

**UNIVERSIDADE FEDERAL DE VIÇOSA
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA**

TIAGO OGIONI COSTALONGA

**INTERFACE HOMEM-MÁQUINA UTILIZANDO
SENSOR KINECT PARA CONTROLE DE
MOVIMENTO DE UM QUADROTOR**

**VIÇOSA
2014**

TIAGO OGIONI COSTALONGA

**INTERFACE HOMEM-MÁQUINA UTILIZANDO
SENSOR KINECT PARA CONTROLE DE
MOVIMENTO DE UM QUADROTOR**

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 490 - Monografia e Seminário e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Orientador: Prof. Dr. Alexandre Santos Brandão

VIÇOSA
2014

TIAGO OGIONI COSTALONGA

**INTERFACE HOMEM-MÁQUINA UTILIZANDO SENSOR KINECT
PARA CONTROLE DE MOVIMENTO DE UM QUADROTOR**

Monografia apresentada ao Departamento de Engenharia Elétrica do Centro de Ciências Exatas e Tecnológicas da Universidade Federal de Viçosa, para a obtenção dos créditos da disciplina ELT 490 - Monografia e Seminário e cumprimento do requisito parcial para obtenção do grau de Bacharel em Engenharia Elétrica.

Aprovada em 25 de abril de 2014.

COMISSÃO EXAMINADORA

Prof. Dr. Alexandre Santos Brandão - Orientador
Universidade Federal de Viçosa

Profa Dra. Ketia Soares Moreira - Membro
Universidade Federal de Viçosa

Prof. Dr. Monael Pinheiro Ribeiro - Membro
Universidade Federal de Viçosa

À Deus e Nossa Senhora, aos meus pais, à minha família e aos meus amigos.

Agradecimentos

Primeiramente, agradeço à Deus e Nossa Senhora por sempre estarem comigo nos bons e maus momentos, me consolando e me dando oportunidades únicas em toda minha caminhada até aqui.

Agradecimento especial à minha mãe, por sempre me mostrar o caminho certo a seguir, mesmo que muitas vezes eu não concordasse, e ao meu pai que me ensinou sempre a viver com dignidade e honestidade nesse mundo e por nunca ter esquecido de depositar o dinheiro nos dias certos. E aos dois, agradeço por sempre estarem do meu lado, me dando tudo o que vocês não tiveram e, sem sombra de dúvidas, tudo o que sonharam para mim.

À minha grande família, que se fosse citar todos, não sobraria espaço para minha monografia, agradeço imensamente o apoio e carinho concedidos, mesmo à distância estávamos juntos. Vocês foram essenciais nessa caminhada.

À toda turma da Engenharia Elétrica 2008 e àqueles que de uma forma ou de outra se integraram nessa turma, agradeço às festas e resenhas que tivemos. Tenho certeza que escrevemos uma página aqui em Viçosa.

Ao orientador Prof. Dr. Alexandre Santos Brandão, pela ajuda, ensinamento e paciência na elaboração dos nossos artigos e desta monografia e pelas soluções milagrosas no MATLAB que, diga-se de passagem, ele resolvia apenas com uma simples observação.

A CNPq/PIBITI pelo apoio financeiro concedido, sem o qual esta pesquisa não teria sido possível.

E, por fim, agradeço à todas as pessoas que fizeram e fazem parte dessa caminhada e da minha vida.

Muito obrigado,

Tiago Ogioni Costalonga

“De pé sem cair e deitado sem dormir”

Valter Costalonga

Resumo

O objetivo desse trabalho é apresentar uma nova e intuitiva forma de controle de um quadrotor, em específico um quadrotor AR Drone Parrot 1.0, através de uma interface homem-máquina (IHM), baseada em gestos. Em outras palavras, os próprios movimentos do usuário são utilizados para comandar o veículo aéreo não tripulado (VANT). Deste modo, o usuário se torna o “controlador”. Como objetivos específicos, tem-se: (1) extração de informações espaciais de partes específicas do corpo; (2) reconhecimento de postura do usuário a partir dessas informações; (3) associar os movimentos realizados aos comandos específicos que serão enviados ao quadrotor. O sensor de movimentos Kinect da Microsoft será utilizado para possibilitar o reconhecimento de gestos e, portanto, controlar um VANT. Este sensor é mais viável economicamente se comparado a outros sensores com a mesma finalidade e funciona como dispositivo de rastreamento das posições do usuário, através de duas câmeras - a de cor (RGB) e a de infravermelho. Após a captura das imagens em profundidade pelo Kinect, o programa identifica o movimento feito pelo usuário, faz o reconhecimento desse gesto e interage com o periférico de entrada e, então, o comando é transferido para o VANT. Portanto, obteve-se os dados de posição espacial tridimensional da mão direita, mão esquerda, cotovelo esquerdo, bem como dos ângulos formados pelos mesmos. A partir dos dados da mão direita foi feita a primeira derivada temporal desta, o que gerava os dados de intenção de movimento, que eram convertidos em dados que pudessem ser enviados ao quadrotor. Pelos dados da mão esquerda era possível controlar a rotação e o comando de pausar do quadrotor. Com isso, o quadrotor recebia esses dados e se movimentava de acordo com a intenção de movimento da mão direita e/ou esquerda do usuário. Como forma de avaliar o desempenho da interface, foi comparado os gráficos de entrada e de resposta do quadrotor obtendo como resultados uma tendência igual e esperada das curvas analisadas. Assim neste trabalho o funcionamento e a aplicação pretendidas foram alcançadas durante todo o processo prático.

Abstract

The aim of this paper is to present a new and intuitive way to control a quadrotor, in specific a quadrotor Parrot AR Drone 1.0, through a human-machine interface (HMI), based on gestures, ie, using their own user's movements to command the unmanned aerial vehicle (UAV). Thus the user becomes the "controller". As specific objectives, it has: (1) extraction of spatial information of specific parts of the body; (2) recognition of the user's posture from these informations; (3) to associate the movements performed to specific commands that will be sent to the quadrotor. To enable gesture recognition and, thus, control a UAV, will be used the motion sensor Microsoft Kinect which is more economically viable compared to other sensors with the same purpose and works as a tracking device of user's positions, since it has two cameras - the color (RGB) and infrared. After capturing the images in depth by Kinect, the algorithm identifies the movement made by the user, do the recognition of this gesture and interact with the input device and, then, the command is transferred to the UAV. Therefore, it was obtained the three-dimensional spatial position's data of the right hand, left hand, left elbow and the angles formed by them. From the data of the right hand was taken the first time derivative thereof, which generated the data of intended movement, which were converted into data that can be sent to quadrotor. By the data of the left hand was possible to control the rotation and the command to pause the quadrotor. Thus, the quadrotor receives these data and moves according to the intention of movement of the right and/or left hand of the user. In order to evaluate the performance of the interface, it was compared the graphics of the quadrotor's input and response getting as results a trend equal and expected of the curves analyzed. So in this paper the operation and intended application were achieved during all the practical process.

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	10
1.1	O Sensor Kinect	12
1.2	Quadrotor ArDrone Parrot	16
1.3	Objetivos	17
2	Materiais e Métodos	19
2.1	Descrição dos Comandos Realizados por Gestos Corporais	20
2.1.1	Comando de Decolagem	21
2.1.2	Comando de Rotação	22
2.1.3	Comando Direcional	22
2.1.4	Comando de Parar/Pausar	23
2.2	Descrição do Método de Controle do Movimento	23
3	Resultados e Discussões	25
3.1	Movimento de Rolagem	25
3.2	Movimento de Arfagem	26
3.3	Movimento Descendente/Ascendente	26
4	Conclusões	31
	Referências	32

Lista de Figuras

1	Aplicações de uma Interface Natural do Usuário.	11
2	O sensor Kinect	12
3	Componentes principais do Kinect	12
4	Exemplo de luz estruturada, com as distorções é possível calcular a profundidade	13
5	Luz Estruturada emitida pelo Kinect	14
6	Funcionamento da captura de uma imagem de profundidade feita pelo Kinect	15
7	Exemplo da utilização de marcadores especiais	15
8	Exemplos de veículos aéreos não tripulados (VANT's).	16
9	Rastreamento do esqueleto do usuário.	20
10	Ângulos de controle do quadrotor.	21
11	Método de Comando. As setas vermelhas atuam apenas uma vez enquanto as setas azuis indicam atuação cíclica.	21
12	Gestos para controle do quadrotor.	22
13	Sistemas de referência.	24
14	Controle do movimento de rolagem do quadrotor.	28
15	Controle de movimento de arfagem do quadrotor.	29
16	Controle de movimento de descendente e ascendente do quadrotor.	30

Lista de Tabelas

1	Correspondência entre os gestos do usuário e os comandos do ArDrone . . .	20
---	---	----

1 *Introdução*

Com o constante avanço da tecnologia nas últimas décadas, grande parte da população mundial passou a ter um contato direto com sistemas computacionais, aumentando o espectro de usuários que, agora, não se restringem apenas aos técnicos em Eletrônica ou Informática, mas incluem, também, pessoas inseridas em qualquer área de conhecimento e de variados graus de conhecimento tecnológico. Nesse contexto, surgiu-se a necessidade de criar sistemas que atendessem às necessidades e capacidades dos indivíduos de forma simples e intuitiva.

A possibilidade de uma interface de controle natural, intuitiva e expressiva de plataformas móveis abre uma grande variedade de aplicações que vão desde entretenimento até segurança [1],[2]. Desse modo, uma Interface Homem-Máquina (IHM) pode ser entendida como um conjunto de processos, diálogos e ações por meio dos quais um ser humano utiliza e interage com um computador [3].

A partir dos anos 70, as IHMs tornaram-se uma preocupação no desenvolvimento de novas tecnologias. Desde esta época, entende-se como Interação (ou interface) Homem-Máquina a maneira na qual o usuário interage com o sistema computacional, ou seja, a IHM consiste em realizar tarefas através da colaboração entre o humano e o componente de automação [4]. Como citado anteriormente, é importante que a IHM tenha funcionalidade e usabilidade, visando atender as necessidades e se acomodar as capacidades das pessoas a quem eles são endereçados [5]. Assim sendo, para o desenvolvimento de uma interface, é fundamental uma análise física e mental do ser humano, buscando entender como ele irá reagir diante da interface. Fisicamente, o usuário se utiliza do toque, da visão e da audição para interagir com o meio, além disso, ele mentalmente utiliza do raciocínio, percepção, memória e linguagem (fatores cognitivos) [6].

Dentre os vários tipos de uma IHM, existe uma interface que pode ser controlada através de movimentos corporais realizados pelo usuário. Esse método de controle, conhecido como Interface Natural do Usuário (INU), associa várias modalidades de comunicação

através de interações físicas humanas, tendo como foco as habilidades das pessoas tais como toque, voz, visão e funções cognitivas mais complexas [7]. Como exemplos de aplicação de INU pode-se citar o desenvolvimento de um sistema para TV interativa em que a mão e os dedos do usuário são rastreados por uma câmera o que permite, assim, controlar os canais e volume da televisão através de comandos gestuais da mão [8]. Tem-se também a possibilidade de manipulação de um robô projetor que é posicionado através de um gesto pré-determinado com as mãos e a direção da projeção pode ser controlada através de um gesto com o braço do usuário que aponta para a superfície na qual será projetada a imagem, outras funções como “zoom in/out” também são disponíveis [9]. Outro exemplo de uma INU é a implementação de uma interface “multi-touch” para planejamento urbano colaborativo, isto é, o desenvolvimento de uma tela sensível a variados toques simultâneos de modo que cada tipo de toque corresponde a uma função da interface tais como: zoom, rotação, movimento vertical/horizontal, entre outros [10]. Os dois últimos exemplos podem ser vistos na Figura 1. Todas essas aplicações mostram uma nova forma de interação com sistemas eletrônicos visando criar uma interface mais intuitiva e natural para os usuários.



(a) Robô projetor.

(b) Tela “multi-touch”.

Figura 1: Aplicações de uma Interface Natural do Usuário.

Deste modo, o uso de gestos para controle de plataformas móveis, comparado às interfaces tradicionais pode oferecer inúmeras oportunidades de interação [11]. Novas soluções baseada em gestos tem sido largamente introduzidas em vários cenários de interação como, por exemplo, a navegação de conteúdos multimídia e de mundos virtuais. Segundo [12], a construção de sistemas baseados em gestos terá um importante papel nas futuras tendências da IHM.

1.1 O Sensor Kinect

Conhecido, anteriormente ao seu lançamento em 2009, por Projeto Natal, o sensor Kinect, mostrado na Figura 2, foi apresentado pela empresa Microsoft como um acessório do console de jogos XBOX360 e ganhou o gosto de um público muito diversificado, que engloba os jogadores no campo de entretenimento e os pesquisadores nos campos tecnológicos.



Figura 2: O sensor Kinect

No mundo dos vídeos-game, o Kinect revolucionou a maneira como o jogador interage com os jogos. Sua tecnologia permite que se elimine completamente o uso dos joysticks uma vez que o controlador passa a ser o corpo do usuário, ou seja, os comandos para os jogos são dados por gestos ou posturas corporais do jogador. Para o mundo científico, o aparelho inovou a interação com as máquinas e computadores, levando o aparelho desde um laboratório até uma sala de cirurgia [13].

Em sua estrutura mostrada na Figura 3 e detalhadas em [14], o Kinect apresenta as seguintes especificações:



Figura 3: Componentes principais do Kinect

- Uma câmera RGB comum, como uma webcam, com uma resolução de até 640x480 pixels a uma taxa de 30 fps, é responsável pela captura da imagem colorida do ambiente;

- Um emissor de infravermelho e um sensor de profundidade também infravermelho baseados na tecnologia Light Coding, da empresa israelense PrimeSense, com uma resolução de 320x240 pixels a uma taxa de 30 fps, os quais possibilitam a medida de profundidade da cena;
- Quatro microfones para captura de sons, os quais permitem a gravação do audio do ambiente e a determinação da localização da fonte de som e a direção da onda sonora;
- Um motor na base do sensor para inclinação do mesmo.

A tecnologia Light Coding citada anteriormente, talvez seja a responsável pelo sucesso científico do sensor. Seu funcionamento é baseado em luz estruturada, ou seja, uma luz com uma estrutura bem definida (como, por exemplo, quadriculado, linhas horizontais ou verticais) que permite a obtenção de informações sobre uma determinada cena baseado na forma como essa luz é refletida. A Figura 4 ilustra um exemplo de uma luz estruturada projetada em uma superfície. Note que é possível retirar informações de profundidade acerca do tamanho (reflexão da luz mais próxima é maior que da luz mais distante) ou distorções da luz (a luz se distorce em superfícies irregulares) [15, 16].



Figura 4: Exemplo de luz estruturada, com as distorções é possível calcular a profundidade

A maioria dos sensores usados para medição de profundidade utilizam, ao invés da luz estruturada, as chamadas câmeras “time-of-flight” (TOF) cuja função é, basicamente, medir o tempo em que a luz deixa essa câmera, reflete em um objeto e é novamente captada pela câmera, desse modo é possível obter a distância em que o objeto se encontra da mesma. Porém essa tecnologia exige a utilização de um software e hardware dedicados que medem o tempo de cada pixel emitido e recebido o que faz o custo ser elevado. Por

exemplo, a empresa suíça Mesa Imaging produz uma linha de câmeras TOF voltadas para pesquisa científica há vários anos, chamada SwissRanger. O melhor modelo é a SwissRanger SR4000, que custa 9000 dólares. Ela tem uma resolução de 176x144 e opera entre 30 e 50 frames por segundo.

Já o Kinect, diferentemente das câmeras TOF, utiliza luz estruturada que é projetada com luz infravermelha (uma nuvem de pontos), mostrada na Figura 5, invisível à olho nu e que não necessita de iluminação externa, ou seja, funciona no escuro. Como qualquer câmera digital é capaz de detectar infravermelho, então é dispensável o uso de um hardware ou software dedicado de alto custo. Assim pode-se obter os dados de profundidade do ambiente de maneira “simples” e com baixo custo, porém seu uso fica restrito a ambientes internos e pouco iluminados, uma vez que sofre efeitos da radiação emitida pelo sol.

Essa tecnologia, juntamente com a câmera comum RGB, permite ao Kinect enviar informações espaciais tridimensionais de determinado objeto ou pessoa, como mostra a Figura 6. Deste modo, o sensor ainda apresenta uma ferramenta extremamente valiosa: o rastreamento do usuário em tempo real.

No caso específico de reconhecer uma pessoa em uma cena, é necessário conhecer quais dados de uma imagem tridimensional corresponde à pessoa em um ambiente. Deve-se levar em consideração que um ambiente não é composto apenas por indivíduos, mas também por objetos, móveis e/ou animais. Neste caso, o problema não é confundir o objeto com o fundo, já que informações de profundidade são conhecidas, mas sim perceber que um conjunto de dados é uma pessoa em determinada posição.

Assim, para realizar o rastreamento de uma pessoa, a Microsoft usou técnicas de



Figura 5: Luz Estruturada emitida pelo Kinect

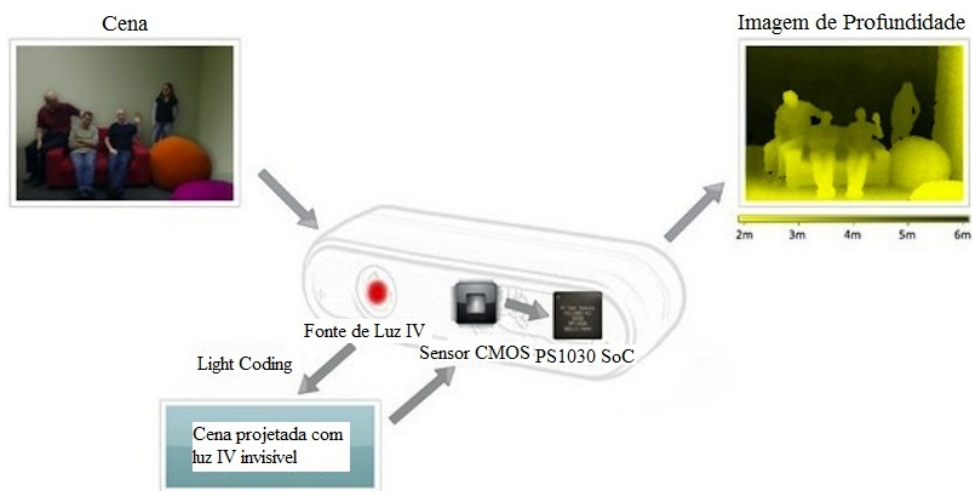


Figura 6: Funcionamento da captura de uma imagem de profundidade feita pelo Kinect

aprendizado de máquina, ou seja, ao contrário de escrever um programa que resolve o problema, escreveu-se um programa que aprende a resolver o problema. Faz-se uma filmagem utilizando marcadores especiais, como exemplificado pela Figura 7, para detectar a posição corporal e, simultaneamente, é feito o mesmo vídeo utilizando o Kinect. Os dois dados obtidos são fornecidos ao algoritmo de aprendizado. Após utilizar a técnica com vários exemplos e combinando com algumas observações como, por exemplo, o tamanho constante das partes do corpo de uma pessoa, abertura entre os joelhos e cotovelos e a análise de captura de antigos jogos, chegou-se a um programa capaz de detectar a posição do corpo a partir da imagem tridimensional vinda do Kinect, independente do tamanho da pessoa, da posição e ação da mesma [17].

Nesse contexto, percebeu-se a grande possibilidade de utilizar uma tecnologia de baixo custo porém de satisfatório desempenho para o desenvolvimento de pesquisa, projetos e



Figura 7: Exemplo da utilização de marcadores especiais

tecnologia em outras áreas que não fossem o de desenvolvimento de jogos [13].

1.2 Quadrotor ArDrone Parrot

Muitas tarefas cotidianas podem ser realizadas por robôs. Em especial, os veículos aéreos não tripulados (VANTs), como aqueles mostrados na Figura 8, que têm atraído a atenção da comunidade científica e industrial devido às aplicações e aos desenvolvimentos tanto militares quanto civis [18].

Existem VANTs de pás rotativas, como helicópteros e quadrotores, e os de asas fixas,



(a) Mini-helicóptero.



(b) ArDrone.



(c) Drone Israelense.

Figura 8: Exemplos de veículos aéreos não tripulados (VANT's).

como aviões e planadores. As aeronaves de pás rotativas são mais versáteis se comparados com os de asa fixa, devido a sua capacidade de decolar, planar e aterrissar em pequenos ambientes, como laboratórios, galpões, dentre outros [19].

Os VANTs de pás rotativas constituem umas das mais complexas máquinas voadoras devido a sua habilidade de decolar e aterrissar verticalmente, realizar um voo pairado, mover-se adiante (arfagem) e lateralmente (rolagem), enquanto se mantém em uma mesma cota vertical, além da possibilidade de poder mudar completamente sua direção de voo e de deter seu movimento abruptamente [20, 21].

Os quadrotores são mais estáveis e apresentam todos os benefícios obtidos pelos helicópteros, porém apresentam uma baixa capacidade de carga em voo [22], quando comparados a helicópteros de dimensões similares.

Existem várias aplicações nas quais um quadrotor é usado, como, por exemplo, em auxílio para operações de busca e salvamento, recolhimento de informações para previsão meteorológica, vigilância de ambientes, dentre outras [23].

O quadrotor ArDrone Parrot foi, originalmente, fabricado para jogos de realidade aumentada, entretanto seu conjunto de sensores e uma interface de software chamou a atenção de universidades e instituições de pesquisa. Ele possui uma estrutura de suporte em fibra de carbono, corpo plástico, quatro motores de alta eficiência, duas câmeras, sensores e placa de controle [24].

Os sensores presentes na versão 1.0 do quadrotor são compostos por: 1) Transmissor e receptor de ultrassom para medir a altura do VANT; 2) Giroscópio para obtenção da orientação do drone; 3) Acelerômetro para auxiliar a medida de orientação da aeronave. Além dos sensores citados, a comunicação entre o quadrotor ArDrone Parrot 1.0 e um computador é feita através de uma conexão USB e/ou Wi-Fi 802.11b/g. O tempo de amostragem do drone é de, aproximadamente, 33 ms [24].

1.3 **Objetivos**

Este trabalho tem como objetivo geral criar uma IHM que permita ao usuário controlar um veículo aéreo não tripulado (VANT), em específico um quadrotor ArDrone Parrot 1.0, utilizando gestos. Em outras palavras, o usuário é o “controlador”. Como objetivos específicos, tem-se: (1) extração de informações espaciais de partes específicas do corpo; (2) reconhecimento de postura do usuário a partir dessas informações; (3) associar os

movimentos realizados aos comandos específicos que serão enviados ao quadrotor.

O uso do Kinect aplicado ao controle de quadrotores é relativamente novo, tem-se exemplos os trabalhos apresentados em [25, 26, 27]. Em especial, o instituto ETH Zurique tem feito pesquisas com controle e comando desses veículos. O método utilizado por eles propõe uma interação dinâmica com quadrotor baseado na posição do braço, ou seja, o mapa tridimensional do braço do usuário é mapeado para as coordenadas locais (também tridimensional) do quadrotor. Entretanto, a proposta deste trabalho é diferente, em termos do método de controle, apesar de usar gestos para efetuar os comandos ao quadrotor, o método consiste em mapear partes específicas do corpo e aplicar o conceito de derivada (intenção de movimento), de modo a realizar o movimento do quadrotor.

2 *Materiais e Métodos*

Para possibilitar o reconhecimento de gestos e, portanto, controlar um VANT, utilizou-se o sensor de movimentos Kinect da Microsoft que é mais viável economicamente se comparado a outros sensores com a mesma finalidade e funciona como dispositivo de rastreamento das posições do usuário [28]. Tal sensor possui duas câmeras: a de cor (RGB) e a de infravermelho, detalhes podem ser encontrados em [29]. Na presente proposta, após a captura das imagens em profundidade pelo Kinect, o movimento feito pelo usuário é identificado, o gesto realizado é reconhecido e a interação com o VANT é realizada.

Inicialmente, o esqueleto do usuário é identificado através do Kinect. A Figura 9 indica os pontos específicos do corpo a serem rastreados. Entre os pontos selecionados estão as posições relativas à mão direita e esquerda, dadas em coordenadas cartesianas, e os ângulos β (formado pelos pontos da mão, cotovelo e ombro esquerdo) e α (formado pela mão e ombro esquerdo e um ponto qualquer que tenha o mesmo valor da coordenada x do ombro esquerdo). A Figura 10 ilustra os ângulos citados, os quais são necessário na etapa de controle do quadrotor. Vale destacar que a imagem capturada pelo sensor kinect não é apresentada como aquela tirada por uma câmera convencional, e sim espelhada, como pode ser visto nas Figuras 9 e 10.

A Figura 11 descreve a metodologia adotada para controle do quadrotor através dos gestos a serem detectados através do sensor kinect. Cada uma das etapas é descrita de forma sucinta na sequência:

Inicialização do Kinect e rastreamento do esqueleto: esta etapa inicializa o Kinect e a partir de seus sensores realiza a identificação do esqueleto do usuário, para realização do rastreamento;

Identificação e aquisição de dados do esqueleto do usuário: uma vez detectado o usuário, as informações pertinentes a cada parte do corpo são guardadas e analisadas a cada frame;

Instruções para decolagem e controle: uma série de instruções é apresentada ao usuário para que ele possa realizar os comandos iniciais de controle;

Controle de movimento e direção de voo: o algoritmo de comando, através da posição da mão esquerda do usuário e da derivada temporal da posição da mão direita do usuário, realiza o controle do movimento do quadrotor.

2.1 Descrição dos Comandos Realizados por Gestos Corporais

Este projeto focou no comando de decolagem, comando de rotação, comando direcional e no comando de pausar, mostrados na Tabela 1 e descritos nas próximas subseções.

Tabela 1: Correspondência entre os gestos do usuário e os comandos do ArDrone

Gesto	Comando
Mão direita para cima	Subir
Mão direita para baixo	Descer
Mão direita para esquerda	Ir para esquerda
Mão direita para direita	Ir para direita
Mão direita para frente	Ir para frente
Mão direita para trás	Ir para trás
Mão esquerda para cima e $\beta < 90^\circ$	Decolagem
$35^\circ < \alpha < 50^\circ$	Pause
Mão esquerda para esquerda	Guinada anti-horária
Mão esquerda para direita	Guinada horária



Figura 9: Rastreamento do esqueleto do usuário.

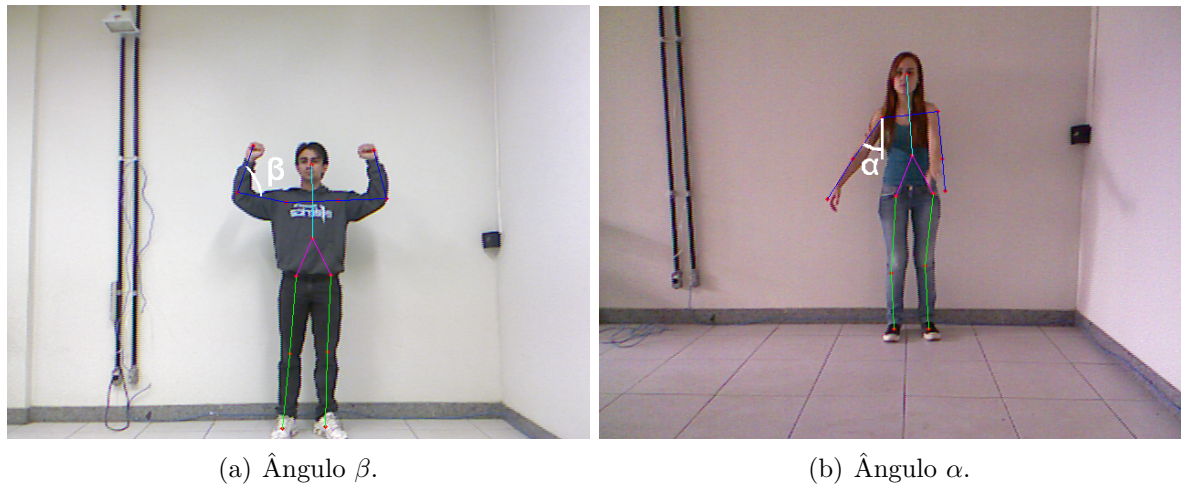


Figura 10: Ângulos de controle do quadrotor.

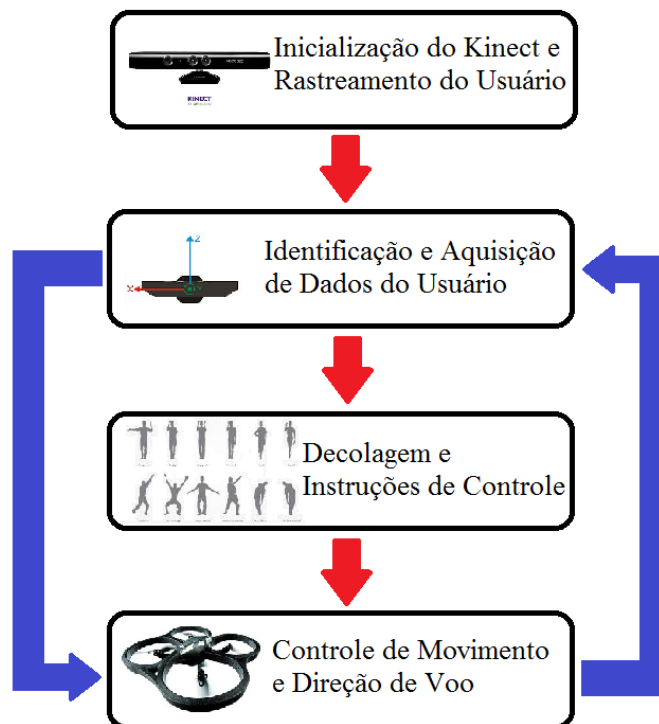


Figura 11: Método de Comando. As setas vermelhas atuam apenas uma vez enquanto as setas azuis indicam atuação cíclica.

2.1.1 Comando de Decolagem

Após o usuário ter sido reconhecido e seu esqueleto já estar sendo rastreado pelo Kinect, é feito o comando de decolagem que consiste em levantar a mão esquerda, mantendo o cotovelo esquerdo estático e a mão direita abaixo da linha do quadril, como mostra a Figura 12(a). Ao realizar este movimento o sistema envia o comando de decolagem ao quadrotor, o qual realiza a decolagem e aguarda novos comandos.

2.1.2 Comando de Rotação

Após a decolagem, o usuário tem a possibilidade de rotacionar o quadrotor em torno de seu eixo normal, para isso a mão esquerda deve estar acima do cotovelo esquerdo e a mão direita deve estar abaixo da linha do quadril esquerdo (ver Figura 12(b)). Desse modo ao se movimentar a mão esquerda para o lado direito ou esquerdo o quadrotor rotaciona para o sentido horário ou anti-horário com uma velocidade constante predeterminada.

2.1.3 Comando Direcional

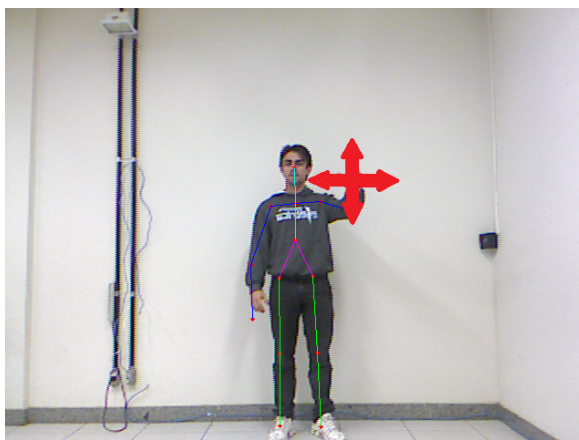
Também após a decolagem, o usuário, agora com sua mão direita acima da linha do quadril, conforme ilustrado na Figura 12(c), pode comandar o quadrotor para frente e para trás, para direita e para esquerda e para cima e para baixo de acordo com o movimento da sua mão direita.



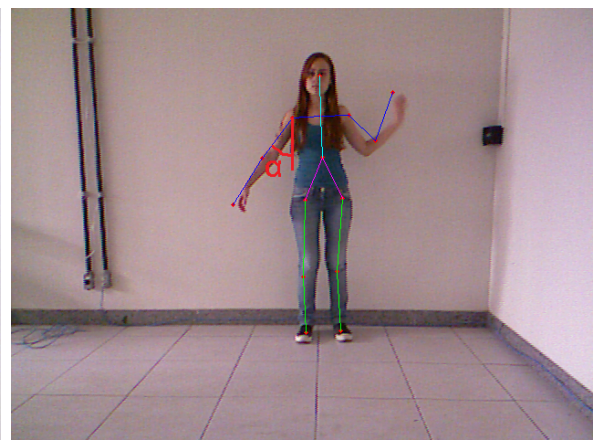
(a) Comando de decolagem.



(b) Comando de rotação.



(c) Comando direcional.



(d) Comando de pausar.

Figura 12: Gestos para controle do quadrotor.

2.1.4 Comando de Parar/Pausar

Depois de realizada a decolagem, o usuário fica livre para comandar o quadrotor, porém se fez necessário a utilização de um comando “parar”, como mostrado na Figura 12(d). Este comando faz com que todos os outros comandos sejam desabilitados, deixando o quadrotor em modo de voo estacionário, não podendo se mover em nenhum sentido mesmo que um comando direcional seja dado.

2.2 Descrição do Método de Controle do Movimento

Para realizar controle do movimento do quadrotor, foi calculada a primeira derivada temporal da posição tridimensional (em x , y e z) da mão direita, dada por

$${}^k\dot{x} = \frac{dx_{md}}{dt}, \quad (2.1)$$

$${}^k\dot{y} = \frac{dy_{md}}{dt}, \quad (2.2)$$

$${}^k\dot{z} = \frac{dz_{md}}{dt}. \quad (2.3)$$

segundo o referencial mostrado na Figura 13(a). Vale mencionar que o índice k corresponde aos valores relacionados ao referencial do Kinect e o índice md denota mão direita.

O referencial do quadrotor ilustrado na Figura 13(b) é distinto daquele associado ao Kinect, mostrado na Figura 13(a). Logo, faz-se necessário adequar os sinais de controle de movimento obtidos no referencial do Kinect para aqueles relacionados à atuação do quadrotor. Sendo assim, tem-se que

$${}^q\phi = K_{x\phi} {}^k\dot{x}, \quad (2.4)$$

relaciona a variação no eixo x do referencial $\langle k \rangle$ e o ângulo de rolagem do quadrotor,

$${}^q\theta = K_{z\theta} {}^k\dot{z}, \quad (2.5)$$

relaciona a variação no eixo z do referencial $\langle k \rangle$ e o ângulo de arfagem do quadrotor, e

$${}^q\dot{z} = K_{yz} {}^k\dot{y}, \quad (2.6)$$

relaciona a variação no eixo y do referencial $\langle k \rangle$ e a velocidade ascendente/descendente vertical do quadrotor. Com essas relações é possível realizar uma comparação entre o comando enviado ao quadrotor (entrada) e a sua ação desenvolvida (saída). Note que K_{ij} são constantes de proporcionalidade.

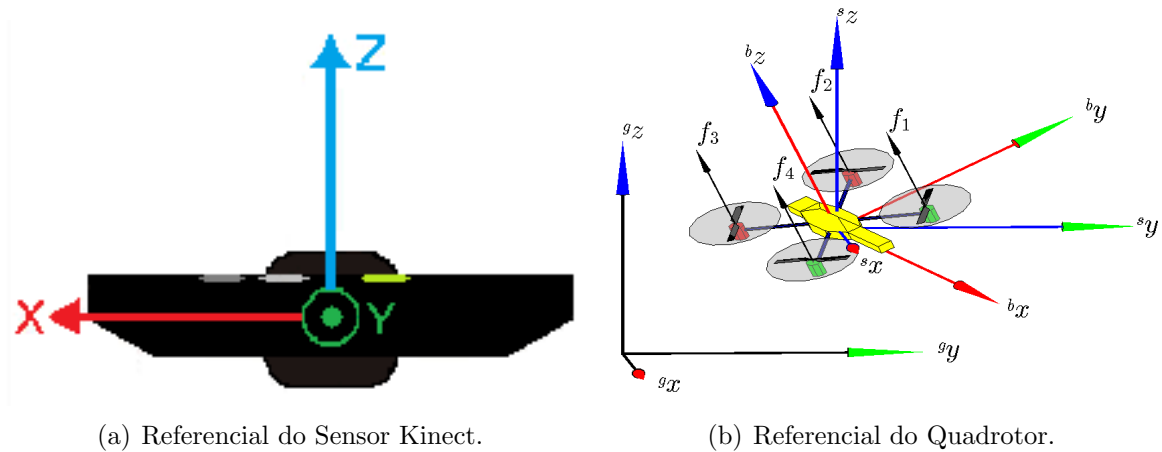


Figura 13: Sistemas de referência.

3 *Resultados e Discussões*

Nesta seção será feito um estudo de desempenho de resposta do quadrotor tomando os referenciais do Kinect e do quadrotor.

Os resultados experimentais foram obtidos através de três sub-testes. No primeiro, comandos para o movimento de rolagem foram enviados ao quadrotor. No segundo, comandos para o movimento de arfagem foram dados ao VANT. Por fim, no terceiro teste, comandos para movimento ascendente/descendente foram registrados para o deslocamento da aeronave. As seções a seguir destacam os resultados obtidos em cada um desses testes e as discussões pertinentes.

3.1 Movimento de Rolagem

A Figura 14(a) mostra o deslocamento lateral (direita/esquerda) da mão direita do usuário. Os instantes de tempo inferiores a zero indicam a etapa de identificação e reconhecimento do usuário. A partir do instante $t = 0$, a aeronave está apta a receber um comando.

A Figura 14(b) apresenta uma comparação entre o sinal enviado ao quadrotor (${}^k\dot{x}$) e o sinal de saída do mesmo (${}^q\phi$). Como dito anteriormente, para $t < 0$, o ArDrone não recebe nenhum comando direcional, daí pode-se visualizar que a partir de $t = 0$, há um acompanhamento dos dois sinais indicando que o VANT seguiu a tendência do comando. Note que o sinal de saída (quadrotor) possui sobrepassos em relação ao sinal de entrada, para $t > 0$, devido à inércia de seu movimento. É possível visualizar que o quadrotor segue a tendência de movimento solicitada pelo usuário. O erro médio quadrático entre os sinais de entrada e saída é de 9.1% com desvio padrão de 0.301.

A Figura 14(c) ilustra a evolução da velocidade do quadrotor em seu referencial. Percebe-se que sua velocidade também é controlada pelo usuário, de forma que a velocidade varia de acordo com a velocidade na qual o gesto é realizado.

3.2 Movimento de Arfagem

A Figura 15(a) mostra o deslocamento “para frente e para trás” da mão direita do usuário. Novamente para $t < 0$ tem-se o reconhecimento do usuário e comando de decolagem, de modo que o quadrotor só recebe comando direcional para $t \geq 0$.

Assim como na Figura 14(b), também foi feita uma comparação entre o sinal de entrada e o sinal de saída do quadrotor. Porém, agora, foram associados os comandos em $^k\dot{z}$ e as informações de arfagem. A Figura 15(b) ilustra os resultados obtidos neste teste experimental. Percebe-se também que para $t \geq 0$ o quadrotor começa a seguir o sinal de entrada (usuário), novamente, sua inércia de movimento leva o quadrotor a obter valores de pico maiores que o sinal de entrada, porém nota-se um desempenho também satisfatório em relação ao movimento realizado, como pode ser notado pelo valor do erro médio quadrático entre os sinais de entrada e saída que é de 12.38% com desvio padrão de 0.3502.

A Figura 15(c) ilustra a velocidade do quadrotor em seu próprio sistema de referência, a qual apresenta uma correspondência direta com os comandos de arfagem enviados pelo usuário, gerados a partir do movimento de sua mão direita.

3.3 Movimento Descendente/Ascendente

A Figura 16(a) mostra a evolução da posição da mão do usuário ao realizar o movimento “para cima e para baixo” com a mesma. Sendo que o comando direcional é enviado ao quadrotor apenas para $t \geq 0$.

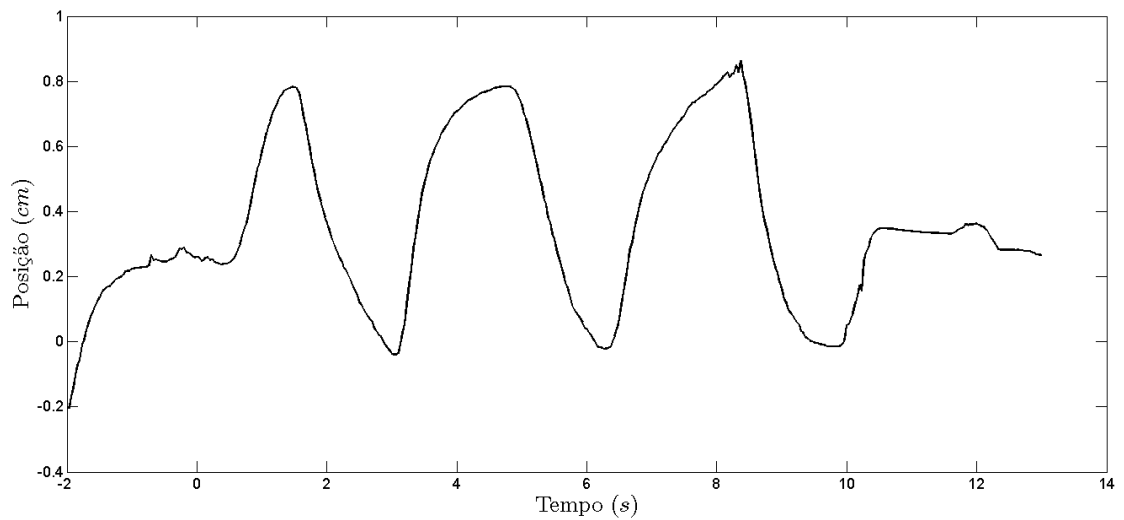
A Figura 16(b) mostra a comparação entre o sinal de entrada (usuário) e o sinal de saída (quadrotor). Note que o movimento do quadrotor para subida e descida sofre influência da aceleração da gravidade e, por isso, ao subir, a velocidade do quadrotor não consegue atingir um valor maior ou igual ao sinal de entrada como nos outros dois movimentos citados anteriormente. Por sua vez, ao descer, verifica-se que a resposta do sistema apresenta um sobrepasso negativo em relação ao sinal de entrada. Desse modo espera-se um erro entre os sinais de entrada e saída maior em relação aos dois outros movimentos anteriores. Assim, o erro médio quadrático entre os sinais de entrada e saída é de 15.65% com desvio padrão de 0.2146. Como nos casos anteriores, a aplicação dos sinais de controle no quadrotor somente acontece para $t \geq 0$.

Na Figura 16(c) observa-se o deslocamento da altura do quadrotor. Como explicado

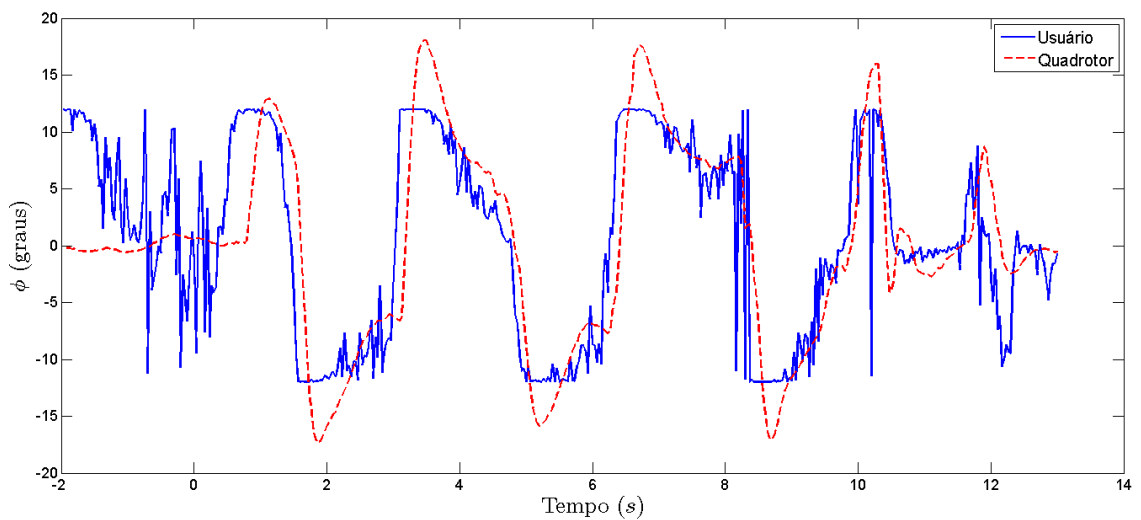
anteriormente, há uma tendência maior de movimento descendente por causa da aceleração da gravidade, logo o deslocamento para baixo é maior que o para cima, resultando em uma perda de altitude.

Vale ressaltar que experimentos modificando a guinada da aeronave também foram realizados, entretanto optou-se por não apresentar tais resultados, dado que as entradas de controle são valores constantes, conforme explicitado na Tabela 1 e na Seção 2.1.2.

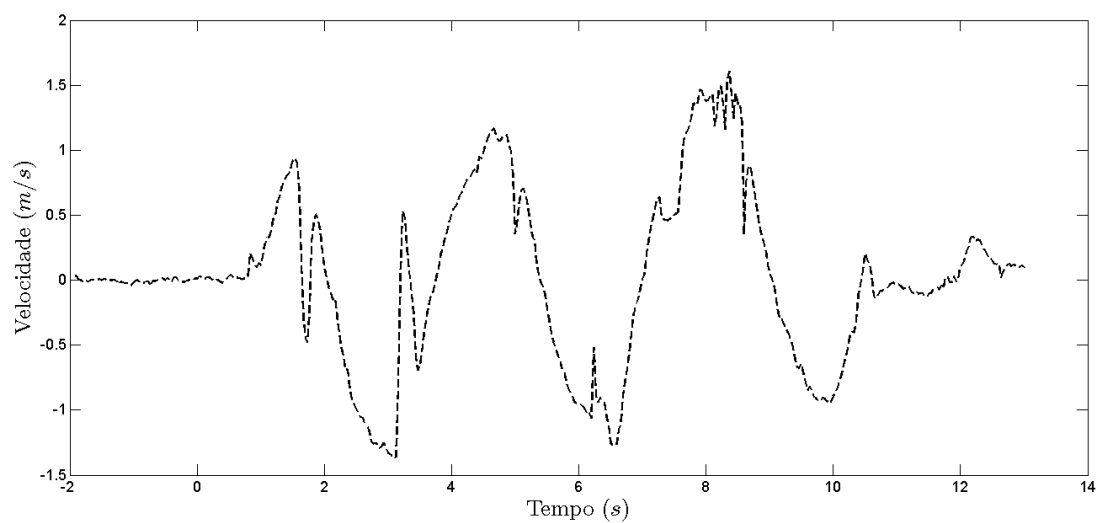
Em adição, também optou-se por não mostrar um movimento complexo 3D executado através dos comandos da mão direita, pois em nosso entendimento a composição (ação conjunta) dos movimentos mais simples aqui descritos podem ser facilmente utilizados para reproduzi-los.



(a) Gráfico da Posição da Mão Direita do Usuário no Eixo-X do Referencial do Kinect.

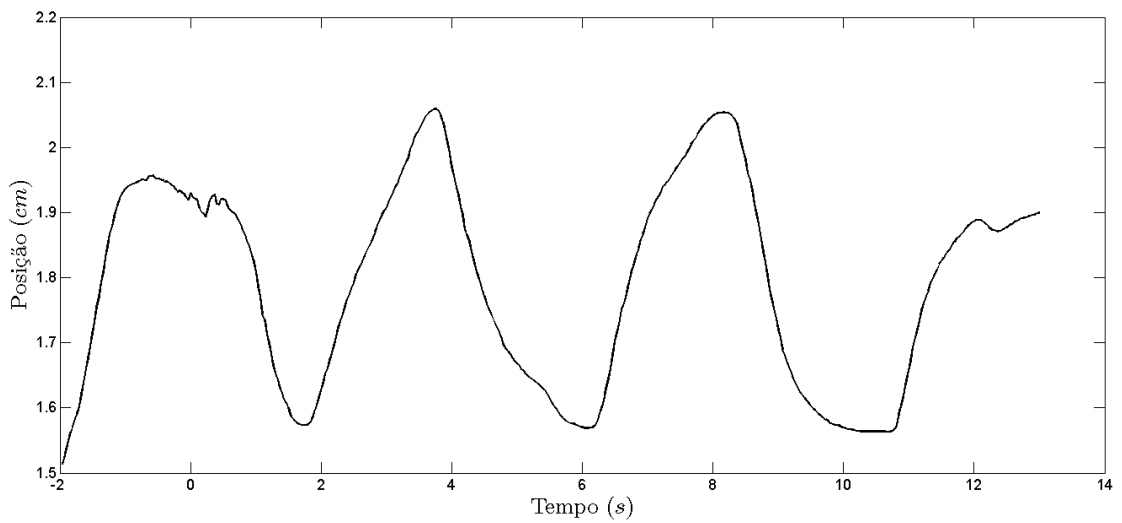


(b) Gráfico do Sinal de Entrada (Usuário) e de Saída (Quadrotor) para Movimento Lateral.

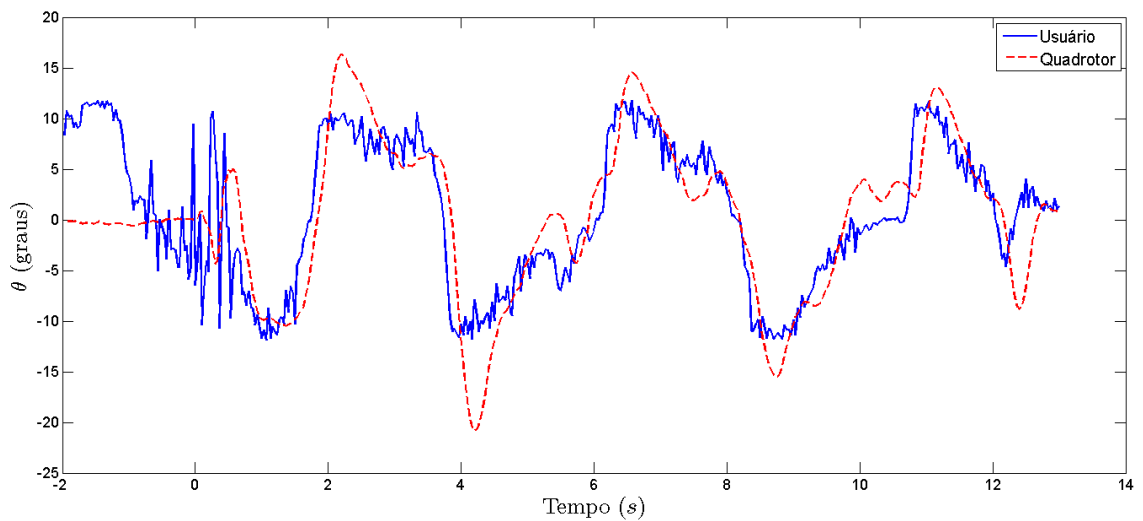


(c) Gráfico da Velocidade do Quadrotor no Eixo-Y do Referencial do Quadrotor.

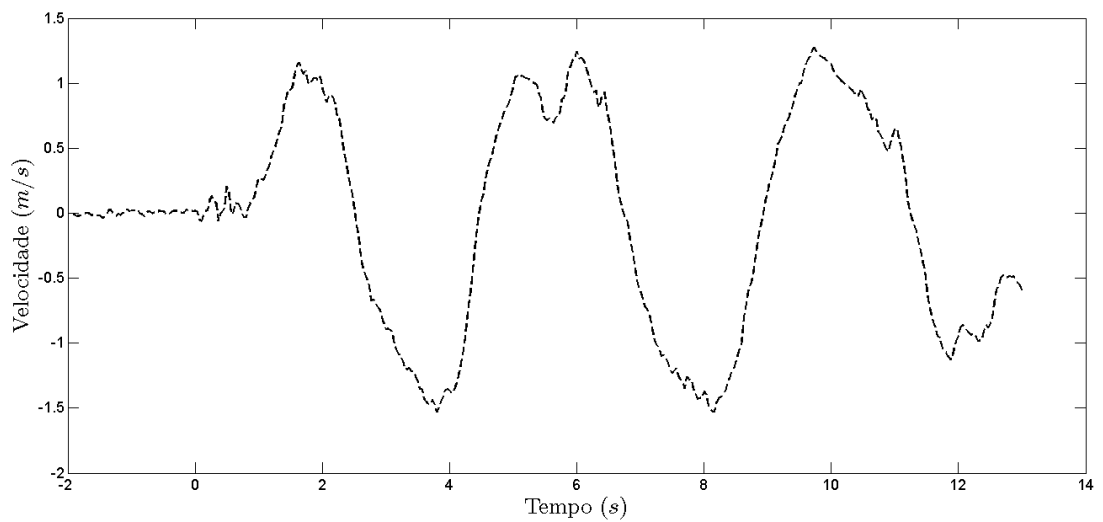
Figura 14: Controle do movimento de rolagem do quadrotor.



(a) Gráfico da Posição da Mão Direita do Usuário no Eixo-Z do Referencial do Kinect.

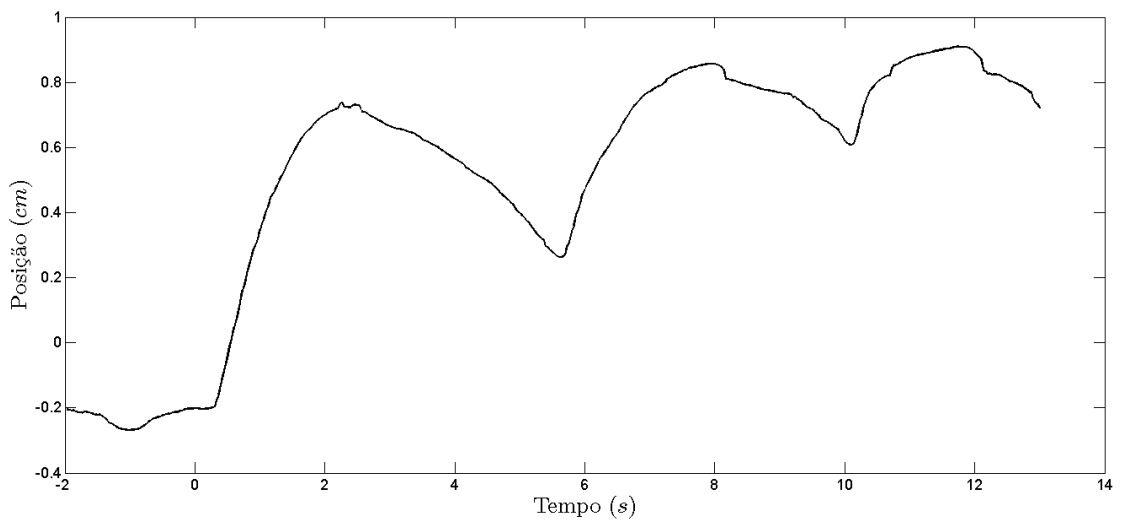


(b) Gráfico do Sinal de Entrada (Usuário) e de Saída (Quadrotor) para Movimento de Frente/Trás.

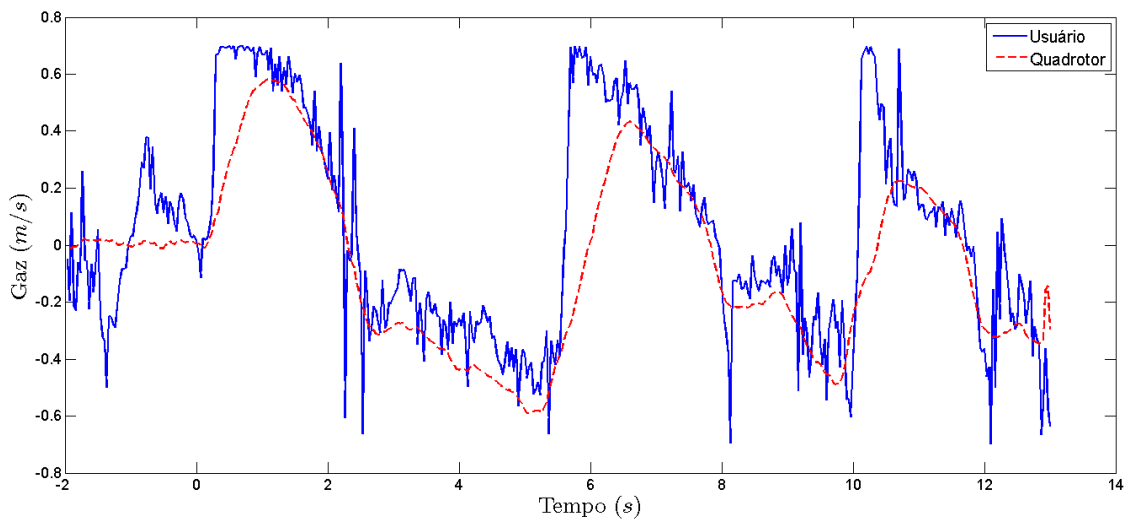


(c) Gráfico da Velocidade do Quadrotor no Eixo-X do Referencial do Quadrotor.

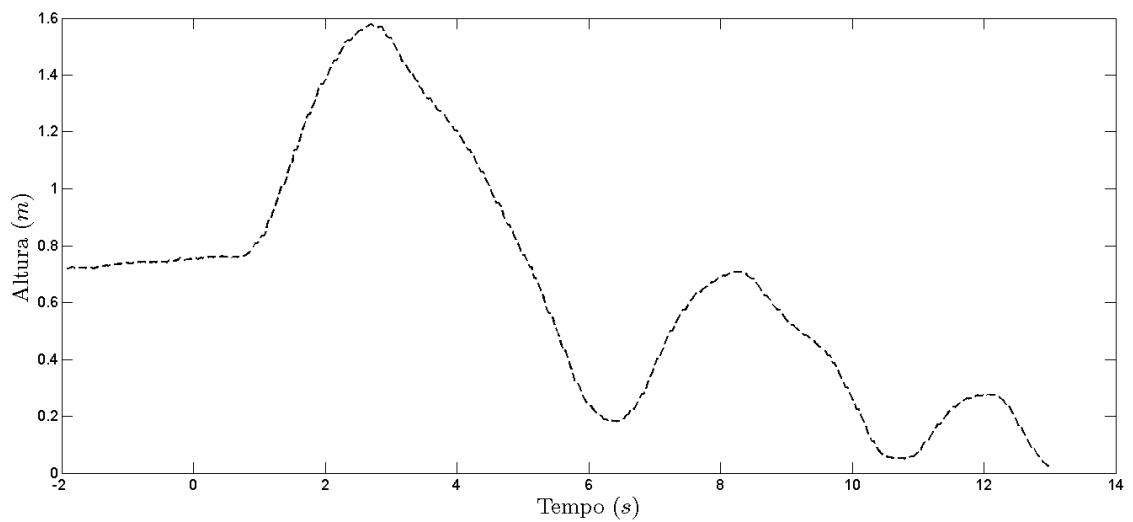
Figura 15: Controle de movimento de arfagem do quadrotor.



(a) Gráfico da Posição da Mão Direita do Usuário no Eixo-Y do Referencial do Kinect.



(b) Gráfico do Sinal de Entrada (Usuário) e de Saída (Quadrotor) para Movimento de Subida e Descida.



(c) Gráfico da Altura do Quadrotor no Eixo-Z do Referencial do Quadrotor.

Figura 16: Controle de movimento de descendente e ascendente do quadrotor.

4 *Conclusões*

Uma interação entre homem e máquina utilizando o sensor Kinect foi implementado nesse trabalho. Inicialmente, fez-se o rastreamento do usuário, de modo que os gestos feitos pelo mesmo pudessem ser interpretados e enviados ao quadrotor em forma de comandos básicos e intuitivos.

Na sequência, o uso do processo de derivação temporal da posição para se obter o controle de movimento de um quadrotor mostrou-se simples, viável e aplicável. Os sinais de saída da aeronave mantiveram a tendência do sinal de entrada, desse modo, pode-se afirmar que a interface homem-máquina proposta, apresentada e desenvolvida nessa monografia sustentou seu objetivo de fazer o controle de movimento de um quadrotor ArDrone Parrot de forma satisfatória.

O uso do sensor Kinect como um novo instrumento de interação homem-máquina mostrou um alto desempenho para rastrear o usuário, devido a sua tecnologia de reconhecimento de pessoas, e para retornar os dados da posição espacial do mesmo através de seu sensor de profundidade, o que teve papel principal na realização deste trabalho para que a interação com o quadrotor pudesse ter êxito.

Pelos resultados obtidos, percebeu-se que o quadrotor respondeu aos comandos enviados pelo Kinect em tempo real. Para os movimentos de rolagem e arfagem, os resultados apresentaram erros de seguimento inferiores a 13%, para o caso de movimento ascendente/descendente, esses erros foram de aproximadamente 16%, devido a ação externa (gravidade) observada sobre a aeronave. Vale destacar que para este trabalho não foi contemplado o controle automático de estabilização de altitude.

Para continuidade deste trabalho, propõe-se a realização de um controle do quadrotor baseado em posição do braço frente ao sensor kinect, o que seria diferente da presente proposta, a qual é baseada em intenção de movimento.

Referências

- [1] PREECE, J.; ROGERS, Y.; SHARP, H. *Interaction Design: Beyond Human-Computer Interaction*. 2nd ed.. ed. USA: Wiley, 2006. ISBN 978-0-470-01866-8.
- [2] BARBOSA, S. D. J.; SILVA, B. S. *Interação Humano-Computador*. Série sbc. Brasil: Editora Campus-Elsevier, 2010.
- [3] BAECKER, R. M.; BUXTON, W. *Human-computer interaction: a multidisciplinary approach*. [S.l.]: Morgan Kaufmann Publishers Inc., 1987.
- [4] RAMASWAMY, A.; MONSUEZ, B.; TAPUS, A. Model driven software development for human-machine interaction systems. *HRI '14 Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, p. 270–271, March 2014.
- [5] GARCÍA, L. S. A interação humano-computador e o design da interface-usuário. 2003.
- [6] MATHEW, A. R.; HAJJ, A. A.; ABRI, A. A. Human-computer interaction (hci): An overview. *Computer Science and Automation Engineering (CSAE)*, v. 1, p. 99–100, June 2011.
- [7] LIU, W. Natural user interface-next mainstream product user interface. *2010 IEEE 11th International Conference on Computer-Aided Industrial Design and Conceptual Design (CAIDCD)*, v. 1, p. 203–205, November 2010.
- [8] FAN, Y.-C.; CHEN, C.-L.; HUANG, S.-S. Natural user interface for interactive television. *2013 IEEE 17th International Symposium on Consumer Electronics (ISCE)*, p. 189–190, June 2013.
- [9] CHOI, S.-W.; KIM, W.-J.; LEE, C. H. Interactive display robot: Projector robot with natural user interface. *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, p. 109–110, March 2013.
- [10] YAO, J.; FERNANDO, T.; WANG, H. A multi-touch natural user interface framework. *2012 International Conference on Systems and Informatics (ICSAI)*, p. 499–504, May 2012.
- [11] PAVLOVIC, V. I.; SHARMA, R.; HUANG, T. S. Invited speech:”gestural interface to a visual computing environment for molecular biologists”. In: *FG '96 Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition (FG '96)*. Washington, DC: IEEE Computer Society, 1996. p. 30.
- [12] SELKER, T. Touching the future. *Communications of the ACM - Surviving the data deluge*, v. 51, p. 14–16, December 2008.

- [13] MICROSOFT. *Efeito Kinect*. <http://www.xbox.com/pt-BR/Kinect/Kinect-Effect>. Acessado em 09/04/2014.
- [14] MICROSOFT. *Kinect for Windows Sensor Components and Specifications*. <http://msdn.microsoft.com/en-us/library/jj131033.aspx>. Acessado em 09/04/2014.
- [15] ARIELLI, Y. et al. *Depth mapping using projected patterns*. USA: Google Patents, abr. 3 2012. US Patent 8,150,142.
- [16] SPEKTOR, E.; MOR, Z.; RAIS, D. *Integrated processor for 3D mapping*. USA: Google Patents, jun. 4 2013. US Patent 8,456,517.
- [17] NARDY, M. A. *Predição da Trajetória Parabólica em Lançamentos Livres usando o Sensor Kinect*. Monografia (Graduação) — Universidade Federal de Viçosa, Minas Gerais, 2013.
- [18] BECKER, M.; BOUABDALLAH, S.; SIEGWART, R. Desenvolvimento de um controlador de desvio de obstáculos para um mini-helicóptero quadri-rotor autônomo-lafase: Simulação. In: *Proceedings of the CBA-Congresso Brasileiro de Automática (CBA 2006)*. Juiz de Fora: SBA, 2006. p. 1201–1206.
- [19] TOURNIER, G. P. et al. Estimation and control of a quadrotor vehicle using monocular vision and moire patterns. In: *Proceedings of the AIAA Guidance, Navigation and Control Conference and Exhibit*. Keystone, Colorado: AIAA, 2006. p. 21–24.
- [20] KIM, S. K.; TILBURY, D. M. Mathematical modeling and experimental identification of an unmanned helicopter robot with flybar dynamics. *Journal of Robotic Systems*, Wiley Online Library, v. 21, n. 3, p. 95–116, 2004.
- [21] CASTILLO, P.; LOZANO, R.; DZUL, A. E. Modelling and control of mini-helicopters. *Modelling and Control of Mini-Flying Machines*, Springer, p. 81–119, 2005.
- [22] KONDAK, K. et al. Autonomously flying vtol-robots: Modeling and control. In: IEEE. *Proceedings of the IEEE International Conference on Robotics and Automation 2007*. Roma, Italy, 2007. p. 736–741.
- [23] LOPES, L. L. *Modelagem em Caixa Branca e Caixa Preta de um Quadrirotor Ar-drone*. Monografia (Graduação) — Universidade Federal de Viçosa, Minas Gerais, 2013.
- [24] KRAJNÍK, T. et al. Ar-drone as a platform for robotic research and education. In: *Proceedings of the Research and Education in Robotics-EUROBOT 2011*. [S.l.]: Springer, 2011. p. 172–186.
- [25] SANNA, A. et al. A kinect-based natural interface for quadrotor control. In: _____. *Intelligent Technologies for Interactive Entertainment*. Genova, Italy: Springer Berlin Heidelberg, 2012. v. 78, cap. 3, p. 179–186.
- [26] PFEIL, K.; KOH, S. L.; LAVIOLA, J. Exploring 3d gesture metaphors for interaction with unmanned aerial vehicles. In: *IUI '13 Proceedings of the 2013 international conference on Intelligent user interfaces*. New York, USA: ACM New York, 2013. p. 257–266.

-
- [27] BOUDJIT, K.; LARBES, C.; ALOUACHE, M. Control of flight operation of a quad rotor ar. drone using depth map from microsoft kinect sensor. *International Journal of Engineering and Innovative Technology (IJEIT)*, v. 3, September 2013.
- [28] KAR, A. Skeletal tracking using microsoft kinect. Department of Computer Science and Engineering, IIT Kanpur. 2010.
- [29] STOWERS, J.; HAYES, M.; BAINBRIDGE-SMITH, A. Altitude control of a quadrotor helicopter using depth map from microsoft kinect sensor. In: *Proceedings of the 2011 IEEE International Conference on Mechatronics*. Beijing, China: IEEE, 2011. p. 358–362.