**UNIVERSIDADE PAULISTA**

**DANILO DE OLIVEIRA DOROTHEU**

**THIAGO GUY MOZOL VICENTE**

**MARCIO FERNANDES CRUZ**

**DIEGO DA SILVA SANTANA**

**DESENVOLVIMENTO DE UMA APLICAÇÃO DE VISÃO COMPUTACIONAL PARA O RECONHECIMENTO DE GESTOS**

**São Paulo**

**2015**

**UNIVERSIDADE PAULISTA**

**DANILO DE OLIVEIRA DOROTHEU**

**THIAGO GUY MOZOL VICENTE**

**MARCIO FERNANDES CRUZ**

**DIEGO DA SILVA SANTANA**

**DESENVOLVIMENTO DE UMA APLICAÇÃO DE VISÃO COMPUTACIONAL PARA O RECONHECIMENTO DE GESTOS**

Trabalho de conclusão de curso para obtenção do título de graduação em Ciências da Computação, apresentado à Universidade Paulista – UNIP.

Orientador: Profº: Gley Fabiano

**São Paulo**

**2015**

**SUMARIO**

[1. INTRODUÇÃO 4](#_Toc419923307)

[2. ESCOLHA DO TEMA 5](#_Toc419923308)

[3. DEFINIÇÃO DO PROBLEMA 6](#_Toc419923309)

[4. DEFINIÇÃO DA HIPÓTESE 6](#_Toc419923310)

[5. DEFINIÇÃO DOS OBJETIVOS 7](#_Toc419923311)

[5.1 OBJETIVO GERAL 7](#_Toc419923312)

[5.2 OBJETIVO ESPECÍFICO 7](#_Toc419923313)

[6. JUSTIFICATIVA 7](#_Toc419923314)

[7. REFERENCIAL TEÓRICO 8](#_Toc419923315)

[8. METODOLOGIA 9](#_Toc419923316)

[9. CRONOGRAMA 10](#_Toc419923317)

[10. SUMARIO PROVISÓRIO 11](#_Toc419923318)

**REFERÊNCIAS..........................................................................................................13**

# INTRODUÇÃO

Para se estabelecer a comunicação entre duas pessoas, é necessário que ambas utilizem um denominador comum para que possam ser compreendidas as informações transmitidas. Para a Linguagem Gestual, os diversos sinais e movimentos exigem do interlocutor, prévio aprendizado, assim como os idiômas, onde é necessário ter conhecimento prévio da gramática da Linguagem cujo qual será usada para estabelecer comunicação.

A aplicação da Visão Computacional permite o desenvolvimento de mecanismos que possibilitam a interação humano-computador baseado na visão, onde são utilizadas câmeras para a captura de imagens que posteriormente serão manipuladas por técnicas de processamento e tratamento, a fim de se obter uma determinada informação (TRUYENQUE, 2005).

O problema encontrado nesta pesquisa é como desenvolver um método de comunicação entre duas pessoas, sendo uma delas deficiente auditiva e a outra, sem qualquer conhecimento em uma linguagem de sinais, como LIBRAS por exemplo.

Através das técnicas da visão computacional, é possível desenvolver uma ferramenta computacional para que possa capturar através de uma câmera ligada ao computador imagens de uma mão livre, ou seja, sem nenhum acessório, identificando sinais da linguagem LIBRAS e traduzindo-a em linguagem natural.

O objetivo geral desta pesquisa é estudar os conceitos de visão computacional, como segmentação, aplicação de filtros que visam adquirir, processar e interpretar imagens, e o objetivo específico, é desenvolver um método que faça o reconhecimento de gestos de sinais em LIBRAS e interpretá-los em uma linguagem natural.

Um dispositivo computacional com entrada de dados baseados de gestos possibilitaria uma interação mais acessível e intuitiva com o computador. Sua importância em desenvolver um trabalho científico visando o tema “visão computacional” possibilita não só uma maior interação entre o homem e o computador, mas também na questão da acessibilidade, auxiliando na facilidade de comunicação entre a comunidade deficiente auditiva e a sociedade.

A pesquisa utilizada neste trabalho é do tipo exploratório por motivo que a área da visão computacional se considerada pouco explorada e repleta de problemas pelos motivos de sua complexidade (RIOS, 2010, p. 7). O método de pesquisa a ser utilizado neste trabalho é o dedutivo, onde pesquisaremos os conceitos gerais da visão computacional até o objetivo especifico de desenvolver um método que reconheça gestos de sinais em LIBRAS.

# ESCOLHA DO TEMA

A linguagem Gestual é uma ferramenta que viabiliza a comunicação para pessoas com deficiências auditivas ou qualquer outra limitação de transmitir informação através da fala. Esta ferramenta vem sendo adotada por instituições ou empresas, como é o exemplo da própria Universidade Paulista, que disponibiliza o estudo desta em seus cursos de Graduação.

É muito comum empresas que atuam no segmento público adotar esta linguagem também, disponibilizando até mesmo cursos aos funcionários, visando aderir à acessibilidade da informação.

Porém, uma das limitações da Linguagem Gestual é a necessidade do seu prévio aprendizado. Diferente de qualquer idiôma (onde muitos aplicativos suprem a necessidade do orador dominar o idioma estrangeiro), o ramo do reconhecimento de gestos como LIBRAS através de uma câmera conectada ao computador é uma área de investigação recente.

“A Língua Gestual é uma língua complexa. Tem a sua própria sintaxe e é composta por diversos elementos organizados no espaço. A complexidade do gesto realizado com as mãos, a posição e trajetória que estas fazem no espaço tridimensional, a expressão facial durante a execução do movimento e até o posicionamento do corpo é, todos em conjunto, o que trazem significado a um gesto. A natureza multi-modal do gesto na língua, assim como a detecção dos ligeiros e precisos movimentos das mãos, é dos principais desafios no reconhecimento da Língua Gestual [...]” Correia (2013, p.2).

Portanto, desenvolver um método de tradução de gestos através da visão computacional pode trazer inúmeros benefícios para a comunidade deficiente auditiva, e sendo assim, este será o tema deste trabalho.

# DEFINIÇÃO DO PROBLEMA

Como desenvolver um método de comunicação entre duas pessoas, sendo uma delas deficiente auditiva e a outra, sem qualquer conhecimento em uma linguagem de sinais, como LIBRAS, por exemplo?

# DEFINIÇÃO DA HIPÓTESE

A interação da tecnologia no ramo de reconhecimento computacional pode auxiliar na identificação de elementos com maior precisão, como reconhecer objetos em um determinado espaço e extrair a informação necessária. Então, um dispositivo com esta capacidade de visão pode ser usado em diversos campos, entre eles, a comunicação. Um exemplo de aplicação neste ramo é a identificação de placas dos veículos que ultrapassam a velocidade permitida em uma via, onde o receptor (no caso, o radar) envia a foto capturada no ato da infração à central, com o fim de aplicar os diversos tipos de filtros até que se identifique o responsável pelo veículo para aplicar as medidas cabíveis.

Outro exemplo de aplicação desta ferramenta é o Microsoft Kinect, onde seu objetivo é substituir o controle que até então era fundamental na interação do usuário com o console pelas mãos (sem qualquer tipo de sensor) e as ações que estas geram (CORREIA, 2013)*.*

Portanto, podemos utilizar a visão computacional para identificar elementos da comunicação gestual, mesclando as funcionalidades dos dois exemplos citados acima. O sistema deve ser capaz de capturar a imagem das mãos do usuário e aplicar técnicas que visam tratar a imagem a fim de facilitar seu entendimento, além de processar este resultado e interpretá-lo como uma informação válida. Então, uma pessoa com deficiência auditiva pode utilizar esta ferramenta para transmitir informação com qualquer pessoa, sem que estas conheçam outro tipo de comunicação que não utilize a fala.

# DEFINIÇÃO DOS OBJETIVOS

Elaborar uma pesquisa científica, focado na interação Humano-Computador, baseado em alguns gestos de uma mão livre, sem acessório algum, para que possa interagir com o computador.

# OBJETIVO GERAL

É investigar as técnicas da visão computacional, que visam adquirir, processar e interpretar imagens a fim de estabelecer um meio de visão ao computador.

# OBJETIVO ESPECÍFICO

É desenvolver um método que reconheça gestos da linguagem libras por meio de uma câmera convencional utilizando as técnicas da visão computacional.

# JUSTIFICATIVA

Este é um trabalho de uma aplicação cuja função é estabelecer um meio de visão ao computador para que seja factível uma leitura gestual vinda, mais precisamente, de um deficiente surdo-mudo. Com a finalidade de se desenvolver um tipo de tradutor de gestos, a tecnologia ainda poderia ser utilizada para outros objetivos como, comandos mais iterativos para jogos e atalhos de computadores, a substituição de um mouse, o reconhecimento de objetos, enfim, há muitas utilidades que seriam viáveis, porém o mais prestadio entre eles é o tradutor, pois envolve a questão da acessibilidade para deficientes, que é, no entanto, uma prioridade para sociedade.

A importância em desenvolver um trabalho científico visando o tema “visão computacional” possibilita não só uma maior interação entre o homem e o computador, mas também na questão da acessibilidade, auxiliando na facilidade de comunicação entre a comunidade deficiente auditiva e a sociedade.

Uma das vantagens da tecnologia é permitir o desenvolvimento de soluções que viabilizam o processo de comunicação sem obrigar o prévio aprendizado da linguagem, como uma espécie de tradutor. Diversas ferramentas (como o Google Translator, por exemplo), identificam o conteúdo a ser traduzido (seja ele em texto, voz e até mesmo, imagem) e interpretam as palavras correspondentes no idiôma escolhido. Mas e quando estas informações partem de um interlocutor que necessita traduzir algum tipo de linguagem gestual? Uma das soluções seria desenvolver um dispositivo computacional com entrada de dados baseadas em gestos para seu reconhecimento e interpretação por meio das funções típicas da visão computacional em conjunto com técnicas especificas de reconhecimento de gestos que se agrega com as de processamento de imagens e da computação gráfica.

A visão computacional possui muito material teórico que possibilita iniciar uma pesquisa para desenvolvimento de um trabalho acadêmico especifico ao reconhecimento de gestos de uma mão humana.

Enfim, é viável desenvolver o método, devido nossas pesquisas cuja bibliografia está anexada neste projeto, que há a possibilidade de criar uma interação humana com o computador.

# REFERENCIAL TEÓRICO

A visão computacional pode ser entendida como um mecanismo inteligente que localiza e identifica através de técnicas e cálculos matemáticos, objetos contidos em uma imagem digital, podendo ser obtida através de uma câmera que, segundo Truyenque (2005, p. 21), não há restrições de características, porém há características que podem aprimorar seus resultados em termos de qualidade.

Além disso, não seria apenas localizar e identificar um determinado objeto, também seria extrair informações uteis a fim de resolver problemas reais, como por exemplo, o reconhecimento de gestos, que representa um caso mais específico da visão computacional. Neste caso, o reconhecimento de gestos estaria se relacionando com a iteração humano-computador, tornando assim uma maneira mais intuitiva e também, ao mesmo tempo, uma solução para a interpretação da comunicação de gestos em tempo real.

De acordo com Truyenque (2005, p.18), os sistemas baseados em visão utilizam técnicas de reconhecimento de padrões nas imagens, ou seja, essas técnicas são cruciais para a localização dos objetos, e sem elas não há a possibilidade de extrair alguma informação útil para o caso dos gestos ou de muitos outros sistemas que possuem como alvo de sua informação de interesse um objeto especifico.

‘’Diversos métodos de computação visual são empregados, combinando aspectos de computação gráfica, processamento de imagem e inteligência artificial. ‘’ (MONTEIRO E CONCI, 2003, p. 2). Para reconhecimento de placas, são utilizados métodos de inteligência artificial, no qual não é utilizado para reconhecimento de gestos, por tanto, seria um assunto a mais a ser analisado. Os métodos de processamento de imagens e computação gráficas foram os mais abordados dentre os muitos trabalhos relacionados à visão computacional.

# METODOLOGIA

De acordo com Rios (2010, p. 8) “[...] a visão computacional é um campo que, se já traz benefícios extraordinários com o pouco que se sabe, trará muito mais com o universo que ainda falta ser explorado”. Com base nesta afirmação, devemos levar em consideração os inúmeros problemas envolvidos numa área ainda pouco evoluída. Portanto a metodologia de pesquisa será do tipo exploratório a fim de focar principalmente na descoberta e explicação dos problemas envolvidos.

O método a ser utilizado é o dedutivo, pois a princípio será abrangida na pesquisa a parte que aborda os conceitos gerais da visão computacional visando estabelecer um meio de visão ao computador. Em seguida será um caso mais especifico, buscando apresentar um método cuja função é reconhecer movimentos e objetos do mundo real que, pode ser considerada uma solução para o problema da comunicação de gestos.

Este trabalho será realizado por meio de pesquisas bibliográficas, com consulta de obras de autores sobre os assuntos pertinentes ao tema, além de artigos acadêmicos e pesquisa documental. Desta forma, dividiu-se o projeto em diferentes fases, definido como:

1. Levantamento e documentação referente ao tema;  
  
2. Análise de tecnologias para obtenção dos gestos;  
  
3. Análise de elementos manuais, características da Linguagem Gestual Brasileira;  
  
4. Implementação do software de reconhecimento de gestos;  
  
5. Análise de resultados obtidos pelo software;  
  
6. Documentação das etapas 2, 3, 4 e 5;  
  
7. Documentação da conclusão geral do projeto;

# CRONOGRAMA

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Atividade/Mês | JUN | | JUL | | AGO | | SET | | OUT | | NOV | | DEZ | |
| 01-15 | 16-30 | 01-15 | 16-31 | 01-15 | 16-31 | 01-15 | 16-30 | 01-15 | 16-31 | 01-15 | 16-30 | 01-15 | 16-31 |
| Levantamento e documentação de informações referentes à visão computacional |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pesquisa de tecnologias e técnicas para obtenção dos gestos |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Análise de elementos manuais, característicos da Linguagem Gestual Brasileira. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Implementação do software de reconhecimento de gestos |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Análise de resultados obtidos pelo software |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Documentação das etapas 2, 3, 4 e 5. |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Documentação da conclusão geral do projeto |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Apresentação do Trabalho Final |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# 10. SUMARIO PROVISÓRIO

**REFERÊNCIAS**

AZEVEDO, Eduardo; CONCI, Aura. **Computação Gráfica**: geração de Imagens. São Paulo: Elsevier Editora LTDA, 2003.

CORREIA, Miguel. **Reconhecimento de Elementos Gestuais com *Kinect***: preparação de mestrado. 2013. Disponível em: <http://paginas.fe.up.pt/~ee06160/thesis/wpcontent/uploads/2013/03/RelatorioPDI\_MiguelCorreia\_ee06160.pdf>. Acesso em: 7 de Março de 2015.

FERNANDO, Luiz. **Reconhecimento ótico de caracteres (ORC) e analise de sistema ORC baseados em códigos abertos.** 2012. 44 f.Monografia para graduação em processamento de dados. Universidade (FATEC-SP), São Paulo, 2012.

JUNIOR, Iális Cavalcante de Paula: **Técnicas de Processamento Digital de Imagens com Java**. Disponível em: <http://www.ppgia.pucpr.br/~facon/ComputerVisionBooks/2009ProcessamentoImagensComJava.pdf>. Acesso em: 8 de março de 2015.

OPENKINECT, 2011. **Wiki Open *Kinect***. Disponível em: <http://openkinect.org/>. Acesso em: 13 de abril de 2015.

RIOS, Luiz Romário Santana. **Visão Computacional**. Departamento de Ciência da Computação, Universidade Federal da Bahia (UFBA) 2010. Disponível em: <http://homes.dcc.ufba.br/~luizromario/Apresenta%C3%A7%C3%A3o%20de%20IA/Artigo%20%28final%29.pdf‏>. Acesso em 14 de maio de 2015.

SANTOS, Ana; MOSER, Andre; CRUZ, Felipe; ODA, Lucas; ROCCO, Thiago. **KINLIB: Protótipo de Ensino de Libras Utilizando o Kinect**. Disponível em: <http://engenharia.anhembi.br/tcc-11/cco-02.pdf>. Acesso em: 7 de Março de 2015.

SIOLA, Felipe breve. **Desenvolvimento de um software para reconhecimento de sinais em libras através de vídeo**. Artigo de graduação em computação, universidade federal do ABC, 2010. Disponível em: < http://cmcc-bcc.ufabc.edu.br/bcc/docs/PG3-Final-FelipeSiola-Q3-2010.pdf>. Acesso em 8 de março de 2015.

SIQUEIRA, Mozart Lemos. **Reconhecimento automático de padrões em imagens ecocardiográficas**. 2010. 107 f. Dissertação (Doutorado Ciência da Computação) – Universidade Federal Do Rio Grande DO Sul, Porto Alegre, 2010.

TRUYENQUE, Michel Alain Quintana. **Uma Aplicação de Visão Computacional que Utiliza Gestos da Mão para Interagir com o Computador.** 2005. Dissertação (Mestrado informática) – Universidade PUC, São Paulo, 2005.