

## **ROBÔ ACIONADO POR COMANDO DE VOZ**

Gustavo Madeira da Silva -

Gustavo.madeira@alunos.unis.edu.br Marcos Basílio da Silva -

Marcos.basilio@alunos.unis.edu.br Matheus Basílio da Silva -

Matheus.basilio@alunos.unis.edu.br Rafael de Almeida

Oliveira - Rafael.oliveira2@alunos.unis.edu.br

### **RESUMO**

O presente trabalho tem como escopo a apresentação do funcionamento e construção de um robô autônomo móvel, com a capacidade de interpretar comandos de voz e relacioná-los a ações, visando, de forma primordial, facilitar a realização de atividades humanas, isso é, a mão de obra humana, podendo-se citar a título de exemplo, trabalhos perigosos, bem como aqueles que expõem o ser humano à situações de risco. A navegação do robô autônomo e o reconhecimento dos comandos são realizados por meio da utilização do Arduino UNO, sendo este um microcontrolador de placa única, com a finalidade de percorrer uma trajetória determinada pelo comando. Para tanto, a metodologia utilizada visou sustentar a construção de um robô comandado por meio de voz, almejando a sustentação, ou melhor, o embasamento para a construção de possíveis trabalhos futuros nesta área. Pode-se ressaltar que a técnica de pesquisa foi voltada para a leitura, em primeiro momento, exploratória, assim como, seletiva, em acervo bibliográfico, bem como documental e sítios pertinentes quanto ao assunto constante do recorte desta pesquisa. Por fim, analisa-se que os resultados alcançados com a construção do referido robô são satisfatórios, demonstrando êxito do método escolhido, ou melhor, utilizado.

**Palavra-Chave:** robô, robô autônomo móvel, comandos de voz, Arduino UNO

## **INTRODUÇÃO**

Desde os primórdios de sua origem, o ser humano sempre utilizou ferramentas e utensílios para auxiliar a realização de tarefas cotidianas relacionadas às suas necessidades de sobrevivência (Romano, V.; Dutra, M, 2002). Ao longo dos séculos, esse auxílio e facilitação foram obtidos sem sua grande maioria com a utilização da tecnologia, sendo esta sinônimo de robótica, nas últimas décadas.

Um robô é definido, segundo a norma ISO (International Organization for Standardization) 10218, como “uma máquina manipuladora, com vários graus de liberdade, controlada automaticamente, reprogramável, multifuncional, que pode ter base fixa ou móvel para utilização em aplicações de automação industrial”.

Na área da robótica, é evidente o potencial de aplicações dos robôs autônomos móveis, em especial em tarefas perigosas, sujas ou desagradáveis. As aplicações propostas para robótica avançada abrangem grande parte da atividade humana, o que incluem: luta contra o fogo, salvamento de emergência, prevenção de desastres, patrulha de segurança, limpeza industrial e doméstica, serviço doméstico, manipulação de pacientes, operações de procura em ambientes distantes e/ou inóspitos, etc. Uma característica comum destas aplicações é a exigência de um veículo móvel ou plataforma, que sirva como base para qualquer tarefa exigida. A exigência comum para tal veículo, é a habilidade para navegar de uma posição conhecida a uma nova localização (Jácomo, Justo Emilio Alvarez, 2001).

Diante do exposto, é imprescindível salientar que o presente trabalho tem como objetivo a descrição do projeto e implementação de um sistema robótico com a capacidade de interpretar comandos de voz, sendo capaz de se mover de uma posição conhecida a uma nova localização determinada pelo comando.

## **2- REVISÃO DE LITERATURA**

### **2.1- RECONHECIMENTO DE VOZ**

A voz humana é a forma mais simples, natural e eficaz do ser humano expressar seus pensamentos e transmitir informações. Neste contexto, a capacidade de percepção e compreensão da referida voz por equipamentos visando facilitar e agilizar a comunicação homem-máquina,

tem sido amplamente desenvolvida nas últimas décadas.

O funcionamento do reconhecimento de voz pode ser resumido basicamente como: a voz humana é percebida como variações da pressão sonora, e deve ser convertida em sinal elétrico por meio de um transdutor eletroacústico, como o microfone. Após passar por um circuito condicionador, que garante nível e cortes de frequências adequados, o sinal é convertido de analógico para digital por um conversor A/D (CARDOSO, SERGIO A., 2010).

### **2.1.1- MÉTODOS DE RECONHECIMENTO**

O reconhecimento da fala é subdividido no Reconhecimento Automático do Locutor (RAL) e no Reconhecimento Automático da Voz (RAV), sendo que aquele busca diferenciar os indivíduos tanto pela identificação como pela verificação de locuções, por meio da extração de características distintas que cada indivíduo apresenta (VALIATI, Joao Francisco, 2000), enquanto este, por sua vez, deve compreender automaticamente uma elocução, que pode ser tratada sob 3 aspectos: o reconhecimento de palavras isoladas, palavras conectadas e fala contínua.

Deve ser ressaltado que uma das aplicações do RAV condiciona-se à dependência ou não do locutor no reconhecimento, sendo que sistemas dependentes do locutor são capazes de reconhecer a fala de um único locutor ao qual foi treinado. Este será o sistema de reconhecimento automático de voz utilizado neste trabalho e os usuários deste tipo de sistema devem sempre treiná-lo antes de utilizá-lo. O treinamento envolve alguns fatores relativos, quais sejam: a escolha de um vocabulário; cada sistema é treinado a fim de reconhecer a fala de um determinado locutor e seu vocabulário específico; a necessidade que cada usuário repita o vocabulário de palavras ordenadas diversas vezes para ser criado um padrão e por fim, o sistema deve ser treinado em um ambiente apropriado.

Sendo assim, após a conclusão da fase de treinamento pelo usuário, este será habilitado pelo sistema ao seu vocabulário específico.

Segundo os autores Valiati e João Francisco (2000), o citado sistema possui algumas vantagens e desvantagens. Entre as vantagens, destacam-se: A flexibilidade do vocabulário (o usuário é capaz de adicionar ou alterar palavras para um vocabulário simples, bastando-lhe um treinamento prévio); o tamanho do vocabulário (o sistema pode suportar um vocabulário extenso). Por outro lado, as desvantagens observadas em sistemas deste tipo são: a limitação do número de usuários; a necessidade de uma grande quantidade de processamento antes que o

sistema venha a ser utilizado e o fato de certas aplicações poderem englobar muitos usuários, pode ocasionar uma elevada capacidade de armazenamento do sistema para os parâmetros dos diversos locutores; a fase de treinamento é desagradável ao usuário e devido a fatores como, cansaço, doença, posição e tipo de microfone, a voz do mesmo locutor poderá variar.

## **APLICAÇÕES DO RECONHECIMENTO DE VOZ**

### **INTERFACES EM SISTEMAS DE COMPUTAÇÃO**

As aplicações deste tipo utilizam o reconhecimento de fala como uma interface para sistemas de computador, assim como para aplicativos que executam sobre esses sistemas. A maior parte destes, utilizam a fala para a manipulação de menus, janelas, caixas de diálogo, palhetas e outros componentes.

### **AUXÍLIO A DEFICIENTES**

Os deficientes, em sua grande maioria, com impedimento motor ou visual, usam o reconhecimento de fala para realizar tarefas que os tornem mais independentes em sua vida pessoal, sendo o surgimento de tal reconhecimento, principalmente direcionado ao auxílio das pessoas com deficiências físicas.

### **CONTROLE DE AMBIENTES**

O controle de cadeiras de roda, camas de hospital, luzes e demais itens do ambiente pessoal são úteis para auxiliar em tarefas desta natureza por meio da voz, tornando o ambiente adequado às necessidades de cada usuário.

Além dessas três principais aplicações, o comando de voz está amplamente difundido no cotidiano, estando presente em *smartphones*, podendo citar, a título de exemplificação,

o *Google* assistente, a Siri, casas com expoentes como o *google home* e Alexa, bem como vários eletrodomésticos e carros fabricados com comando acionado por voz.

### 3 -METODOLOGIA

Para atingir a capacidade de interpretar comandos de voz, sendo capaz de obedecê-los e se mover de acordo, foi construído um robô, que consiste em um Arduino modelo UNO R3, módulo de reconhecimento de voz modelo V3.0, ponte H modelo L9110S, dois motores DC e fonte de tensão independente.

#### -FERRAMENTAS

Para o presente trabalho foi escolhido desenvolver um robô de plataforma retangular sobre três rodas, sendo duas delas movidas por motores e a terceira de apoio, conforme demonstrado pela figura 3.1. A construção foi escolhida da forma exposta, visando facilitar a realização de curvas.

Figura 3.1- Aspectos estruturais

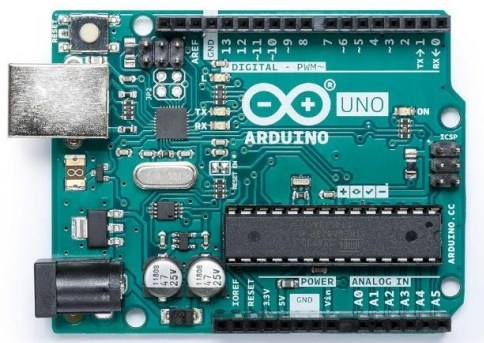


Fonte: [filipeflop.com](http://filipeflop.com)

Sendo assim, o sistema de controle está baseado no microcontrolador Arduino UNO R3, conforme demonstrado na figura acima.

O Arduino é uma plataforma de código aberto baseada em *hardware* e *software* de usar. A plataforma é baseada na linguagem C/C++ e neste trabalho, sendo ele responsável por processar as informações enviadas pelo módulo de voz e aplicá-las no programa, controlando o movimento do robô.

Figura 3.2- Arduino UNO R3



Fonte: Amazon.com

A interpretação dos comandos de voz fica sob a responsabilidade de outro *hardware*, o módulo de voz, Módulo de reconhecimento de Voz V3 da empresa ELECHOUSE, conforme retratado na imagem, tendo em vista que este aparelho ao interpretar os comandos, os transforma em dados que são enviados ao Arduino e integrados no programa.’

Figura 3.3- Módulo de reconhecimento de Voz V3



Fonte: elechouse.com

Para controlar a velocidade e o sentido de giro dos motores por PWM, do inglês *Pulse Width Modulation* e em português Modulação por Largura de Pulso (MLP), utilizou-se a ponte H modelo L9110S, como analisado na figura 3.4, a qual possui entrada para dois motores, que podem ser de até 12V e 800mA. O dispositivo, além de possibilitar o controle dos motores, permite que sejam controlados de forma independente.

Figura 3.4- Ponte HL9110S



Fonte:multipecas.curitiba.br

Além das componentes citados anteriormente, também foram utilizados para realização deste trabalho outros equipamentos, sendo eles: 2 Motores DC(3~6v), 1 Bateria 9v., 4 Pilhas alcalinas 1,5v., 1 Suporte para 4pilhas, 1 Clip para alimentar Arduino utilizando bateria 9v., 2 Rodas deborracha, 1 Roda boba(universal), 15 Jumpers, 1 Mini Protoboard 170pontos, 1 Jogo deparafusos.

### **- MÉTODO DE DESENVOLVIMENTO DA PROGRAMAÇÃO**

O objeto a ser trabalhado será capaz de obedecer inicialmente a quatro comandos de voz, os quais dois são para distância de locomoção e dois de mudança de trajetória. Os comandos de locomoção serão os de se mover em distâncias de 1 metro e para se locomover em uma trajetória determinada.

Diante disso, pondera-se que a mudança de trajetória será estabelecida com os comandos de virar à esquerda e virar à direita e para realização deste trabalho, é necessário um algoritmo capaz de alcançar o objetivo estabelecido.

Com o objetivo de alcançar o programa final, o desenvolvimento do programa foi dividido em 3 etapas. Cada etapa focará em uma parte importante visando entendimento dos componentes utilizados, sendo a 1 delas, a que consiste em utilizar no lugar dos motores dois LEDs, que devem ser acionados por 4 botões, que simulam os comandos de voz. O botão 1 aciona os dois LEDs por 1 segundo, o botão 2 aciona somente o LED 1 por 1 segundo e o botão 3 faz o mesmo com o LED 2, o botão 4 aciona os dois LEDs por 1 segundo, posteriormente aciona somente o LED 1 por 1 segundo e volta a acionar os dois LEDs por 1 segundo.

#### **Programaetapa 1:**

```

int led2=3;
int b1=4;
int b2=5;
int b3=6;
int b4=7;

void setup() {

  pinMode(led1, OUTPUT);]
  pinMode(led2, OUTPUT);
  pinMode(b1, INPUT);
  pinMode(b2, INPUT);
  pinMode(b3, INPUT);
  pinMode(b4, INPUT);
  }

void loop() {

  if(digitalRead(b1)==HIGH){
    digitalWrite(led1, HIGH);
    digitalWrite(led2, HIGH);
    delay(1000);
  }
}

```

#### **4 - RESULTADOS EDISCUSSÕES**

Como apresentado na revisão literária, o módulo de reconhecimento de voz (V3) usado, utiliza uma das aplicações do RAV, no qual condiciona-se à dependência do locutor no reconhecimento. Dito isso, foi escolhido um vocabulário específico, em que foram escolhidas as palavras: deslocar, direita, esquerda e trajeto.

Sendo assim, como o reconhecimento ficou limitado a um único locutor, este foi responsável pelo treinamento do módulo no reconhecimento, sendo dito várias vezes as palavras usadas.

#### **-PROTÓTIPO**

Com o módulo reconhecimento de voz devidamente preparado, assim como todo o



programa, foi construído, em seguida, um protótipo, conforme se observa na figura 4.1. Com ele, foi obtido o tempo necessário para os deslocamentos escolhidos.

Figura 4.1 – Protótipo

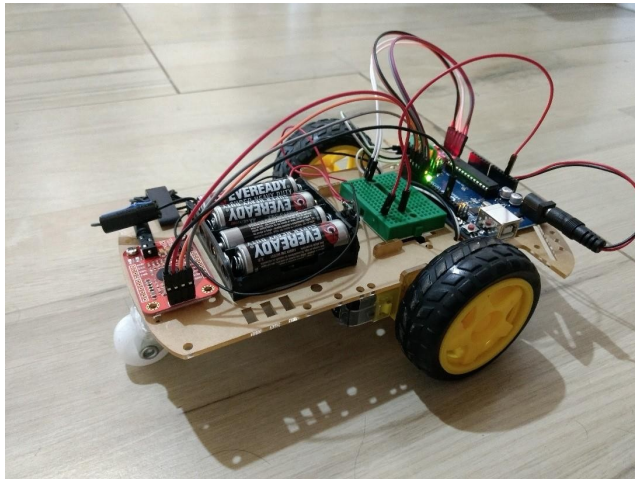
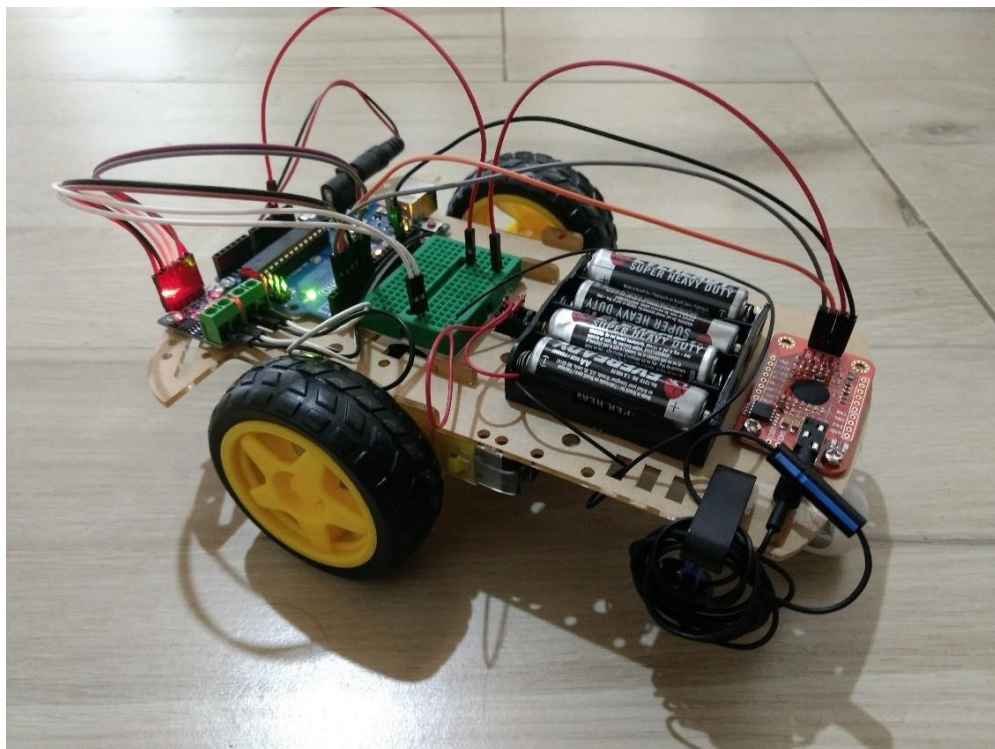


Figura 4.3 – Protótipo



Para determinar o tempo necessário para alcançar 1 metro, foi demarcado no solo, bem como cronometrado tempo que o objeto leva para alcançar 1 metro, como demonstrado nas figuras abaixo.

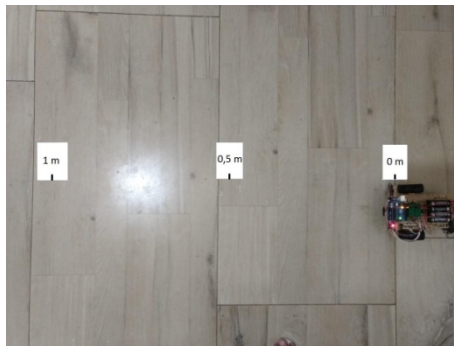


Figura 4.4 – cálculo do tempo de deslocamento

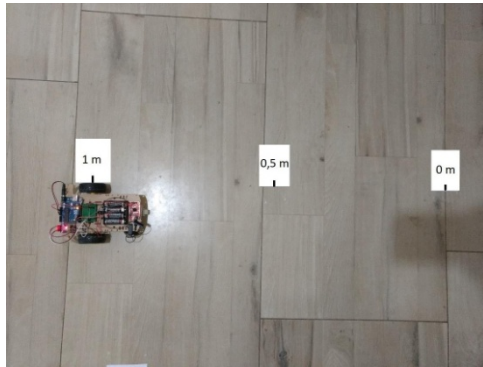


Figura 4.5 - cálculo do tempo de deslocamento

Logo, para determinar o tempo necessário para virar 90 graus, foi demarcado no solo o marcador de tempo que o objeto leva para virar, como pode-se analisar nas figuras abaixo.

Figura 4.6 - cálculo do tempo de rotação

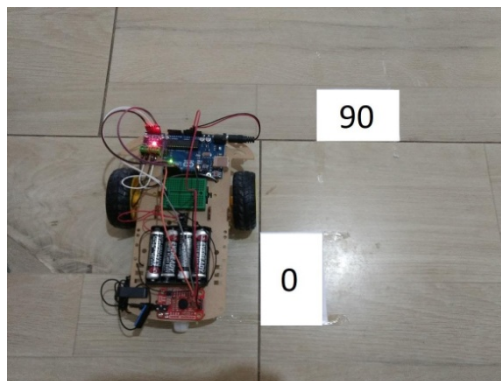
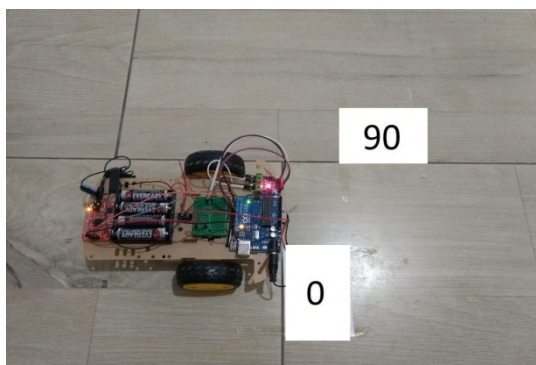


Figura 4.7 - cálculo do tempo de rotação



## - DIFICULDADES NO DESENVOLVIMENTO

Para o desenvolvimento do protótipo, as principais dificuldades foram apresentadas no reconhecimento de voz, tendo em vista que sua precisão está diretamente ligada ao microfone utilizado, ao tom de voz do locutor e ao nível de ruído ao qual o objeto está inserido. Estes três fatores oscilam constantemente, fato este, que dificultou consideravelmente o projeto, o que teve como influência direta no trabalho realizado. Isto poderá ser amenizado em trabalhos futuros com a utilização de métodos de reconhecimento de voz melhor elaborados, como por exemplo, os desenvolvidos pelo Google, que são aplicados principalmente em *smartphones*.

Verificou-se ainda, a existência de um outro problema, sendo este, no tempo de deslocamento, uma vez que há o oscilamento de acordo com o nível de carga das pilhas utilizadas e com o tipo de solo no qual o objeto está inserido.

## 5 - CONCLUSÃO

Como demonstrado, o reconhecimento de voz é amplamente explorado pela tecnologia e sua aplicação na robótica é formidável, principalmente em robôs autônomos móveis. Com isso, o aprendizado adquirido no desenvolvimento de projetos iniciais, como este, é necessário para o desenvolvimento de projetos futuros com uma aplicação mais complexa, como por exemplo, o auxílio em tarefas ariscadas e em caso de impossibilidade de locomoção.

Pode-se ponderar então, que o presente trabalho trata-se de um ótimo ponto de partida para o desenvolvimento destes projetos.

Neste trabalho, todos os objetivos apresentados foram alcançados e pôde ser concluído que o sistema desenvolvido tem a capacidade de interpretar comandos de voz, sendo capaz de se mover de uma posição conhecida a uma nova localização determinada pelo comando.

## REFERÊNCIAS BIBLIOGRÁFICAS

CARDOSO, SERGIO A. et al. Sesame: sistema de reconhecimento de comandos de voz utilizando pds e rna. In: **Anais do XVIII Congresso Brasileiro de Automática**. 2010. p.1316-1323.

VALIATI, Joao Francisco. Reconhecimento de voz para comandos de direcionamento por meio de redes neurais. 2000.

JÁCOBO, Justo Emilio Alvarez. Desenvolvimento de um Robô Autônomo Móvel Versátil utilizando Arquitetura Subsumption. **Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas**, 2001.

ROMANO, V.; DUTRA, M. Introdução a robótica industrial. **Robótica Industrial: Aplicação na Indústria de Manufatura e de Processo, São Paulo: Edgard Blücher**, p. 1-19, 2002.