

USO DA TECNOLOGIA DE COMANDOS DE VOZ PARA DOMÓTICA

Marco Antonio Carvalho¹

Carlos Eduardo de França Roland²

Resumo

A busca por facilidade e comodidade é inerente à raça humana. A evolução da espécie é baseada nas mudanças provocadas pelo desenvolvimento de ferramentas, equipamentos e processos que garantam a perpetuação da espécie e melhorem a qualidade de vida. A partir do desenvolvimento das tecnologias digitais a humanidade passou a ter acesso a dispositivos eletrônicos que mudaram a forma como os relacionamentos ocorrem entre pares e com as máquinas, mais especificamente com os computadores. Em função dessas mudanças, surgem novas áreas do conhecimento humano que propõem formas mais adequadas de se conviver neste cenário. Uma das áreas de estudo é do interfaceamento com máquinas eletrônicas digitais por Reconhecimento de Voz. Neste contexto, percebe-se a busca por utilização de tecnologia de reconhecimento de fala aplicada a processos de automação e controle, especialmente residenciais denominada Domótica, que motivou o desenvolvimento da pesquisa apresentada neste trabalho. Com base em pesquisa bibliográfica exploratória, buscou-se entender quais os conceitos envolvidos na aplicação de reconhecimento de voz para automação e controle baseado em plataformas micro controladas, identificando-se equipamentos e componentes necessários para o desenvolvimento de sistemas operados por voz. Tomou-se como exemplo a implementação de um sistema com um módulo de *hardware* para reconhecimento de fala disponível no mercado, associado a um controlador de arquitetura Arduino para controlar movimentos de um robô, para servir como base de aplicação à domótica. O resultado dos estudos mostra que, apesar da complexidade envolvida no reconhecimento de fala, a tecnologia digital está suficientemente evoluída tornando mais fácil e acessível a implementação de sistemas até por pessoas sem formação específica em eletrônica e computação.

Palavras-chave: Arduino. Automação Residencial. Comandos por Voz. Domótica. Reconhecimento de Voz.

Abstract

The search for ease and convenience is inherent in the human race. The evolution of the species is based on changes brought about by the development of tools, equipment and processes that ensure the perpetuation of the species and improve the quality of life. Since the development of digital technologies, humanity has gained access to electronic devices that have changed the way relationships occur between peers and with machines, more specifically with computers. Due to these changes, new areas of human knowledge emerge that propose more appropriate ways to live in this scenario.

¹ Graduando em Análise e Desenvolvimento de Sistemas pela Fatec Dr Thomaz Novelino – Franca/SP. Endereço eletrônico: [marco.acarvalho76@gmail.com].

² Docente no curso de Análise e Desenvolvimento de Sistemas da Fatec Dr Thomaz Novelino – Franca/SP. Endereço eletrônico: [carlos.roland@fatec.sp.gov.br].

One of the areas of study is interfacing with Digital Electronic Speech Recognition machines. In this context, one can see the search for the use of speech recognition technology applied to automation and control processes, especially residential, called Domotics, which motivated the development of the research presented in this paper. Based on exploratory bibliographic research, we sought to understand the concepts involved in the application of voice recognition for automation and control based on micro controlled platforms, identifying equipment and components necessary for its implementation. An example of implementation of a system with a commercially available speech recognition hardware module, coupled with an Arduino architecture controller to control the movements of a robot was used to serve as a basis for application to home automation. The result of the studies shows that, despite the complexity involved in speech recognition, digital technology is sufficiently evolved making it easier and more accessible to implement systems even for people with no specific background in electronics and computing.

Keywords: *Arduino. Domotics. Home Automation. Voice Commands. Voice Recognition.*

1 Introdução

A espécie humana se diferencia das demais espécies vivas pela evolução proporcionada por mudanças buscadas com o objetivo de facilitar sua existência, garantir sua perpetuação no processo de seleção natural, e aumentar sua qualidade de vida.

Desde os pré-australopitecos, evoluindo pelos gêneros *Homo*, até o homem moderno, a busca por desenvolvimento de ferramentas, equipamentos e processos tem sido contínua, e no último século com crescimento exponencial. As Tecnologias da Informação e Comunicação (TIC) oferecem uma plataforma de equipamentos digitais (*hardware*) que podem ser programados (*software*) para realizar os mais diferentes processos de automação que têm, consistentemente, aumentado suas capacidades de processamento e armazenamento e diminuído seus preços, tornando-se a cada dia mais acessível à população.

Neste cenário de busca por maior qualidade e facilidade de vida, a interação humana com dispositivos computacionais tem recebido especial atenção para a interface de comunicação e controle por voz.

Assim nasceu a proposta de pesquisa de utilização de interface de voz para operação de processos de automação e controle residenciais, baseada em plataformas de sistemas digitais micro controlados. A partir da definição deste tema, realizou-se uma pesquisa bibliográfica exploratória para se conhecer mais

profundamente os elementos e os processos envolvidos nessa área do conhecimento humano, o que permitiu a definição das questões de pesquisa norteadoras do projeto. Buscou-se então identificar dispositivos de *hardware* e de *software* que permitam o desenvolvimento de sistemas de automação, controle e segurança residenciais, caracterizados como sistemas Domóticos, e como implementá-los sob os aspectos técnico e econômico.

Neste trabalho são apresentados, na seção 2, os conceitos fundamentais dos elementos envolvidos no tema e nas questões de pesquisa; na seção 3, os elementos de *hardware* e de *software* como um exemplo prático da utilização componentes micro controlados para aplicação de domótica operada por comandos de voz; em seguida são discutidos os resultados dos estudos (seção 4); e tecidas considerações sobre a realização do projeto de pesquisa. Por fim são relacionadas as Referências utilizadas que suportam conceitos, definições e características dos elementos estudados.

2 Embasamento teórico dos componentes da pesquisa

Nesta seção são apresentados os elementos que contextualizam o tema, as questões norteadoras da pesquisa e a hipótese de solução às questões elencadas.

2.1 Domótica

A Domótica é uma ciência interdisciplinar que estuda o uso de tecnologias para melhorar a qualidade de vida e conforto em um ambiente habitacional. A palavra é originada do termo em latim *domus*, que significa casa. A Domótica requer a integração de áreas do conhecimento humano como engenharia, arquitetura, tecnologia da informação, automação, eletrônica, engenharia elétrica, telecomunicações e desenho industrial, para tornar a habitação inteligente (SULTAN e NABIL, 2019).

Os autores afirmam que à medida que dispositivos digitais de controle se tornaram mais acessíveis, tanto econômica quanto tecnicamente, a demanda por seu uso em processos de automação residencial aumentou, buscando-se tornar os ambientes habitacionais mais inteligentes e econômicos. Por décadas, profissionais da construção civil buscaram embarcar, em projetos habitacionais, dispositivos eletrônicos digitais para automação e controle, disponibilizando um conjunto de funções inteligentes desde o gerenciamento do uso de energia elétrica, aspectos de

autossustentabilidade, controle de segurança, entretenimento e assistência a deficientes à automação residencial ou doméstica.

Ainda segundo Sultan e Nabil (2019) a utilização de sistemas digitais para automação em ambientes organizacionais como edifícios de escritórios, comerciais ou industriais não tem encontrado dificuldade em função de existirem nestes contextos equipes técnicas de Tecnologia da Informação (TI) para projetar, implementar, e administrar tais sistemas. Entretanto quando se trata do uso doméstico dessas tecnologias, o foco principal em termos de arquitetura e estrutura da construção é de oferecer facilidade de uso e adaptabilidade. Casas inteligentes devem não apenas direcionar esforços de pesquisa e desenvolvimento de controles ambientais de temperatura e iluminação, mas ampliar o espectro de soluções com projetos adaptativos e sensíveis ao contexto de uso potencialmente abarcados por interesses de profissionais de arquitetura e de especialistas em interação humano-computador.

Neste cenário de requisitos ampliado encaixa-se a domótica operada por voz. O reconhecimento de voz é um componente chave no conjunto de soluções de domótica. Basicamente, é um tipo de tecnologia que permite que sistemas de informação computadorizados entendam palavras fornecidas por fala.

Controlar um dispositivo eletrônico por voz permite tanto facilidade de uso quanto aumento de eficiência e produtividade. Sistemas de reconhecimento de voz também auxiliam usuários a executarem duas ou mais tarefas simultaneamente enquanto trabalhando com computadores ou outros dispositivos (VATANSEVER, KUSCU, e TUNA, 2018).

2.2 Reconhecimento de Fala

Os sistemas modernos de reconhecimento de fala combinam tecnologias interdisciplinares tais como Processamento de Sinais, Reconhecimento de Padrões, Linguagem Natural, e Linguística em um *framework* estatístico unificado. Tais sistemas, que têm aplicação em um amplo espectro de problemas de processamento de sinais, representam uma revolução no Processamento Digital de Sinais (PDS). Como um campo dominado por processadores orientados a vetores e matemática baseada na álgebra linear, os sistemas PDS são baseados em sofisticados modelos estatísticos implementados usando complexos paradigmas de *software*. Tais sistemas são capazes de entender entrada de fala contínua usando-se vocabulários de milhares de palavras em ambientes operacionais (PICONE, 1996).

Considerando a abordagem do estado da arte da tecnologia, a produção de fala humana, assim como o processo de reconhecimento é modelado em quatro estágios: geração de texto, produção de fala, processamento acústico, e decodificação linguística. Dada a complexidade da interface humano-computador, fica evidente a necessidade de protocolos de avaliação que enderecem um grande número de diferentes tipos de sistemas de línguas faladas. As tarefas de reconhecimento de fala são classificadas em quatro categorias de acordo com os critérios de comunicação homem-homem ou homem-computador, e se essas comunicações são dialógicas ou monológicas (FURUI, 2007).

Nos sistemas de automação por comandos de voz, a comunicação é categorizada como homem-computador monológica definida como Ditado, segundo o autor. Nesta categoria, os enunciados falados são feitos com a expectativa de que serão convertidos exatamente em textos com caracteres exatos e por isso a espontaneidade da fala é mais simples, facilitando os processos computacionais de reconhecimento. Por se tratarem de comunicações por palavras isoladas, o início e o fim de cada palavra podem ser detectados diretamente a partir da energia do sinal recebido pelo sistema, tornando a detecção dos limites das palavras mais fácil do que em situações de discurso contínuo. Entretanto, há que se considerar que em ambientes reais de uso de sistemas de reconhecimento de fala, a comunicação é contaminada por ruídos que dificultam a identificação das palavras.

Algoritmos de reconhecimento de fala são baseados em Processamento de Linguagem Natural (PLN) que é um campo da Ciência da Computação que auxilia a interpretação dos comandos de voz emitidos pelo usuário. As soluções de reconhecimento de fala com PLN envolvem recursos de Inteligência Artificial e Linguística Computacional para permitir que sistemas computacionais entendam e respondam a comandos dados em linguagem natural humana. O principal desafio da PLN é entender a linguagem humana e habilitar computadores a derivar significado a partir de comandos por voz emitidos em linguagem natural. A maioria dos algoritmos de PLN são baseados em aprendizagem da máquina, mais especificamente em aprendizagem de máquina estatística (RANI et al., 2017).

O Reconhecimento de Voz é o subcampo interdisciplinar da linguística computacional que desenvolve metodologias e tecnologias que permitem o reconhecimento e a tradução da linguagem falada em texto por computadores. Também é conhecido como Reconhecimento Automático de Fala (ASR do termo em

inglês Automatic Speech Recognition), Reconhecimento de Fala por Computador ou apenas Fala para Texto (STT – Speech To Text). Ele incorpora conhecimento e pesquisa nos campos de linguística, ciência da computação e engenharia elétrica (BOYD, 2018; KINCAID, 2018; e ROUSE, 2018).

A pesquisa bibliográfica dos autores citados afirma que alguns sistemas de reconhecimento de fala exigem treinamento (também chamado de inscrição) em que um falante individual lê texto ou vocabulário isolado para configurar o sistema. O sistema analisa a voz específica da pessoa e a utiliza para afinar o reconhecimento da fala dessa pessoa, resultando em maior precisão. Sistemas que não usam treinamento são chamados de sistemas independentes de alto-falante, e os sistemas que usam treinamento são chamados de dependentes de alto-falante.

Aplicativos de reconhecimento de voz disponibilizados em diversas plataformas digitais incluem interfaces de usuário que recebem comandos de voz como nos *smartphones* para fazer discagem por voz (por exemplo com o comando: ligar para casa); para roteamento de chamadas com um comando fazer chamada a cobrar, por exemplo; como nos eletrodomésticos para controles domóticos; para operação de microcomputadores em pesquisas como para encontrar um *podcast* onde palavras específicas foram ditas; para entrada de dados simples como a inserção de um número de cartão de crédito; para preparação de documentos estruturados como um laudo de radiologia; e processamento de fala para texto (por ex.: operação de editores de textos ou gerenciadores de mensagens de e-mails) (BOYD, 2018; KINCAID, 2018; ROUSE, 2018).

O termo Reconhecimento de Voz refere-se à identificação do falante, e não ao que ele está dizendo. O reconhecimento do falante pode simplificar a tarefa de traduzir a fala em sistemas que foram treinados em uma voz de uma pessoa específica ou pode ser usado para autenticar ou verificar a identidade de um falante como parte de um processo de segurança (SINGH, 2014).

Do ponto de vista tecnológico, o reconhecimento de voz tem uma longa história com várias ondas de grandes inovações. Mais recentemente, o campo se beneficiou dos avanços em *deep learning* e *big data*. Os avanços são evidenciados não apenas pelo surgimento de trabalhos acadêmicos publicados no campo, mas principalmente pela adoção na indústria mundial de uma variedade de métodos de *deep learning* para projetar e implementar sistemas de reconhecimento de voz. Dentre as empresas da indústria da fala estão Google, Microsoft, IBM, Baidu, Apple, Amazon, Nuance,

SoundHound e iFLYTEK, muitos dos quais divulgaram a tecnologia básica em seus sistemas de reconhecimento de fala como sendo baseada em *deep learning* (VELDE, 2019).

No contexto do desenvolvimento Faça Você Mesmo, mais conhecido pelo acrônimo do termo em inglês – Do It Yourself (DIY) – a plataforma Arduino oferece um conjunto de componentes e acessórios que facilitam e viabilizam tanto técnica quanto economicamente o desenvolvimento de sistemas comandados por voz.

A seguir são apresentados alguns destes produtos, bem como características das plataformas micro controladas, amplamente disponíveis no mercado.

2.2.1 EasyVR Shield 3.0

O EasyVR Shield 3.0 é a mais popular ferramenta para reconhecimento de voz da atualidade, é um *shield* Arduino com uma seleção de 26 comandos integrados já disponíveis em inglês dos EUA, italiano, japonês, alemão, espanhol e francês prontos para executar ações, e um espaço para outros 32 comandos que precisam ser definidos pelo usuário e treinados no sistema. A placa possui um conector adicional de saída de linha e fone de ouvido, e acesso aos pinos de I/O do módulo (FORTEBIT, 2019).

2.2.2 ELECHOUSE Voice Recognition Module V3

O ELECHOUSE é um painel de reconhecimento de voz compacto e fácil de controlar que suporta até 80 comandos de voz no total, mas permite que no máximo 7 comandos de voz possam funcionar ao mesmo tempo. Antes de ser implementado, os usuários precisam treinar o módulo para permitir que ele reconheça qualquer comando de voz e qualquer som pode ser treinado como um comando.

Esta placa tem 2 formas de controle: Porta Serial (função completa), e Pinos de Entrada Geral (parte da função). Os pinos de saída gerais na placa podem gerar vários tipos de ondas enquanto o comando de voz correspondente é reconhecido (ELECHOUSE, 2019).

2.2.3 Grove Speech Recognizer

A Seeed Studio criou este dispositivo em forma de kit que pode reconhecer 22 comandos, incluindo iniciar, parar, reproduzir música, dentre outros. Toda vez que ele reconhecer um comando, ele retornará um valor e, em seguida, o alto-falante

conectado a ele repetirá o comando.

Esse valor pode ser usado para controlar outros dispositivos como um motor, tocador de música, ou outro atuador conectado ao sistema. Testes mostram que o equipamento possui uma alta taxa de reconhecimento e um falso *trigger* muito baixo (SEED, 2019a).

2.2.4 ReSpeaker Core

Outro acessório da Seeed Studio é o ReSpeaker Core que suporta reconhecimento de fala e de texto para fala funcionando tanto *off-line* quanto *online*. Com um microfone de extensão ele consegue perceber os comandos de voz emitidos à distância, mesmo com um alto ruído de fundo (SEED, 2019b).

O projeto foi iniciado no Kickstarter, a plataforma de financiamento coletivo (*crowdfunding*), em 23 de agosto de 2016 tendo em 1 mês superado a meta e levantado perto de US\$ 200 mil de investidores (KICKSTARTER, 2017).

Os acessórios apresentados são componentes que são ligados a plataformas micro controladas para a implementação de sistemas de automação e controle. Na seção seguinte, são apresentados os conceitos e as principais características desses dispositivos.

2.3 Micro Controladores

O micro controlador é um pequeno componente eletrônico, dotado de uma inteligência programável, utilizado no controle de processos lógicos, é pequeno porque em uma única pastilha de silício encapsulada existem todos os componentes necessários ao controle de um processo. Tais dispositivos são conhecidos como Sistemas em Uma Pastilha (SOC - do termo em inglês System On a Chip)

Dotado de inteligência programável porque possui Unidade Central de Processamento e Unidade Lógica Aritmética, onde todas as operações matemáticas e lógicas são executadas e memória principal onde toda a lógica é estruturada na forma de programas que são gravados na área de programa do componente. Assim toda vez que o micro controlador for energizado o programa interno é executado.

São utilizados para controle de processos em associação com periféricos como *leds*, *displays*, relés, sensores, motores, etc., e são chamados de controles lógicos porque a operação do sistema baseia-se nas ações lógicas que devem ser executadas, dependendo do estado dos periféricos de entrada e saída.

2.3.1 Arduino

O Arduino é uma plataforma de *hardware* e *software* de código aberto, desenvolvida a partir do projeto da plataforma Wiring apresentada na tese de mestrado de Hernando Barragán pelo Instituto de Design Interativo do IVREA (Itália) em 2003 (BARRAGÁN, 2019). Pesquisadores do IVREA adotaram o micro controlador ATmega8 numa plataforma que batizaram de Arduino que passaram a produzir e a comercializar por USD\$50. Assim a comunidade Maker (pessoas que desenvolvem seus próprios equipamentos e dispositivos eletromecânicos interativos sem necessariamente ter formação técnica nessas áreas) passaram a ter acesso simples e fácil a sistemas eletrônicos digitais para automação e controle. Tais sistemas interativos podem detectar e controlar objetos no mundo físico e digital com o uso de sensores e atuadores facilmente conectados à placa. O *software* para desenvolvimento da lógica de operação do sistema é licenciado sob a Licença Pública Geral Menor GNU (LGPL) ou a Licença Pública Geral GNU (GPL), e seu *hardware* sobre a licença Creative Commons Attribution Share-Alike 2.5, permitindo a fabricação de placas como as Arduino e a distribuição de *software* por qualquer pessoa, fazendo com que o projeto adquirisse uma comunidade de usuários que tornaram a plataforma Arduino muito maior do que a empresa (ARDUINO, 2019).

Os projetos da placa Arduino usam uma variedade de microprocessadores e controladores. As placas são equipadas com conjuntos de pinos digitais e analógicos de entrada e/ou saída (I/O) que podem ser interligados a várias placas de expansão ou *breadboards* (matrizes de contato para prototipação) e outros circuitos. As placas possuem interfaces de comunicação serial, incluindo USB em alguns modelos, que também são usadas para carregar programas a partir de computadores pessoais. As placas não possuem recursos para uso em redes locais, porém é comum combinar um ou mais acessórios (*shields*) para expandir suas funcionalidades. Os micro controladores são tipicamente programados usando um dialeto de recursos das linguagens de programação C++. O projeto Arduino fornece um ambiente de desenvolvimento integrado (IDE) baseado no projeto de linguagem Wiring, mas outros exemplos de linguagens que podem comunicar-se com a conexão serial são: Max/MSP, Pure Data, SuperCollider, ActionScript e Java.

Em 2010 foi realizado um documentário sobre a plataforma chamado Arduino: The Documentary (YOUTUBE, 2010).

2.3.2 NodeMCU

NodeMCU é uma plataforma para Internet das Coisas (IoT do termo em inglês Internet of Things) de código aberto.

A Internet das Coisas (IoT) é uma rede de dispositivos físicos fixos ou móveis, tais como veículos, eletrodomésticos e outros objetos que possuem tecnologia embarcada, sensores e conexão com rede capazes de coletar e transmitir dados. Ela emergiu dos avanços de várias áreas como sistemas embarcados, microeletrônica, comunicação e sensoriamento. De fato, a IoT tem recebido bastante atenção tanto da academia quanto da indústria, devido ao seu potencial de uso nas mais diversas áreas das atividades humanas (LUETH, 2014).

A Internet das Coisas, em poucas palavras, é uma extensão da internet atual, que proporciona aos objetos do dia a dia, quaisquer que sejam, capacidade computacional e de comunicação de dados para se conectarem à Internet. A conexão com a rede mundial de computadores viabiliza, primeiro, controlar remotamente os objetos, e segundo, permitir que os próprios objetos sejam acessados como provedores de serviços. Estas novas habilidades dos objetos comuns, geram muitas oportunidades tanto no âmbito acadêmico quanto no industrial. Todavia, estas possibilidades apresentam riscos e acarretam amplos desafios técnicos e sociais (LUETH, 2014).

Para tornar a implementação de sistemas de IoT mais fáceis e acessíveis, foram criadas as plataformas NodeMCU que são compostas por um módulo micro processado associado a um outro para comunicação de dados. O NodeMCU inclui *firmware* que funciona com o módulo de comunicação sem fio ESP8266 Wi-Fi SoC da Espressif Systems, e *hardware* baseado no módulo ESP-12. O termo NodeMCU, por padrão, refere-se ao *firmware* em vez dos kits de desenvolvimento. O *firmware* usa a linguagem de *script* Lua e baseia-se no projeto eLua, tendo sido construído no SDK Espressif Non-OS para ESP8266 (YUAN, 2017).

No verão de 2015, os criadores abandonaram o projeto do *firmware* e um grupo de colaboradores independentes, mas dedicados, assumiu o controle. No verão de 2016, o NodeMCU incluiu mais de 40 módulos diferentes. Devido a restrições de recursos, os usuários precisam selecionar os módulos relevantes para seu projeto e criar um *firmware* adaptado às suas necessidades (YUAN, 2017).

2.4 Sistemas Embarcados

Diferente de um sistema operacional generalista, como os que normalmente encontramos em microcomputadores, um sistema embarcado é um sistema computacional programado com uma função dedicada dentro de um sistema mecânico ou elétrico maior, geralmente com restrições de computação em tempo real. Ele é incorporado como parte de um dispositivo completo, muitas vezes incluindo peças mecânicas e de *hardware*. Sistemas embarcados controlam muitos dispositivos em uso comum atualmente, e podem abranger desde processadores de 8-bits, até processadores de 64-bits multicore e ASICs.

Os sistemas embarcados modernos são frequentemente baseados em micro controladores, mas também são comuns microprocessadores usando chips externos para circuitos de memória e periféricos, especialmente em sistemas mais complexos. Em ambos os casos, os processadores utilizados podem ser de tipos variados, desde fins gerais até aqueles especializados em determinada classe de cálculos, ou até mesmo customizados para o aplicativo em questão.

Os sistemas embarcados variam de dispositivos portáteis, como relógios digitais e MP3 *players*, até grandes instalações fixas, como semáforos, controladores de fábrica e sistemas amplamente complexos, como veículos híbridos, ressonância magnética e aviônicos. A complexidade varia de baixa, com um único *chip* micro controlador, até altíssima, com várias unidades, periféricos e redes montadas dentro de um grande chassi ou gabinete.

3 Reconhecimento de Voz para Automação e Controle

Nesta seção é apresentado um exemplo de implementação de sistema de controle por voz, utilizando o *shield* EasyVR Shield 3.0 apresentado no Referencial Teórico e mostrado na Figura 1. O exemplo comentado, adaptado de Nasir (2013), é do projeto de controle de movimentos de um robô usando comandos de voz. Para isso, foram definidos os comandos AVANÇA, VOLTA, DIREITA, ESQUERDA e PARE.

Figura 1 – EasyVR Shield 3.0



Fonte: Maker.io (2019)

Para a implementação do sistema, o autor indica que se divida em partes para se entender os conceitos envolvidos no projeto. Em primeiro momento, criam-se os comandos de operação através do *software* EasyVR Commander fornecido pelo fabricante do *shield*, e a partir dele gravam-se os códigos de reconhecimento de voz na memória de dados da placa.

Então, procede-se ao interfaceamento do *shield* EasyVR com uma placa Arduino UNO (Figura 2), por razões de compatibilidade de pinagem entre as duas placas. Outros modelos de Arduino podem ser usados, mas as ligações elétricas entre eles deverão ser feitas com uso de cabos e matriz de contato (*proto-board*).

Para a gravação dos códigos para reconhecimento dos comandos de voz no *shield* montado sobre o Arduino UNO, é necessário se mudar o *jumper* J12 para a posição PC para que ele receba os dados do *software*. Então se conecta a porta USB do Arduino UNO ao computador que executa o EasyVR Commander e procede-se à sequência de operações para criação e gravação dos comandos na placa. Detalhes dessa operação são apresentados em Nasir (2013). O processo de gravação da voz com os comandos é feito em duas fases para melhores resultados de reconhecimento e cada comando é associado a um número inteiro, por exemplo 0 para o comando AVANÇA, 1 para o comando VOLTA, etc. que será transmitido à placa controladora.

Figura 2 – Arduino UNO



Fonte: Arduino.cc (2019)

O conjunto montado para o projeto é apresentado na Figura 3.

Figura 3 – Conjunto de placas montadas para implementação do projeto



Fonte: Nasir (2013)

Tendo criado os comandos, encerra-se a execução do EasyVR Commander e muda-se a posição do *jumper* J12 do *shield* para a posição SW, para que ele opere como dispositivo reconhecedor de comandos de voz.

A etapa seguinte é desenvolver o *sketch* que é o conjunto de instruções de programação para o Arduino UNO receber os dados do EasyVR, identificar qual comando foi enviado ao sistema e realizar a operação solicitada. Para esta etapa utiliza-se o Ambiente de Desenvolvimento Integrado (IDE do termo em inglês

Integrated Development Environment) disponibilizado no site do projeto Arduino gratuitamente. Para se beneficiar do conjunto de bibliotecas também gratuitas disponibilizadas pela comunidade Arduino, procede-se à instalação da EasyVR Shield Library a partir do site oficial da placa EasyVR.

Neste trabalho são comentadas partes do código disponibilizado por Nazir (2013) com as devidas adaptações. O resultado da operação pelo Arduino pode ser verificado pelo Monitor Serial disponível na IDE. Para que o *sketch* apresente mensagens operacionais no monitor, deve-se manter o cabo USB conectado ao computador e ao Arduino e se utilizar comandos de comunicação serial entre os dispositivos no *sketch*. Ao se emitir comandos de voz, o microfone acoplado ao *shield* recebe o sinal sonoro, reconhece o comando e envia para o Arduino o número associado a ele no processo de gravação pelo EasyVR Commander.

As bibliotecas necessárias são incluídas no início do código:

```
#include "Arduino.h"
#include "SoftwareSerial.h"
#include "EasyVR.h"
```

Definire-se quais portas de comunicação serial (12 e 13) são usadas para comunicação entre o *shield* e o Arduino

```
SoftwareSerial port(12,13);
```

e instancia-se o objeto que será usado para a comunicação dos comandos entre o *shield* e o Arduino:

```
EasyVR easyvr(port);
```

em seguida definem-se as variáveis de controle e configuram-se a comunicação serial do Arduino com o computador, os parâmetros do EasyVR, e inicializa-se a variável de controle com os dados do *shield*:

```
int8_t grupo, indice; // define variáveis inteiras de 8 bits

void setup() {
  // configura comunicação serial com o Monitor
  Serial.begin(9600);
  port.begin(9600);
```

```
if (!easyvr.detect()) {  
  Serial.println("EasyVR NÃO detectado!");  
  for (;;) // entra em laço infinito por não poder continuar  
  }  
  
  Serial.println("EasyVR detectado.");  
  
  // define o tempo de timeout em 5 milisegundos  
  easyvr.setTimeout(5);  
  
  // valores possíveis de língua: 0-inglês, 1-italiano, 2-alemão  
  // 3-francês, 4-espanhol, 5-japonês  
  easyvr.setLanguage(0);  
  
  // inicializa grupo com o valor da trigger que habilita o shield  
  // para receber comandos. Valor 1 para a palavra Robô  
  // Para executar comandos, fala-se a trigger seguida do comando  
  grupo = EasyVR::TRIGGER;  
}
```

em seguida define-se a lógica do processo que será executado em um laço infinito no método `loop()` como comentado no código:

```
void loop() {  
  
  // envia mensagem para Monitor Serial para emitir comando  
  Serial.print("Fale um comando do grupo");  
  Serial.println(grupo);  
  easyvr.recognizeCommand(grupo);  
  
  do {  
    // pode executar algum processo enquanto aguarda comando de voz  
  }  
  
  // aguarda finalização do recebimento de comando de voz  
  while (!easyvr.hasFinished());  
  
  // verifica se o comando recebido é uma trigger para receber ação  
  idx = easyvr.getWord();  
  if (idx > 0) {  
    // recebeu trigger diferente da padrão (código 0)  
    // definir bloco de comandos para tratar trigger recebida  
    // no método action()  
    return;  
  }  
  
  // recebeu trigger padrão (0 = ROBO)  
  // obtém comando e retorna seu numero  
  indice = easyvr.getCommand();  
  
  if (indice >= 0) {  
    // cria e inicializa variáveis para reconhecer comandos  
    uint8_t treino = 0;  
    char nome[32];  
  }  
}
```

```

// mostra no Monitor o código do comando recebido
Serial.print("Comando: ");
Serial.print(indice);

// busca na lista de comandos do shield
// qual foi recebido e mostra nome
if (easyvr.dumpCommand(grupo, indice, nome, treino)) {
  Serial.print(" = ");
  Serial.println(nome);
}
else {
  // ocorreu erro
  // soar o sinal de erro de código 0 em volume máximo
  Serial.println();
  easyvr.playSound(0, 31);
  // executar uma rotina de erro a ser definida
  // erro();
}
}
else { // não achou comando ou violou timeout
  // tratar se violou timeout
  if (easyvr.isTimeout()) {
    Serial.println("Esgotou tempo, tente novamente...");
    // obtém código do erro e mostra no Monitor como hexadecimal
    int16_t erro = easyvr.getError();
    if (erro >= 0) {
      Serial.print("Erro ");
      Serial.println(erro, HEX);
    }
  }
  // achou comando, tratar ação do grupo 1 (opção padrão)
  grupo = 1;
}

void action() {

  switch (grupo) {

    // case 0:
    // switch (indice) {
    // inserir código aqui para tratar outras triggers
    // }
    // break; // volta para receber novos comandos

    case 1:
    switch (indice) {
      case 0:
        // tratar ação AVANÇA
        break;
      case 1:
        // tratar ação VOLTA
        break;
      case 2:
        // tratar ação DIREITA
        break;
      case 3:

```

```
// tratar ação ESQUERDA
break;
case 4:
// tratar ação PARE
break;
}
break;
}
}
```

4 Resultados e discussão

A partir dos estudos realizados constatou-se que o reconhecimento de voz, por dispositivos digitais, é complexo especialmente quando se busca o reconhecimento digital de discursos de fala contínua. Entretanto sob o aspecto de reconhecimento de comandos de voz, que são caracterizados por palavras únicas que expressam as ações que devem ser realizadas, a dificuldade é menor e já existe no mercado uma série de dispositivos modulares que abstraem significativamente a aplicação do conjunto *hardware* e *software* para aplicações de automação por comandos de voz.

O *shield* utilizado no estudo exploratório para conhecer o uso de reconhecimento de voz para domótica é um dispositivo que se encontra no mercado brasileiro por, em média, R\$ 160,00 em outubro/2019 e no mercado internacional por R\$85,00.

A facilidade apresentada por esses acessórios está na disponibilização, por licença *open-source* sem reserva de direitos autorais e de patentes, de bibliotecas de *software* para implementação das funcionalidades de reconhecimento de voz nos projetos de automação e controle baseados na plataforma Arduino.

O exemplo tratado na seção 3 se limitou às funcionalidades básicas de um *sketch* para tratamento de 5 comandos, mas as possibilidades de soluções são maiores, especialmente quando se explora as funções mais complexas disponíveis nas bibliotecas fornecidas pelos fabricantes dos *shields* e pela comunidade de desenvolvedores Arduino.

A partir do *software* disponibilizado pelo fabricante do *shield* pode-se configurar comandos de operação de automação e controle residenciais de forma simples e prática, não sendo necessário que se tenha formação específica em eletrônica e computação.

Considerações finais

A humanidade busca, a todo instante, facilitar a realização de suas atividades diárias com o objetivo de aumentar sua qualidade de vida. Como resultado dessa busca o mundo tem passado por mudanças significativas, especialmente aquelas baseadas nas Tecnologias de Informação e Comunicação. Com o aumento exponencial da capacidade de processamento e de armazenamento dos computadores digitais em oposição aos custos de aquisição desses equipamentos que diminuíram significativa e consistentemente, a área de automação e controle tem recebido atenção e investimentos por parte da indústria eletrônica digital favorecendo e simplificando o uso de dispositivos sensores e atuadores em uma ampla variedade de aplicações.

Neste cenário de operação de dispositivos digitais, especialmente os *smartphones*, pesquisa e desenvolvimento de interfaces de operação por voz adquiriram maturidade e produtos de grandes fabricantes disputam os usuários desses dispositivos. Google, Apple, Microsoft, e Amazon oferecem seus assistentes de voz como Google Assistente, Siri, Cortana e Alexa, respectivamente. O desenvolvimento desses produtos disponibilizou uma base de conhecimento que permitiu levar tais recursos às plataformas baseadas em micro controladores para automação de processos.

A Domótica é uma área da indústria de automação e controle que se beneficia desses recursos tendo como complemento de sua utilização aspectos de segurança patrimonial que valorizam as soluções.

Operar dispositivos por voz é uma facilidade que seres humanos buscam, como evidenciado pelo uso cada vez maior de funções de gravação incluídas, por exemplo, em aplicativos de redes sociais. A demanda por esta interface de utilização homem-computador é crescente, e hoje tecnologias de *hardware* e de *software* tornam o desenvolvimento de soluções para essa área mais simples e baratas.

O estudo realizado mostrou que existem várias opções de módulos de *hardware* disponíveis no mercado e uma grande disponibilidade de bibliotecas de *software* para inclusão desses dispositivos em projetos de automação e controle, especialmente para domótica, que viabilizam suas aplicações em situações reais de uso.

Referências

- AFCDUD. **The History of Smart Homes**. 2015. Disponível em: <https://www.afcdud.com/fr/smart-city/422-how-the-history-of-smart-homes.html>. Acesso em: 29.out.2019.
- ARDUINO. **What is Arduino?** Disponível em: <https://www.arduino.cc/>. Acesso em: 28.out.2019.
- BARRAGÁN, H. **The Untold History of Arduino**. 2019. Disponível em: <https://arduinhistory.github.io/>. Acesso em: 30.out.2019.
- BOYD, C.. **Speech Recognition Technology: The Past, Present, and Future**. 2018. Disponível em: <https://medium.com/swlh/the-past-present-and-future-of-speech-recognition-technology-cf13c179aaf>. Acesso em: 28.out.2019.
- ELECHOUSE. **Speak Recognition, Voice Recognition Module V3**. 2019. Disponível em: https://www.elechouse.com/elechouse/index.php?main_page=product_info&products_id=2254. Acesso em: 30.out.2019.
- FORTEBIT. **EasyVR 3 Plus Shield for Arduino**. 2019. Disponível em: <https://fortebit.tech/speech-recognition/>. Acesso em: 29.out.2019.
- FURUI, S. **Speech ans Speaker Recognition Evaluation**. 2007. Disponível em: <https://ccc.inaoep.mx/~villasen/bib/SPEECH%20AND%20SPEAKER%20RECOGNITION%20EVALUATION.pdf>. Acesso em: 28.10.2019.
- KATRE, S. R.; ROJATKAR, D. V. **Home Automation: Past, Present and Future**. 2017. Disponível em: <https://www.irjet.net/archives/V4/i10/IRJET-V4I1061.pdf>. Acesso em: 28.out.2019.
- KICKSTARTER. **ReSpeaker - Add Voice Control Extension To Anything You Like**. 2017. Disponível em: <https://www.kickstarter.com/projects/seeed/respeaker-an-open-modular-voice-interface-to-hack>. Acesso em: 28.out.2019.
- KINCAID, J. **A Brief History of ASR: Automatic Speech Recognition**. 2018. Disponível em: <https://medium.com/descript/a-brief-history-of-asr-automatic-speech-recognition-b8f338d4c0e5>. Acesso em: 30.out.2019.
- LUETH, K. L. **Why the Internet of Things is called Internet of Things: Definition, history, disambiguation**. 2014. Disponível em: <https://iot-analytics.com/internet-of-things-definition/>. Acesso em: 29.out.2019.
- MAKER.IO. **How To Set Up EasyVR Voice Shield on Arduino**. 2019. Disponível em: https://www.digkey.com/en/maker/blogs/2019/how-to-set-up-easyvr-voice-shield-on-arduino?utm_adgroup=General&slid=&gclid=CjwKCAjwxt_tBRAXEiwAENY8hRcsd7jjKjaJ3NLXjk_OUFimcGOTTm2vZEKVDUxW77GJReYGA7n-sBoCfWEQAvD_BwE. Acesso em: 29.out.2019.

NASIR, S. Z. **Voice Recognition Project Using EasyVR Shield**. 2013. Disponível em: <https://www.theengineeringprojects.com/2013/04/voice-recognition-project-using-easyvr.html>. Acesso em: 29.out.2019.

PICONE, J. **Fundamentals os Speech Recognition: a short course**. 1996. Disponível em: http://www.iitg.ac.in/samudravijaya/tutorials/fundamentalOfASR_picone96.pdf. Acesso em: 28.out.2019.

RANI, P.J.; *et al.* **Voice Controlled Home Automation System Using Natural Language Processing (NLP) and Internet of Things (IoT)**. 2017. Disponível em: https://www.academia.edu/36236821/VOICE_CONTROLLED_HOME_AUTOMATION_SYSTEM_USING_NATURAL_LANGUAGE_PROCESSING_NLP_AND_INTERNET_OF_THINGS_IoT. Acesso em: 29.out.2019

ROUSE, M. **Voice Recognition (Speaker Recognition)**. 2018. Disponível em: <https://searchcustomerexperience.techtarget.com/definition/voice-recognition-speaker-recognition>. Acesso em: 29.out.2019.

SEED. **Grove - Speech Recognizer**. 2019a. Disponível em: http://wiki.seeedstudio.com/Grove-Speech_Recognizer/. Acesso em: 28.out.2019.

_____. **ReSpeaker Core**. 2019b. Disponível em: http://wiki.seeedstudio.com/ReSpeaker_Core/. Acesso em: 30.out.2019.

SINGH, N. **A Study on Speech and Speaker Recognition Technology and its Challenges**. 2014. Disponível em: https://www.researchgate.net/publication/308888698_A_Study_on_Speech_and_Speaker_Recognition_Technology_and_its_Challenges. Acesso em: 29.out.2019.

SULTAN, M.; NABIL, K. **Smart to Smarter: Smart Home Systems History, Future and Challenges**. Disponível em: <https://pdfs.semanticscholar.org/18c1/b2bcb167e4f52d5c1dddfecbff6881d4357b.pdf>. Acesso em: 29.out.2019

VATANSEVER, A.; KUSCU, H.; TUNA, G. **Voice Controlled Home Automation Design**. 2018. Disponível em: <https://dergipark.org.tr/tr/download/article-file/528317>. Acesso em: 28.out.2019.

VELDE, N. V. D. **Speech Recognition Technology Overview**. 2019. Disponível em: <https://www.globalme.net/blog/the-present-future-of-speech-recognition>. Acesso em: 30.out.2019.

YOUTUBE. **Arduino: The Documentary**. 2010. Disponível em: <https://www.youtube.com/watch?v=kQXeMCLJtwQ>. Acesso em: 28.out.2019.

YUAN, M. **Getting to know NodeMCU and its DEVKIT board**. 2017. Disponível em: <https://developer.ibm.com/tutorials/iot-nodemcu-open-why-use/>. Acesso em: 28.out.2019.