

GERAÇÃO E SUA VIZÃO DE TRAJETÓRIAS AUTOMÁTICAS PARA UMA SIMULAÇÃO RESIDENCIAL DE UMA CADEIRA DE RODAS COMANDADA POR VOZ

ALEXANDRE L. C. BISSOLI, BRUNO L. S. DA SILVA, LUIS E. M. DE LIMA, RODRIGO V. ANDREÃO.

*Coordenadoria de Engenharia Elétrica, Campus Vitória, Ifes
Av. Vitória, 1729, 29040-780 Vitória, ES, BRASIL
E-mails: xandybissoli.5@hotmail.com, bruno.leg@gmail.com, luisedu@ifes.edu.br,
rodrigova@ifes.edu.br*

Abstract— This article is devoted to the simulation of a robotic wheelchair navigation in a residential environment structured through a map. It is described the trajectory generation algorithm which takes into consideration jerk at the start and at the end position, making the trajectory smooth to the user. The wheelchair is controlled by voice command, which employs a post processing step to improve noise robustness and to reduce command error rate.

Keywords— Speech Recognition, Wheelchair, Intelligent Automation

Resumo— Este artigo trata da simulação da navegação de uma cadeira de rodas em um ambiente residencial estruturado com mapa. É apresentada a geração de trajetórias com análise do jerk na partida e na chegada, tornando a trajetória suave para o usuário. A cadeira é acionada por um sistema de reconhecimento de voz que utiliza uma etapa de pós-processamento para aumento da robustez a ruídos e redução da taxa de comandos errados.

Palavras-chave— Reconhecimento de Voz, Cadeira de Rodas e Automação Inteligente.

1 Introdução

De acordo com o IBGE (Instituto Brasileiro de Geografia e Estatística) a população brasileira atual é de aproximadamente 190 milhões de pessoas (censo 2010), desse total 23,9% possui algum tipo de deficiência e 6,96% (aproximadamente 13,3 milhões de pessoas) possui deficiência motora.

Essa realidade justifica esforços em contribuir para a melhoria da qualidade de vida destas pessoas. Visando essa melhoria faz-se necessário o investimento em pesquisas que melhorem a mobilidade dos cadeirantes (portadores de necessidades motoras especiais).

Neste artigo é apresentada a simulação da navegação de uma cadeira de rodas em um ambiente residencial fictício, com integração de geração de trajetórias a partir de comandos de voz.

Cadeiras de rodas automatizadas podem ser consideradas robôs móveis a rodas, sistema de locomoção do tipo tração-diferencial. Enquanto robôs móveis, um dos desafios atual é dotá-los de autonomia suficiente para o planejamento e execução de trajetórias sem a necessidade de intervenção do usuário, além da definição de origem e destino (posição e orientação, inicial e final).

O Planejamento de Caminhos e a Geração de Trajetórias são duas funções essenciais de um planejador de movimentos para robôs móveis autônomos. A função Planejamento do Caminho realiza o planejamento geométrico do melhor caminho que leva o robô da origem ao destino desejado, podendo ser considerado o mapeamento do ambiente (ambiente estruturado) ou a execução de comportamentos reflexivos para detecção de obstáculos e fuga de colisões. Complementando o planejamento, a função de Gera-

ção da Trajetória define a cinemática para a execução de um determinado caminho, ou seja, as relações entre posição, velocidade e aceleração sob as quais o robô estará submetido em cada instante de tempo para que a execução do caminho seja realizada respeitando a restrição tempo de execução (Cruz, 2011).

Para as trajetórias tratadas neste trabalho são considerados dois tipos de perfis de velocidade, o trapezoidal e o perfil em S. Sendo que os perfis de velocidade em S estabelecem variações de aceleração denominadas “jerk”, que exercem influência sobre o conforto do usuário a ser analisada neste trabalho (Rew, 2010).

Como a interface homem-máquina utilizada neste trabalho é via comandos de voz, processo este definido como a conversão eletrônica de um sinal de áudio em palavras (Anusuya, 2009), vale citar outro desafio que é a obtenção de uma relação sinal-ruído que permita o correto desempenho do sistema. Conforme apontado por Park et al. (2007), apesar da diversidade de sistemas propondo a interface por comando de voz, poucos são aqueles cujo desempenho é satisfatório em ambientes ruidosos. O desafio desses sistemas é oferecer uma baixa taxa de erro, uma vez que comandos errados irão impactar na mobilidade, podendo levar o usuário a situações perigosas.

A integração do comando de voz com a geração de trajetórias suavizadas são tratadas em uma simulação no LabWindows. A simulação é uma etapa que precede a implementação embarcada e que contribui significativamente para o êxito da implementação. Através dela, é possível validar o protocolo de testes, identificar situações não previstas no projeto do sistema, validar a interface homem máquina via

comandos de voz e validar o algoritmo de determinação de trajetórias.

2 Materiais e Métodos

2.1 Reconhecimento de Voz

O processo denominado reconhecimento de voz, realizado por dispositivos eletrônicos, possui diversas aplicações: acionamento de uma cadeira de rodas (Pacnik, 2005); automação residencial e acionamento de dispositivos em geral (Yuksekkaya, 2006); conversor fala-texto.

Uma das principais vantagens deste tipo de tecnologia é a economia de tempo na realização de tarefas do dia-a-dia, além da possibilidade de uma maior independência para indivíduos que necessitam de equipamentos para auxílio em sua locomoção, além de trazer um maior conforto.

2.1.1 Funcionamento de um SRV

O funcionamento de um SRV é ilustrado na Figura 1.

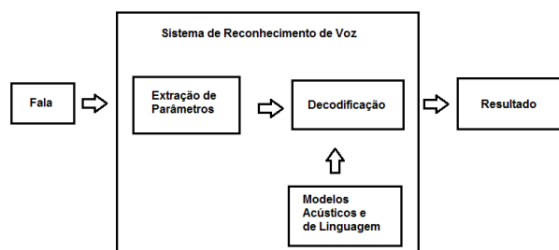


Figura 1. Diagrama de blocos de um SRV

O sinal captado passa por uma etapa de extração de características, onde são retirados os parâmetros necessários para as etapas seguintes. Isto é feito principalmente para reduzir a quantidade de dados que deverão ser processados, e, conseqüentemente, reduzir o custo computacional do sistema.

Tais parâmetros são calculados a partir do processamento do sinal de voz por um conjunto de filtros na escala mel, e são denominados coeficientes mel-cepstrais.

Na etapa de decodificação, é necessária a utilização de modelos acústicos. Tais modelos representam os fonemas do português brasileiro através de Modelos Ocultos de Markov (HMM). Os modelos acústicos são gerados a partir de uma base de dados de voz contendo a transcrição fonética das palavras pronunciadas. Neste trabalho, foram utilizados os modelos acústicos gerados pelo grupo FalaBrasil, da Universidade Federal do Pará (Neto, 2011).

Tais modelos são treinados de modo que não dependam do locutor que o está utilizando. Isto evita a necessidade de uma etapa de adaptação dos modelos ao usuário. Eles levam em consideração diferentes pronúncias de palavras, reconhecendo, assim, as variações da voz de diferentes locutores.

A decodificação consiste então na comparação do vetor de parâmetros extraídos com os modelos

acústicos treinados. Neste ponto, a utilização de modelos de linguagem é importante para a redução de custo computacional. Tais modelos levam em conta o vocabulário da aplicação, fazendo assim com que o sistema direcione suas buscas apenas aos comandos previstos.

2.1.2 Sistema de Reconhecimento de Voz Julius

O Julius é um software de reconhecimento de voz contínua, de código aberto. Ele é desenvolvido no Japão e possui suporte a diversas entradas de áudio, bem como diversas configurações para a etapa de decodificação.

Ele é desenvolvido, principalmente, utilizando a linguagem de programação C, o que torna fácil sua implementação em diferentes sistemas operacionais, como Linux e Windows (Lee, 2010). Por ser um software livre, sua customização é facilitada.

Ele foi configurado neste trabalho no modo de captação em tempo real, com uma taxa de amostragem de 16 kHz, sendo cada amostrada quantizadas em 8 bits.

2.1.3 Interferência de ruídos

O grande problema dos sistemas de reconhecimento de voz é a sua resistência, ou robustez, a ruídos. O ruído pode ser definido como qualquer sinal captado, que não foi destinado ao acionamento do sistema. Como exemplo, temos a respiração do usuário, ou até mesmo o barulho causado pelo acionamento dos motores da cadeira de rodas.

A seguinte tabela mostra como a eficiência do sistema implementado por Yuksekkaya (2006), reduz, ao diminuirmos a relação entre as energias do sinal de voz e a energia do ruído.

Tabela 1. Eficiência de um SRV com ruído

Eficiência do Sistema de Reconhecimento de Voz	
Relação Sinal-Ruído (dB)	Eficiência do Sistema (%)
5	44,3
15	72,8
25	82,3
35	92,1
Acima de 45	98,7

Uma estratégia que influencia no aumento da robustez ao ruído, mas também aumenta a rejeição de comandos, é aquela baseada na utilização de dois modelos de linguagem, chamada de estratégia de fusão de vocabulários (Silva, 2013). Ela faz o sistema levar em consideração a pronúncia correta de cada palavra, para validar o resultado.

Tal estratégia é realizada com a criação de um segundo vocabulário, no qual estão contidas todas as palavras do vocabulário original, adicionadas de palavras com composição fonética parecida com as anteriores. O sistema processará o sinal de áudio

duas vezes, cada uma com um vocabulário diferente, retornando, assim, dois resultados. Quando estes resultados forem semelhantes, o sistema o considera como válido, caso contrário, ele rejeita o comando.

A seguinte tabela nos demonstra o desempenho dessa estratégia, quando comparada com a utilização de apenas um vocabulário, onde a coluna denominada Erro considera os casos em que foi reconhecido um comando diferente do falado e a coluna FN (Falso Negativo) considera os casos onde o sistema rejeita o locutor.

Tabela 2. Comparação da taxa de erros de um SRV para utilização de 1 e 2 vocabulários distintos

Nº de Vocab.	SNR 20 dB		SNR 15dB		SNR 10dB	
	Erro (%)	FN (%)	Erro (%)	FN (%)	Erro (%)	FN (%)
1	1,5	0	4,1	0	9,3	0,74
2	0	15,2	1,5	32,6	1,1	51,9

Através da tabela 2 pode-se ver que este tipo de estratégia consegue reduzir a taxa de erros de reconhecimento do sistema. Em compensação, o mesmo passa a rejeitar mais comandos, devido à necessidade de coincidência dos resultados dos vocabulários.

2.2 Geração de Trajetória

Existem muitas técnicas para a geração de caminhos geométricos a serem percorridos por robôs móveis com rodas. Nessas técnicas parte-se de um ponto inicial, ou origem, (x_i, y_i, θ_i) e chega-se a um ponto final, ou destino, (x_f, y_f, θ_f) (Latombe, 1991). Entretanto, quando o sistema de locomoção destes robôs impõe restrições não-holonômicas (tração-diferencial) aos movimentos a serem executados, muitas destas técnicas não se aplicam, sendo necessárias adequações que viabilizem a geração de caminhos factíveis.

Como este trabalho aborda a locomoção em um ambiente residencial, que normalmente apresenta espaço restrito para manobras, é importante evitar o uso de trajetórias complexas. As trajetórias utilizadas foram compostas por segmentos de reta e rotações que se tornaram viáveis por se tratar de um ambiente estático, além do fato de que trajetórias retilíneas apresentam implementação de baixo custo computacional e conseqüentemente reduzem o tempo de resposta do sistema, características desejáveis para sistemas de controle embarcados.

Embora haja distintos padrões de movimentação, o padrão utilizado neste trabalho é o movimento ponto-a-ponto. No padrão Ponto-a-Ponto a trajetória é geometricamente definida pelas coordenadas dos pontos origem e destino e as respectivas orientações da cadeira nestes pontos, e cinematicamente esta trajetória é definida em três fases que estabelecem um período de aceleração retirando a cadeira do repouso, um período de velocidade constante para translação e um período final de desaceleração com a chegada ao ponto de destino.

2.2.1 Perfis de Velocidade

Os dois perfis de velocidade normalmente utilizados para a caracterização da locomoção ponto-a-ponto são:

- Perfil de Velocidade em S;
- Perfil de Velocidade Trapezoidal.

A figura 2 apresenta o perfil de velocidade trapezoidal e o perfil de aceleração correspondente, ambos divididos em três etapas.

Na primeira etapa o sistema é acelerado, na segunda etapa a velocidade é mantida constante (aceleração nula) e por fim na terceira etapa o sistema é desacelerado.

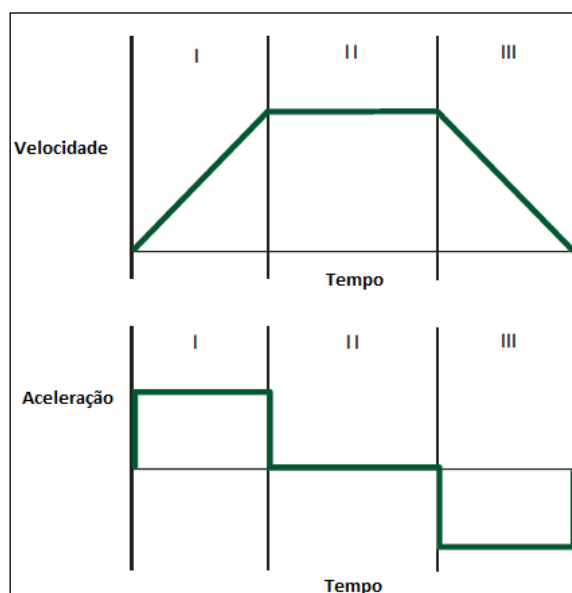


Figura 2. Perfil de velocidade trapezoidal e perfil de aceleração.

A figura 3 apresenta o perfil de velocidade em S e o perfil de aceleração correspondente, sendo que neste caso são divididos em sete etapas.

Na primeira etapa o sistema é acelerado com aceleração crescente (diferente do caso trapezoidal em que o sistema é acelerado com aceleração constante), na segunda a aceleração é constante e na terceira etapa o sistema ainda está sendo acelerado, porém a variação da aceleração (jerk) é negativa.

Essas três primeiras etapas do perfil em S correspondem somente à primeira etapa do caso trapezoidal, a diferença é que nesse caso foram acrescentadas as etapas 1 e 3, onde ocorre uma variação contínua da aceleração, isto é, eliminam-se as descontinuidades do jerk, que estão presentes no caso trapezoidal, tornando o movimento da cadeira mais suave.

A quarta etapa desse perfil corresponde a segunda etapa do perfil em trapézio, onde a aceleração é nula e a velocidade é mantida constante.

As três últimas etapas fazem o papel inverso às três primeiras, na quinta etapa inicia-se a desaceleração da cadeira e com variação negativa do jerk, na sexta etapa a desaceleração continua, mas com jerk

nulo, e por fim na sétima e última etapa a cadeira é desacelerada, mas com jerk positivo.

Nas etapas 5 e 7, onde ocorre uma variação contínua da desaceleração eliminam-se as discontinuidades do jerk, que estão presentes no caso trapezoidal, suavizando os movimentos da cadeira também na frenagem.

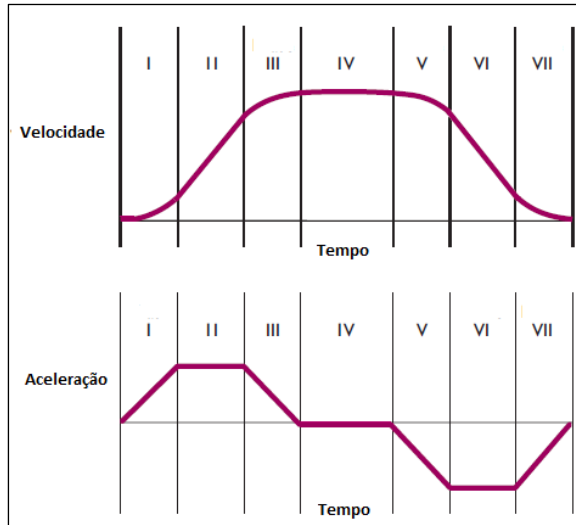


Figura 3. Perfil de Velocidade em S e perfil de aceleração.

A representação algébrica do perfil de velocidade trapezoidal são provenientes da Mecânica Clássica do movimento de corpos (XX), e são apresentadas a seguir:

$$v(t) = \int_{t_0}^t a(t)dt \quad (1)$$

Isto é,

$$v(t) = v(0) + at \quad (2)$$

$$s(t) = \int_{t_0}^t v(t)dt \quad (3)$$

Isto é,

$$s(t) = s(0) + v(0)t + \frac{at^2}{2} \quad (4)$$

Já as equações referentes ao perfil de velocidade em S introduzem a consideração do jerk (Rew, 2010). A seguir é apresentada uma síntese deste equacionamento.

Da definição de jerk, denotado por j:

$$a(t) = \int_{t_0}^t j(t)dt \quad (5)$$

Isto é,

$$a(t) = a(0) + jt \quad (6)$$

Para a velocidade, tem-se:

$$v(t) = \int_{t_0}^t a(t)dt \quad (7)$$

Isto é,

$$v(t) = v(0) + a(0)t + \frac{jt^2}{2} \quad (8)$$

E finalmente para a posição, tem-se:

$$s(t) = s(0) + v(0)t + \frac{at^2}{2} + \frac{jt^3}{6} \quad (9)$$

A utilização do perfil de velocidade em S para caminhos em retas contribui para uma suavização dos movimentos da cadeira proporcionando um maior conforto ao usuário da cadeira de rodas, mesmo que para isso seja necessário um tempo de percurso um pouco maior.

Como foi observado na figura 3, o perfil de velocidade em S é simétrico, porém, atualmente, tem-se observado muitos trabalhos com perfis de velocidade em S assimétricos, com o objetivo de redução do tempo de trajetória e eliminação de vibrações residuais ao final do movimento, sendo assim o jerk é tratado apenas na desaceleração. Vale ressaltar que neste trabalho há interesse em conjugar minimização do tempo de trajetória com conforto para o usuário, o que definiu a escolha pelo perfil em S simétrico.

2.3 Software de Simulação

A simulação foi desenvolvida com o auxílio do software LabWindows CVI, onde são apresentadas as trajetórias executadas pela cadeira em uma interface gráfica representando uma residência.

O reconhecimento de voz, por sua vez, foi implementado através do software Julius. Os comandos de voz são recebidos pelo Julius que, após detectar a validação do comando pelo usuário, os envia ao software principal de simulação, onde são executados.

Para o desenvolvimento do software de simulação, foi construído um mapa de uma residência fictícia na qual a cadeira irá se locomover entre os cômodos. Foram definidos os pontos centrais de cada cômodo, além de pontos auxiliares para a definição dos diversos caminhos. A residência é composta pelos seguintes cômodos: Sala, Varanda, Cozinha, Área de Serviço, Quarto, Corredor, Dormitório e Banheiro, sendo cada um destes, representado por um ponto no mapa. Também foi definido o ponto de entrada da casa.

A cadeira irá se mover entre o seu ponto atual, e o ponto de destino definido com uso de comando de voz e confirmado na tela do software de simulação. É necessária, além da definição do destino, uma confirmação visual deste, feita através do comando responsável por gerar a trajetória da cadeira. Após o programa receber a confirmação do destino da cadeira, ele executa as rotinas referentes à definição de retas e rotações, na sequência adequada para a cons-

tração das trajetórias desejadas. As trajetórias são definidas levando em consideração a análise do jerk, porém, não ficando evidente na interface. Muito embora quando implementada na prática será facilmente percebida.

A integração entre o software de simulação e o software de reconhecimento de voz Julius foi feita através de sockets do tipo UDP, pois o LabWindows fornece a opção de ativação de um servidor de comunicação, que monitora toda a comunicação do tipo UDP numa porta pré-definida.

O Julius foi configurado para enviar o comando reconhecido toda vez que o mesmo for considerado válido (que ocorre quando o resultado dos reconhecimentos de cada vocabulário é idêntico) para o endereço de IP local do computador que está sendo utilizado e a porta configurada no programa da simulação. Uma vez o comando recebido pelo programa de simulação, o mesmo é interpretado e as ações correspondentes são executadas.

3 Experimentos

3.1 Funcionamento da simulação

Ao executar o programa de simulação desenvolvido, uma janela com o mapa do ambiente é apresentada, conforme mostrado na Figura 4. Na entrada do ambiente, aparecerá um quadrado representando a cadeira de rodas.

Em seguida, o usuário deverá dizer o destino que deseja. Para isso, o comando falado deve ser composto pela palavra Destino seguida do nome do cômodo. Por exemplo, se o destino for quarto, o usuário deverá dizer “Destino Quarto”.

Uma vez reconhecido o comando corretamente, o campo Destino na janela da aplicação será automaticamente alterado para “Quarto”. Caso o campo Destino não esteja de acordo com o solicitado, o usuário deverá falar novamente até que obtenha o destino desejado. Nesta situação, o sistema fica a espera do usuário confirmar a trajetória dizendo “Gerar Trajetória”, para então o sistema executar a trajetória.

Ao executar a trajetória, um rastro é traçado ao longo do percurso executado pela cadeira (veja Figura 4). Para evitar poluição visual ou confusão de trajetórias, o usuário pode solicitar que a trajetória seja apagada através do comando “Limpa Tela”, que mantém a cadeira na última posição de destino alcançada.

Toda vez que uma nova trajetória é executada, verifica-se a necessidade de uma rotação inicial para que a cadeira inicie o movimento no sentido e direção corretos de acordo com o destino desejado. As rotações são realizadas sempre percorrendo o menor ângulo. Além disso, não há giros ao final da trajetória, a cadeira sempre chega com o mesmo ângulo que percorreu o último segmento de reta.

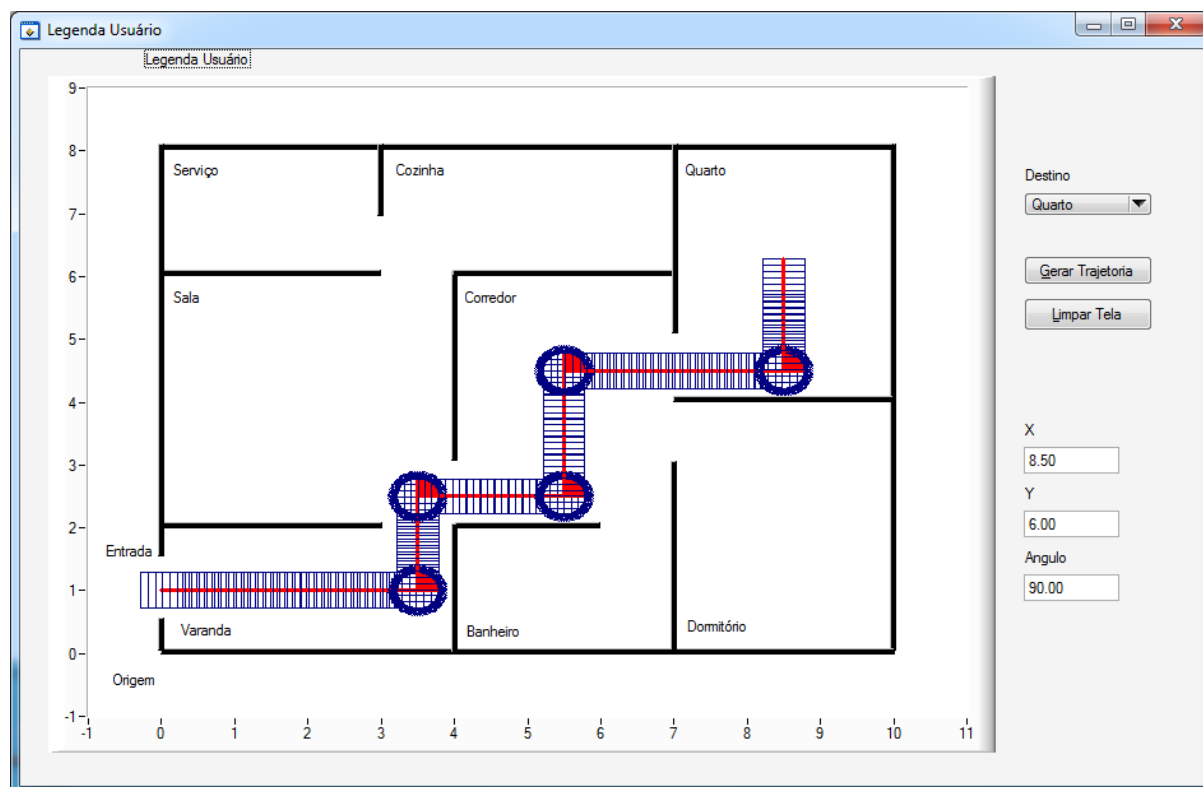


Figura 4. Interface do software desenvolvido com um exemplo de trajetória

3.2 Testes dos comandos de voz do sistema

Os comandos são do formato “destino + cômodo final”, “gera trajetória” e “limpa tela”. Cada comando foi pronunciado dez vezes em um ambiente próximo ao ideal, ou seja, com baixo nível de ruído. O desempenho do sistema reconhecimento de comandos de voz nestas condições é apresentado na Tabela 3.

Nos testes realizados, não foi encontrados erros. A taxa de erros de 0% foi atingida graças a estratégia de rejeição de comandos onde o sistema não tem certeza no resultado, o que gera rejeição. Toda vez que ocorre uma rejeição, o locutor deve falar novamente o comando desejado. Por outro lado, a taxa de acerto indica que de cada 10 comandos, 9 foram reconhecidos corretamente pelo sistema.

Com respeito a influência do ruído, uma análise mais detalhada é realizada no trabalho de Silva (2013).

Tabela 3. Resultados do Teste do Sistema de Reconhecimento de Comandos de Voz

Desempenho do Sistema		
Taxa de Acerto (%)	Taxa de Erros (%)	Taxa de Rejeição (%)
89,1	0	10,9

4 Conclusão

Nesse trabalho preocupou-se principalmente com o conforto do cadeirante. Quando se tem uma pessoa sendo transportada em uma cadeira de rodas existe a necessidade em tornar o movimento confortável para o usuário. Isso foi obtido tratando o jerk no começo e no final da trajetória, mesmo não ficando evidente nesta etapa de simulação. Porém, na prática esta análise será bem perceptível. Além disso, o comando de voz elimina a necessidade de um controle manual.

O ambiente residencial tratado na simulação é estruturado e com mapa, assim possibilitou a utilização de trajetórias retilíneas, que possuem implementação simplificada.

A simulação é uma etapa essencial para o desenvolvimento do projeto embarcado, pois nela foi possível validar os algoritmos de geração de trajetória, detectar erros na implementação e realizar inúmeros testes em menor tempo, evitando assim comportamentos imprevistos na prática e facilitando posteriormente a implementação embarcada.

A estratégia de reconhecimento de voz adotada é robusta a interferência de ruído a, que foram desconsiderados na simulação. Assim, a estratégia da utilização de dois vocabulários distintos para a etapa de reconhecimento de voz apresentou resultados satisfatórios.

Além da estratégia de aumento da robustez ao ruído do sistema de reconhecimento de voz, foi acrescentado uma etapa de confirmação, o que elimina qualquer possibilidade de execução de uma trajetória não solicitada.

Agradecimentos

Ao Ifes pela infraestrutura disponibilizada para a realização do trabalho e a FAPES pelo financiamento do projeto MOBI.

Referências Bibliográficas

- Anusuya, M. A. and Katti, S. K (2009). Speech recognition by machine: a review. *International Journal of Computer Science and Information Security*, Vol. 6, No. 3, pp. 182-205.
- Cruz, C.; Celeste, W.; Bastos, T. F., (2011). A robust navigation system for robotic wheelchairs. *Control Engineering Practice*, Vol. 19, Issue 6, pages 575-590.
- Latombe, J. 1991. *Robot Motion Planning*. Kluwer Academic Press, ISBN 0-7923-912902.
- Lee, Akinobu. 2010. *The Juliusbook*, Disponível em: (<http://keihanna.dl.sourceforge.jp/julius/47534/JJuliusboo-4.1.5.pdf>).
- Neto, N. et al (2011). Free tools and resources for Brazilian Portuguese speech recognition. *Journal of the Brazilian Computer Society*, Vol. 17, No. 1, pp 53-68.
- Pacnik, G; Benkic, K. and Brecko, B (2005). Voice operated intelligent wheelchair – VOIC. *Proceedings of the IEEE International Symposium on Industrial Electronics*, Vol. 3, pp.1221-1226.
- Park, Kwang-Hyun et al (2007). Robotic smart house to assist people with movement disabilities. *Autonomous Robots*, Vol. 22, pp.183–198.
- Rew, K.H.; Kim, K.S., (2010). A Closed-Form Solution to Asymmetric Motion Profile Allowing Acceleration Manipulation. *IEEE Transactions on Industrial Electronics*, Vol. 57.
- Silva, B.L.S.; Andreão, R.V.; Lima; L.E.M., (2013). Reconhecimento de Voz Aplicado ao Acionamento de uma Cadeira de Rodas. *XI Simpósio Brasileiro de Automação Inteligente*, Fortaleza, Brasil. Artigo submetido.
- Yuksekkaya B. et al (2006). A GSM, internet and speech controlled wireless interactive home automation system. *IEEE Transactions on Consumer Electronics*, Vol. 52, No. 3, pp. 837-843.