrevista

Para a segunda entrevista, foi implementado um serviço de reconhecimento de voz, a implementação utilizou o reconhecedor de voz padrão do Android [24]. É um reconhecedor de voz que funciona online, a recepção e transmissão dos dados é feita no aparelho e enviada para o reconhecedor, que retorna o texto correspondente ao que foi falado. Com o conhecimento da ferramenta, a especialista focou a avaliação no reconhecedor de voz e uso da gramática escolhida.

Novas melhorias puderam ser realizadas com o reconhecedor funcionando. As considerações feitas nessa entrevista foram relacionadas a interpretação e reconhecimento de voz. Comandos base tiveram que ser modificados para acomodar facilidades da linguagem natural,

como relaxar comandos que eram somente verbos no infinitivo para substantivos e conjugações mais comuns. Como “sair”, que pode ser pronunciado “sai” ou “saia” que o reconhecedor entende como o mesmo comando. Outra mudança relevante é o controle de legendas laterais, para que o usuário saiba que comandos pode fazer com voz em determinada visão. A Figura 5 exemplifica uma transição com relação à legenda das ações mapeadas para certa visão.

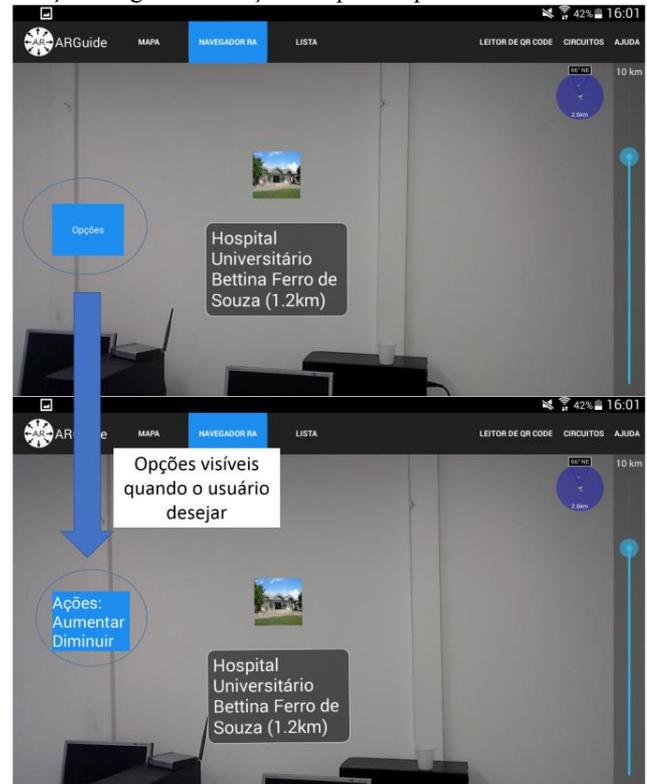


Fig. 5. Alterações realizadas após a segunda entrevista.

O usuário sempre tem disponível uma caixa de texto que abre com um comando de voz para que ele possa visualizar as ações naquela visão. Nesse exemplo, aumentar e diminuir o alcance de pontos de interesse no navegador RA. Em ações muito comuns, como mover uma lista, ou escolher um POI, os próprios marcadores visuais padrão podem indicar esses comandos.

### VI. IMPLICAÇÕES DA INSERÇÃO DE VOZ NA APLICAÇÃO

O serviço de reconhecimento de voz no ARGuide afetou dois pontos principais a aplicação. Um dos pontos já foi discutido, no decorrer das entrevistas algumas mudanças visuais foram necessárias. O outro ponto afeta a arquitetura. A arquitetura do ARGuide é uma variação do modelo MVC, que tem possui uma View com várias visões coordenadas, coordenação essa feita por um Coordinator que intermedia as informações e mídias de determinado POI no Model em uma visão. A Figura 6 mostra a mudança na arquitetura do ARGuide com a adição de comandos de voz.

A inclusão de um serviço de voz é uma interface de coordenação, já que além de acessar os POIs, também

comanda outras visões. Esses comandos são paralelos aos comandos do toque do usuário, então a interação de voz passa a ser coordenada como se fosse o toque.

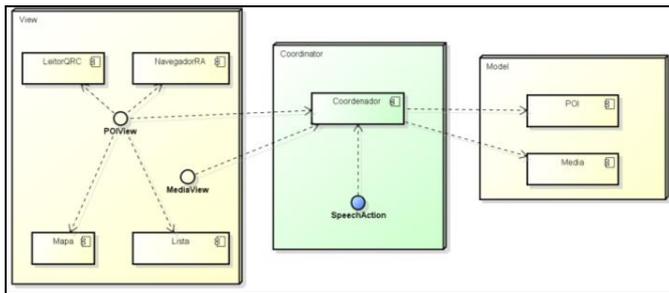


Fig. 6. Mudanças realizadas na arquitetura para acomodar o serviço de voz.

O componente SpeechAction é um conjunto de classes, interfaces e serviços que é responsável pelo conhecimento de voz, ele é o componente novo na arquitetura do ARGuide, e coordena variadas ações do aplicativo.

### VII. SÍNTESE DOS RESULTADOS

As entrevistas realizadas para gerar a implementação das interações de voz com o ARGuide trouxeram um conjunto de considerações que apesar de se aplicarem a uma parte da aplicação podem ser estendidos para toda aplicação.

A Tabela 2 define as principais considerações que foram geradas através desse método. As colunas seguintes mostram em que sentido, se pela Interface Gráfica, ou pela Voz esse procedimento deve ser atendido.

TABELA 2. CONSIDERAÇÕES REALIZADAS PARA INTERFACE E VOZ

Nº	Considerações	Interface Gráfica	Voz
1	Legendas sempre presentes	X	
2	Reconhecedor identificando variações		X
3	Legendas controladas pelo usuário	X	X
4	Ações de configuração sempre acessíveis		X
5	Ações comuns claras para o usuário	X	X
6	Elementos de conteúdo identificados através de texto	X	
7	Ícones devem ser trocados por palavras	X	
8	Ações gerais da aplicação devem ser identificados por palavras já conhecidas pelo público alvo		X

Cada consideração pode ser definida como:

1) **Legenda presente:** A legenda para ações gerais sempre deve estar presente para o usuário.

2) **Reconhecedor:** O reconhecedor deve ter uma gramática que aceita variações, não somente um palavra em cada comando.

3) **Controle de legenda:** O usuário deve controlar a visualização de legendas em cada visão da aplicação.

4) **Ações de configuração:** Para comandos de voz, ações de configuração devem ser sempre acessíveis.

5) **Ações comuns:** Ações comuns da aplicação devem ser facilmente identificados por elementos de interface gráfica.

6) **Identificar conteúdo:** Todo conteúdo disponível por toque deve ter identificação para acessar com voz.

7) **Ícones por palavras:** Substituir ícones por palavras que serão comandos.

8) **Identificar ações gerais:** Ações gerais da aplicação devem ter comandos padrão ou fáceis de memorizar.

A Figura 7 mostra o efeito das considerações em outra parte da aplicação, de maneira mais geral.



Fig. 7. Alterações realizadas após gerar definições gerais.

A troca de ícones que foi feita para as opções foi estendida para a visão de conteúdo do POI. A barra de rolagem sempre presente define a ação de subir ou descer o conteúdo que está sendo mostrado, que serve como exemplo, que mesmo sem legenda é uma ação comum para o usuário mover uma lista para cima ou para baixo, identificada esta ação pela barra de rolagem.

### CONSIDERAÇÕES FINAIS

O uso de voz em aplicações RAM se mostra muito relevante no futuro, já que o uso de linguagem natural para interação com conteúdo é um caminho que está se tornando mais simples com a popularização de aplicações e serviços que utilizam esse tipo de interação.

A contribuição gerada por esse trabalho são as considerações extraídas da implementação do serviço de reconhecimento de voz em uma aplicação de RAM. Essas considerações podem ser estendidas em futuras iterações do

ARGuide, com mais visões ou funções que poderão integrar a aplicação.

A continuidade desse trabalho será feita com testes de usuário e a avaliação dos resultados desse teste para mais uma interação da ferramenta. O teste com outros reconhecedores de voz também está previsto para avaliar a acurácia do reconhecedor na aplicação e como podem afetar e experiência do usuário.

## REFERENCES

- [1] H. Martínez, D. Skournetou, J. Hyppola, S. Laukkanen, and A. Heikkilä, "Drivers and bottlenecks in the adoption of augmented reality applications," *Journal of Multimedia Theory and Application*, 2014.
- [2] NIELSEN, J. "Heuristic Evaluation". Em: NIELSEN, J.; MACK, R. L. *Usability Inspection Methods*. New York, NY: Katherine Schowalter, 1994. Cap. 2
- [3] DE SÁ, M.; CHURCHILL, E. F. "Mobile Augmented Reality: A Design Perspective". Em: HUANG, W.; ALEM, L.; LIVINGSTON, M. A. *Human Factors in Augmented Reality Environments*. New York, NY: Springer, 2013. Cap. 6, p. 139-164.
- [4] Tino Pyssysalo, Tapio Repo, Tuukka Turunen, Teemu Lankila, and Juha Rönning. 2000. "CyPhone—bringing augmented reality to next generation mobile phones." Em *Proceedings of DARE 2000 on Designing augmented reality environments (DARE '00)*. ACM, New York, NY, USA pg. 11-21
- [5] Pulli, P.; Pyssysalo, T.; Metsavainio, J.-P.; Komulainen, O., "CyPhone-experimenting mobile real-time telepresence," in *Real-Time Systems*, 1998. *Proceedings. 10th Euromicro Workshop on*, vol., no., pp.10-17, 17-19 Junho 1998
- [6] Buti Al Delail, Luis Weruaga, and M. Jamal Zemerly. 2012. "CAViAR: Context Aware Visual Indoor Augmented Reality for a University Campus". Em *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology Volume 03 (WI-IAT '12)*, Vol. 3. IEEE Computer Society, Washington, DC, USA, pg. 286-290
- [7] K. G. Srinivasa, Satvik Jagannath, P. S. Akash Nidhi, S. Tejesh, and K. Santhosh. 2014. *Augmented Reality Application: Cloud based Augmented Reality Android application to "know your world better"*. In *Proceedings of the 6th IBM Collaborative Academia Research Exchange Conference (I-CARE) on I-CARE 2014 (I-CARE 2014)*. ACM, New York, NY, USA, , Article 15
- [8] Markov-Vetter, Daniela; Staadt, Oliver, "A pilot study for Augmented Reality supported procedure guidance to operate payload racks on-board the International Space Station," in *Mixed and Augmented Reality (ISMAR)*, 2013 *IEEE International Symposium on*, vol., no., pp.1-6, 1-4 Oct. 2013
- [9] Giesler, B.; Salb, T.; Steinhaus, P.; Dillmann, R., "Using augmented reality to interact with an autonomous mobile platform," in *Robotics and Automation*, 2004. *Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol.1, no., pp.1009-1014 Vol.1, 26 April-1 Maio 2004
- [10] Tang, L.Z.W.; Kian Sin Ang; Amirul, M.; Bin Mohamed Yusoff, M.; Chee Keong Tng; Bin Mohamed Alyas, M.D.; Joo Ghee Lim; Kyaw, P.K.; Foliato, F., "Augmented reality control home (ARCH) for disabled and elderlies," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, 2015 *IEEE Tenth International Conference on*, vol., no., pp.1-2, 7-9 Abril 2015
- [11] Igarashi, T.; Inami, M., "Exploration of Alternative Interaction Techniques for Robotic Systems," in *Computer Graphics and Applications*, IEEE, vol.35, no.3, pp.33-41, Maio-Junho 2015
- [12] Doswell, J.T., "Augmented Learning: Context-Aware Mobile Augmented Reality Architecture for Learning," in *Advanced Learning Technologies*, 2006. *Sixth International Conference on*, vol., no., pp.1182-1183, 5-7 Julho 2006
- [13] Doswell, J.T.; Blake, M.B.; Butcher-Green, J., "Mobile Augmented Reality System Architecture for Ubiquitous e-Learning," in *Wireless, Mobile and Ubiquitous Technology in Education*, 2006. *WMUTE '06. Fourth IEEE International Workshop on*, vol., no., pp.121-123, 16-17 Novembro 2006
- [14] Doswell, J.T., "Context-Aware Mobile Augmented Reality Architecture for Lifelong Learning," in *Advanced Learning Technologies*, 2006. *Sixth International Conference on*, vol., no., pp.372-374, 5-7 Julho 2006
- [15] Goose, S.; Sudarsky, S.; Zhang, X.; Navab, N., "SEAR: towards a mobile and context-sensitive speech-enabled augmented reality," in *Multimedia and Expo*, 2002. *ICME '02. Proceedings. 2002 IEEE International Conference on*, vol.1, no., pp.849-852 vol.1, 2002
- [16] Cawood, S., Fiala, M., "About Augmented Reality", *Augmented Reality: A Pratical Guide*. Pragmatic Bookshelf, USA 2008, cap 1.
- [17] T. Olsson and M. Salo. "Online User Survey on Current Mobile Augmented Reality Applications". *Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality*. Washington: IEE Computer Society. 2011. pp 75-84.
- [18] A. Teixeira et al. "Speech-Centric Multimodal Interaction for Easy-To-Access Online Services – A Personal Life Assistant for the Elderly". *5th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion – DSAI*, 2013.
- [19] B. Shneiderman. "Designing the User Interface: Strategies for Effective Human-Computer Interaction", 3rd Edition, pp 293-295.1998.
- [20] N. Vidakis, M. Syntychakis, G. Triantafyllidis and D. Akoumianakis. "Multimodal Natural User Interaction for Multiple Applications: The Gesture – Voice Example". *International Conference on Telecommunications and Multimedia – TEMU*, 2012.
- [21] L. Xia, K. Kai, W. Xiaochun and W. Dan. *Research and Design of the "Voice-Touch-Vision" Multimodal Integrated Voice Interaction in the Mobile Phone*, 2010.
- [22] Lee, K. B.; Grice, R. A., "The Design and Development of User Interfaces for Voice Application in Mobile Devices," Em *2006 IEEE International Professional Communication Conference*, pp.308-320, 23-25 Outubro 2006
- [23] C. Santos, N. Carneiro, B. Miranda, B. Serique. "Uma Aplicação de Realidade Aumentada Móvel para Ambientes Indoor e Outdoor". *XV Workshop de Realidade Virtual e Aumentada, Marília – SP*, pp 120-126, November 2014.
- [24] Google Inc., "Google Now", <https://www.google.com/landing/now/>. Dezembro 2014.