Gesture Recognition

Tobias Scheidt

3scheidt@informatik.uni-hamburg.de

Contents

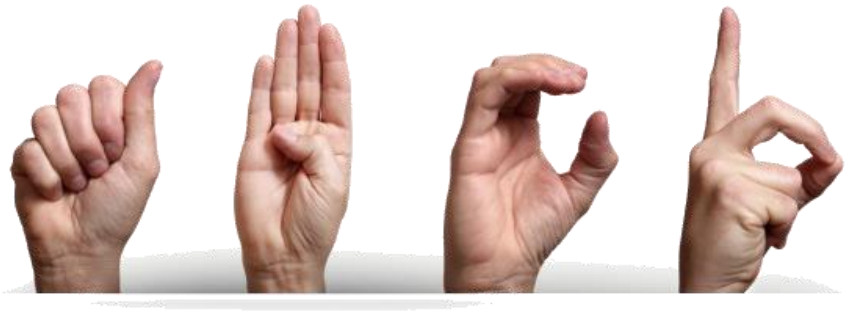
- Introduction
- Motivation
- Devices
- Detection of human bodies and parts of them
- How the Kinect works
- Algorithms for recognize dynamic gestures
- Frameworks

Gesture Recognition

- Interpreting human Gestures with mathematical algorithms
- Communication with the computer

Kinds of gestures

- Static gestures



- Dynamic gestures

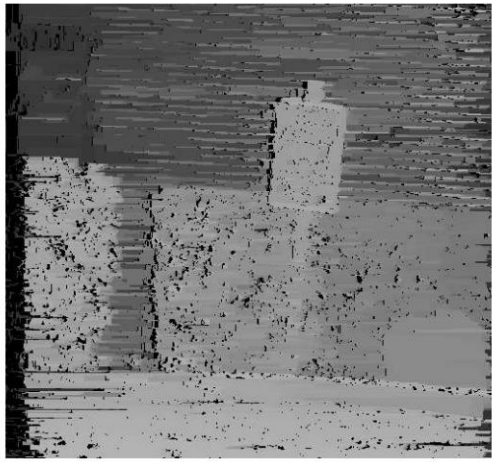
Motivation

- Breaking Human-Computer barrier
- Prevent possible bottlenecks
- Medical issues
  - Detecting of existing health disorder
  - Prevention

Devices

- Wired gloves
  - Computer collects information about the position and rotation of the hand and fingers

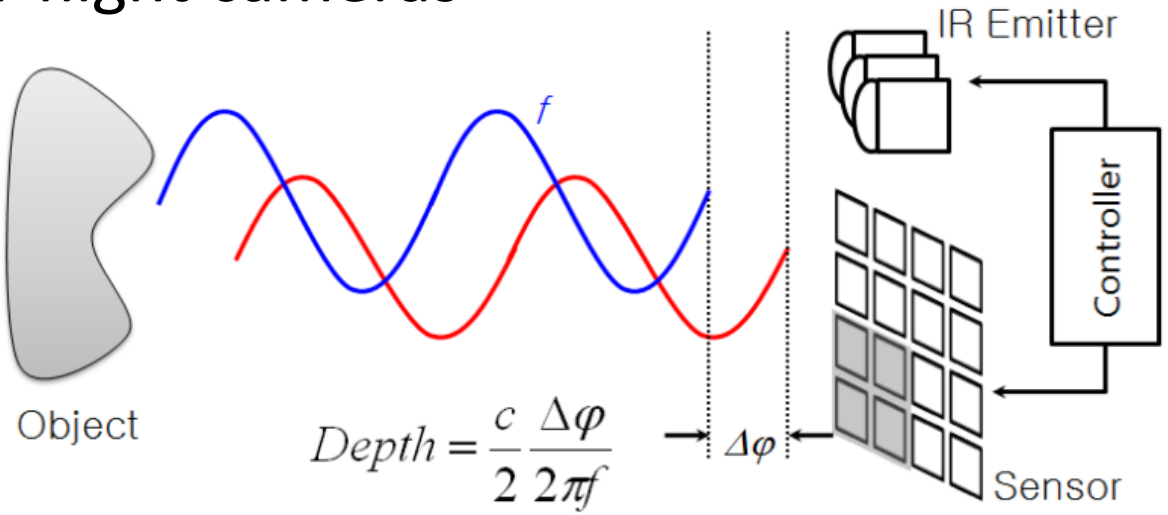
- Stereo cameras



Devices

- Depth-aware cameras
  - structured light cameras

  - time-of-flight cameras



Devices

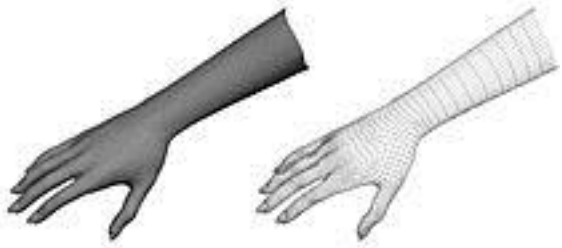
- Controller-based gestures
  - Controllers act as an extension of the body

- Single camera
  - Color based recognition
  - Not anymore state-of-the-art

Modern gesture recognition systems

Detection of human bodies and parts of them

- 3D model-based algorithms

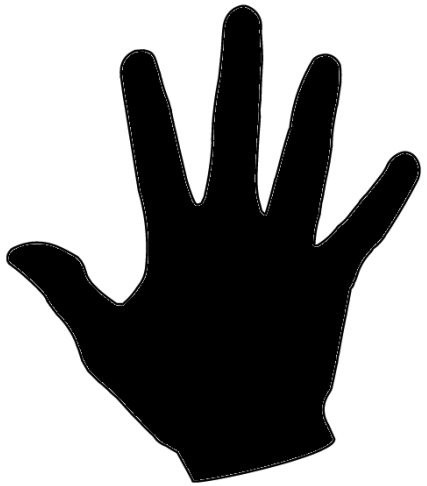


- Skeletal-based algorithms

- Appearance-based models

Appearance-based models

- Does not use spatial representation of the body
- Using a template database
- Mostly used for hand-tracking

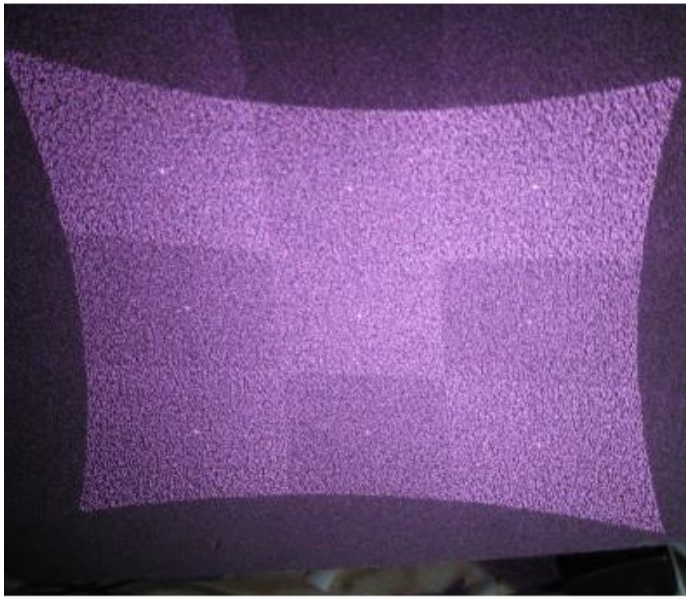


How the Kinect depth sensor works

- Light coding algorithm
  - Patented from PrimeSense
  - Not documented
    - Assumptions due to the patent
    - Studies of the Kinect
  - Uses the infrared emitter and sensor
    - Cheaper than Time-of-Flight depth cameras

How the Kinect depth sensor works

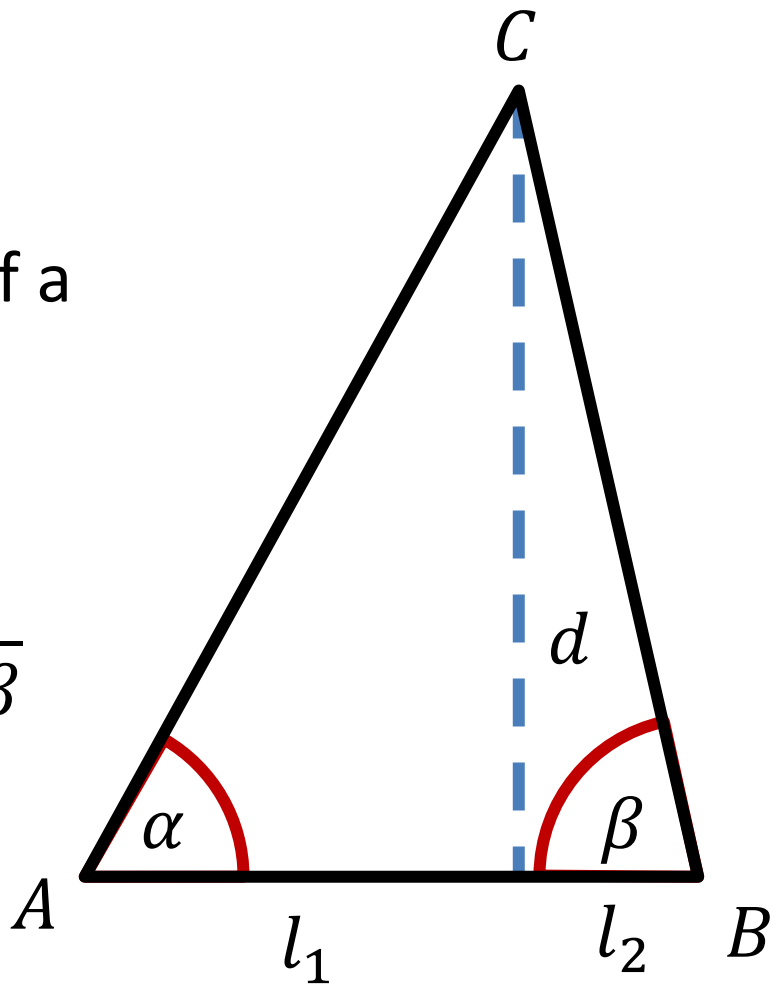
- IR-emitter projecting an alternating dot pattern in the room
- Probably it is a code that simplifies the calculation



How the Kinect depth sensor works

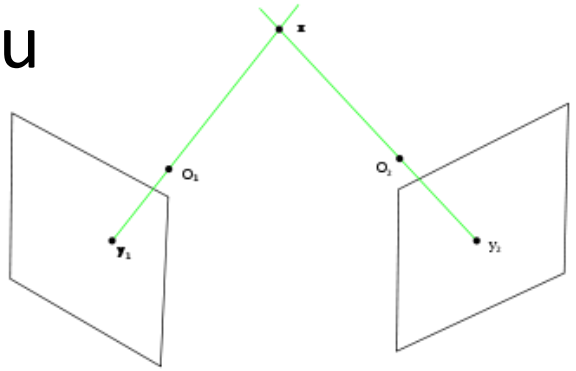
- Triangulation method
  - Determine the location of a point by measuring the angle to it.

$$l = l_1 + l_2 = \frac{d}{tan\alpha} + \frac{d}{tan\beta}$$



How the Kinect depth sensor works

- Kinect has just one IR sensor
- For the triangulation method you need two IR images
- There is a trick:
  - One image is captured by the IR sensor
  - The other picture is invisible and just virtual



- Calculate the horizontal offset of the speckles on the first image relative to hardcoded position

Using multiple Kinect-systems

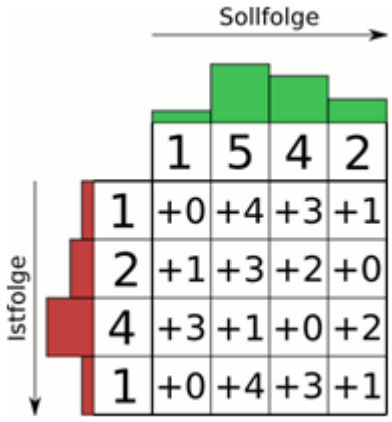
- Great problem of optical systems is the occlusion
- Visible area is just 4 m^2
- Theoretical possible to use multiple Kinect-systems without interference
- Object of current research

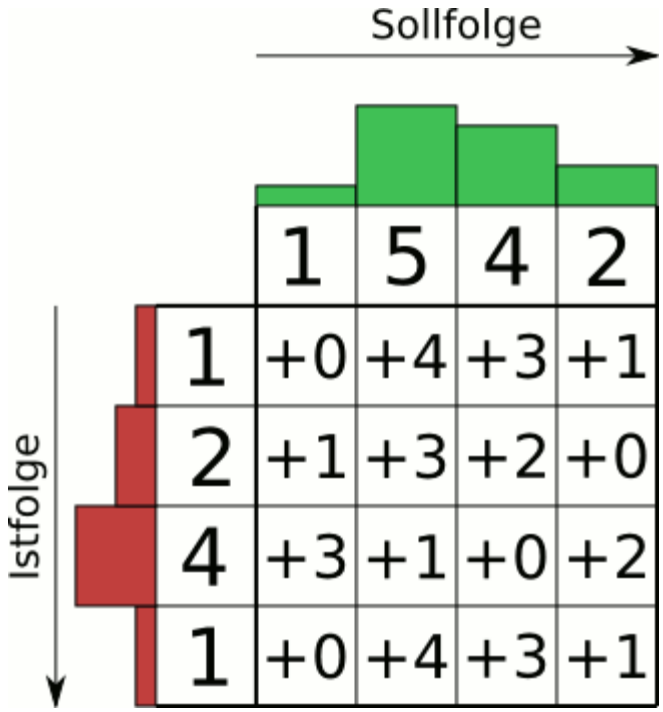
Algorithms for recognize dynamic gestures

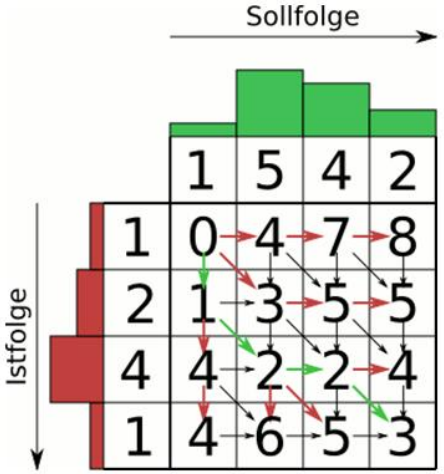
- **Hidden Markov Models**
  - Handling with states and their transition probability
- **Dynamic Time Warping**
  - Algorithm for measuring similarity between two temporal sequences which may vary in time or speed
  - Also used for speech recognition

Dynamic time warping

- Any data which can be turned into a linear sequence can be analyzed with DTW







OpenCV

- Library for real time Computer Vision
- Many purposes → Gesture Recognition
- Released under BSD license
- Supports C++, C, Python, Matlab and Java
- For Windows, Linux, Mac OS, iOS and Android

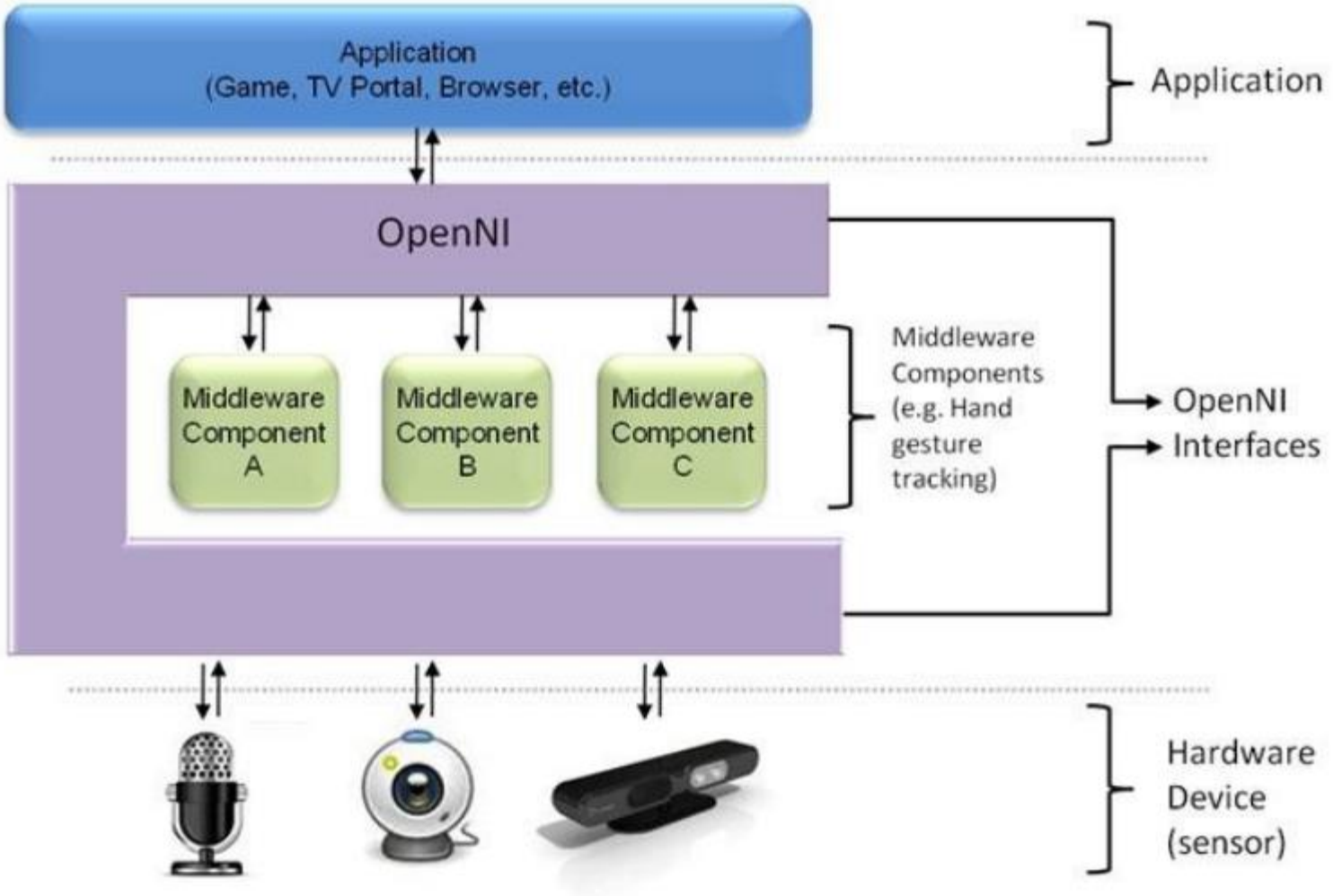
MATLAB

- OpenCV alternative
- High-level scripting language
- Really easy to use
- Standard MATLAB version costs ≈ 1500 €
- OpenCV is free (BSD license)
- Easier to debug

Accessing the Kinect from PC

- OpenNI Framework
- Open source SDK
- Used for the development of 3D sensing
- With the OpenNI SDK you can use the Kinect on the PC
  - Programming your own applications
  - Use a lot of algorithms for body-tracking

OpenNI SDK


Application
(Game, TV Portal, Browser, etc.)
Application
OpenNI
Middleware Component A
Middleware Component B
Middleware Component C
Middleware Components (e.g. Hand gesture tracking)
OpenNI Interfaces
Hardware Device (sensor)

Thank you