



UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO
Trabalho de Formatura Supervisionado

Vitruvius
Um Reconhecedor de Gestos para o
Kinect™

Tiago Andrade Togores

Orientador: Professor Flávio Soares Corrêa da Silva

19 de setembro de 2011

Agradecimientos

Resumo

Sumário

Prefácio	4
I Objetiva	6
1 Introdução	7
1.1 Contextualização	7
1.2 Objetivo	7
2 Tecnologias	9
2.1 Kinect	9
2.2 OpenNI	9
2.3 NITE	9
3 Conceitos	10
3.1 Aprendizado Computacional	10
3.2 Modelo oculto de Markov	10
4 Desenvolvimento	11
5 Resultados	12
6 Conclusão	13
II Subjetiva	14
1 Desafios e Frustrações	15
2 Relação com as disciplinas cursadas	16
3 Trabalhos Futuros	17
A Código	18
B Figuras	19
Referências	20

Prefácio

Esta monografia é fruto de um trabalho realizado na disciplina Trabalho de Formatura Supervisionado (MAC0499), oferecida aos graduandos do curso de Bacharelado em Ciência da Computação (BCC), do Departamento de Ciência da Computação (DCC), do Instituto de Matemática e Estatística (IME), da Universidade de São Paulo (USP), e ministrada no ano de 2011 pelo Professor Doutor Carlos Eduardo Ferreira (~cef).

Todos os produtos deste trabalho podem ser encontrados no sítio:

<http://www.linux.ime.usp.br/~togores/mac499>

O conteúdo do documento está dividido em duas partes. Parte objetiva:

Seção 1

Contextualização, motivação, apresentação do problema e da solução proposta.

Seção 2

Descrição das tecnologias utilizadas e apresentação de trabalhos semelhantes anteriormente realizados.

Seção 3

Definição de conceitos técnicos e matemáticos. Exposição e explicação dos algoritmos de inteligência artificial.

Seção 4

O que e como foi desenvolvido no período, incluindo problemas encontrados, mudanças de planos e metodologias.

Seção 5

Análise dos resultados e comparação com o que era esperado inicialmente.

Seção 6

Considerações finais.

Parte subjetiva:

Seção 1

Desafios e frustrações encontrados durante a realização do trabalho.
Como as dificuldades foram contornadas.

Seção 2

Relevância das disciplinas cursadas no trabalho. Quais os conceitos
que foram aplicados.

Seção 3

Aprimoramento do trabalho e do conhecimento na área.

Parte I
Objetiva

1 Introdução

1.1 Contextualização

Ao longo da história, foram desenvolvidos diversos modos de interação humano-computador, porém sempre com o mesmo objetivo: tornar o controle da máquina pelo homem eficiente, de modo que os resultados desejados fossem obtidos. Entretanto, apenas este aspecto não é suficiente para que o usuário tenha uma experiência agradável.

Interfaces naturais permitem que usuários se comuniquem com dispositivos de forma mais intuitiva, com uma curva de aprendizado baixa. Em pouco tempo, a interface fica imperceptível e a comunicação se torna natural.

Este conceito foi mais disseminado recentemente com o acesso de tecnologias como superfícies multitoque em smartphones e tablets pelo público geral. Anteriormente, já havia o reconhecimento de voz e fala. Em 2010, o Kinect, um periférico originalmente desenvolvido para o console Xbox 360 pela Microsoft, passou a permitir o desenvolvimento de programas que realizam comunicação através de tanto gestos corporais quanto som. Logo após seu lançamento, a comunidade de software livre conseguiu adaptar seu funcionamento para computadores pessoais.

No entanto, uma séria limitação existente para o usuário final é que as aplicações criadas até hoje se baseiam em reconhecer um conjunto fixo de gestos ou permitem a adição de gestos apenas simples (identificando apenas a direção do movimento da mão de uma pessoa, por exemplo).

1.2 Objetivo

Uma das melhores maneiras de contornar esse problema é aplicando conhecimentos da área de aprendizado computacional. Cada gesto seria associado a um modelo estatístico, que devolve uma probabilidade de que um gesto feito seja o que ele (modelo) representa. Ademais, cada modelo é treinado incrementalmente de modo que descreva com maior precisão a ação. Tem-se, assim, duas vantagens:

- Qualquer um pode criar seus próprios gestos;
- Os gestos virtualmente não tem limite de complexidade;

O trabalho desenvolvido consiste, então, em um aplicativo que permite treinamento e reconhecimento de gestos de modo prático para o usuário final, utilizando o Kinect. Para tanto, é necessário o estudo do funcionamento do dispositivo e de suas aplicações e a implementação de uma técnica de inteligência artificial mais adequada.

2 Tecnologias

2.1 Kinect

2.2 OpenNI

2.3 NITE

3 Conceitos

3.1 Aprendizado Computacional

3.2 Modelo oculto de Markov

4 Desenvolvimento

5 Resultados

6 Conclusão

Parte II
Subjetiva

1 Desafios e Frustrações

2 Relação com as disciplinas cursadas

3 Trabalhos Futuros

A Código

B Figuras

Referências