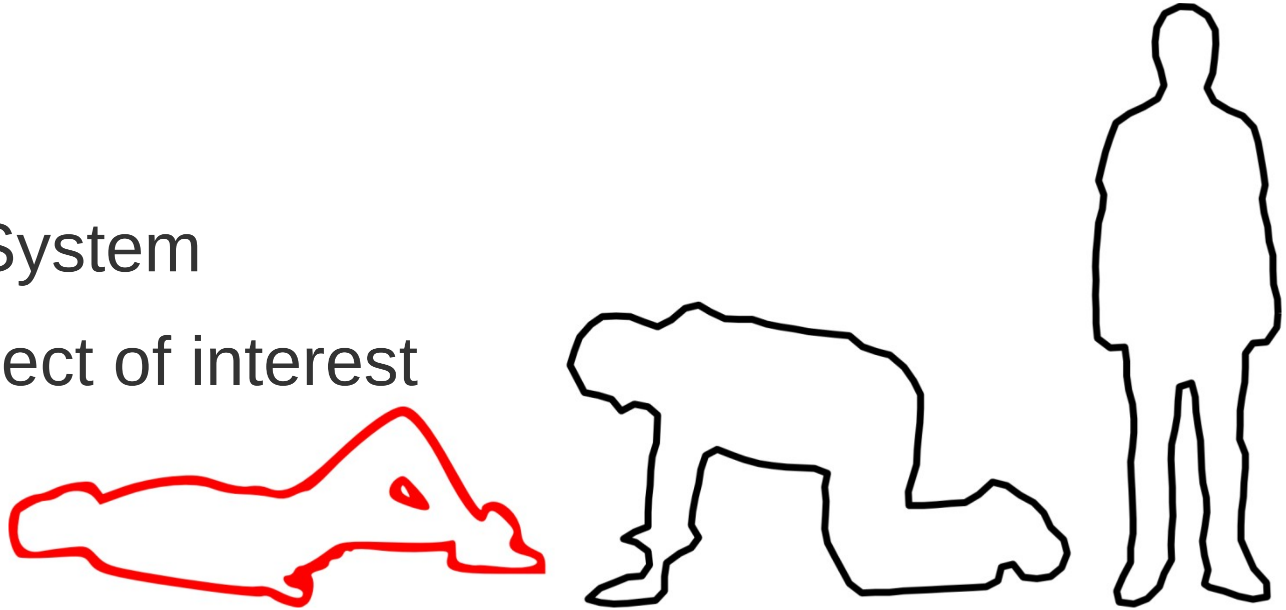
Human Fall Detection by Mean Shift Combined with Depth Connected Components

Michal Kepski and Bogdan Kwolek – Rzeszow University of Technology

Depth is very useful cue to achieve reliable fall detection since humans may not have consistent color and texture but must occupy an integrated region in space. In this work we demonstrate how depth images are extracted by low-cost Kinect device. The person undergoing monitoring is extracted through mean-shift clustering. The fall alarm is triggered on the basis of the distance of the person's gravity center to the altitude at which the Kinect is placed. The experimental results indicate high effectiveness of fall detection in indoor environments and low computational overhead of the algorithm.

Agenda

- Fall detection problem
- Approaches
- Our approach
- The Fall Detection System
- Extraction of the object of interest
- Experiments
- Summary



Fall detection

- Fall Detection isolates falls from activities of daily living (ADLs).

- The goal of fall detection technology is to detect the fall occurrence as soon as possible and generate an alert.

Fall detection: facts

- High percentage of injury-related hospitalizations for seniors are the results of falls.
- From 20 to 30 percent of those who have fallen have medium to severe injuries.
- Half of those, who have fallen can not get up without help.

Fall detection: primary challenges

- Reach high performance of fall detection
- Reduce number of false alarms
- Generate alarm as quickly as possible
- Preserve user privacy

Fall detection: approaches

Since falls are usually characterized by larger acceleration compared with ADL, existing solutions mainly use accelerometers and gyroscopes for detection.

- Several ADLs have similar kinematic motion patterns with real falls (false alarms).
- Inadequate to be worn during the sleep.

Fall detection: approaches

Attempts have been made to detect falls using vision system, consisting of: single camera, multiple cameras or omnidirectional cameras.

- CCD-camera-based solutions require time for installation, camera calibration and they are not generally cheap.
- Can not work in night-light or low light conditions.

Our approach

- Vision device – Microsoft Kinect
- To preserve user privacy utilize only depth images
- Use segmentation and tracking algorithm with low computational overhead
- Run system on embedded platform PandaBoard ES

Vision device – Microsoft Kinect

- Combines depth and RGB camera
- IR projector and camera makes a stereo pair
- Measurement of depth as a triangulation process:
  - the laser source emits beams captured by camera
  - captured pattern is correlated against reference pattern
  - disparity values are obtained
- Distance from device could be calculated:

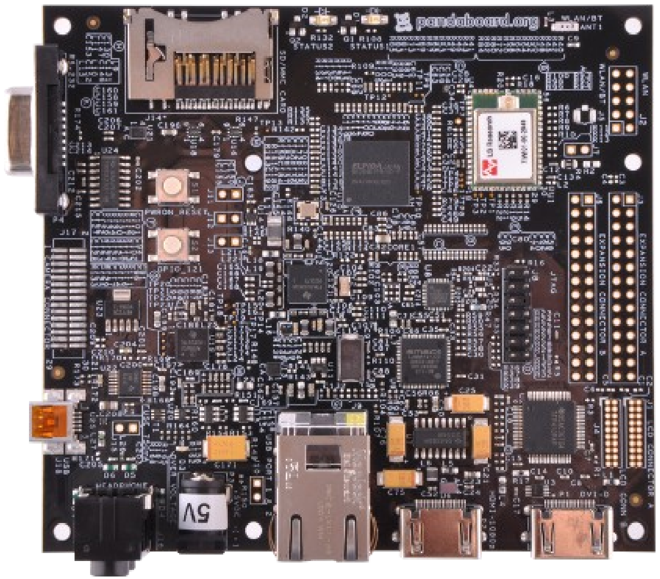
$$Z_k = \frac{Z_o}{1 + \frac{Z_o}{fb} d}$$

Vision device – Microsoft Kinect

- Set of API for communicating with the device is provided (OpenNI)
- API for implementing Natural-Interaction User Interface is provided (NITE) which allows:
  - full body tracking in 3D
  - hand point tracking
  - gesture recognition
- NITE is avalible only for x86/64 architecture

Mobile platform – PandaBoard ES

- ARM architecture
- Dual-core ARM Cortex-A9 1.2 GHz
- 1 GB RAM
- Linux OS
- Dimensions: 114.3 x 101.6 mm



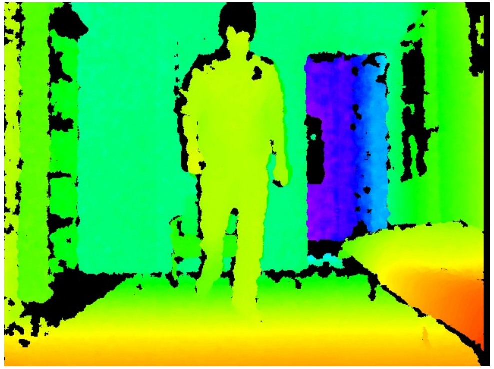
Mean shift

- Nonparametric estimator of probability density
- Treats points in d-dimensional feature space as an empirical probability density function
- Dense regions (clusters) correspond to the modes of underlying distribution
- Does not require prior knowledge of the modes or clusters

$$m(x)=\frac{\sum_{k=1}^{n} x_k\, g\left(\left\|\frac{x-x_k}{h}\right\|^2\right)}{\sum_{k=1}^{n} g\left(\left\|\frac{x-x_k}{h}\right\|^2\right)}-x$$

Mean shift

Extraction of the Object of Interest:

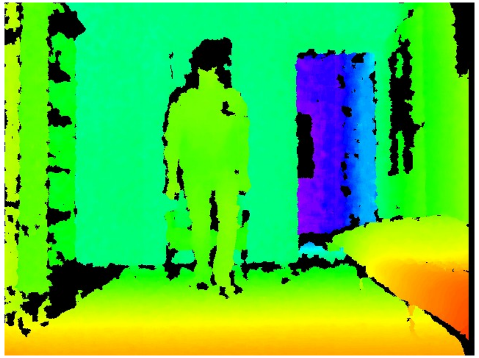


Center of gravity calculation:

$$X_k = -\frac{Z_k}{f}(x_k - x_o + \delta x) \quad Y_k = -\frac{Z_k}{f}(y_k - y_o + \delta y) \quad c(x,y) = \left(\frac{\sum_{i=1}^{n} X}{n}, \frac{\sum_{i=1}^{n} Y}{n}\right)$$

Mean shift

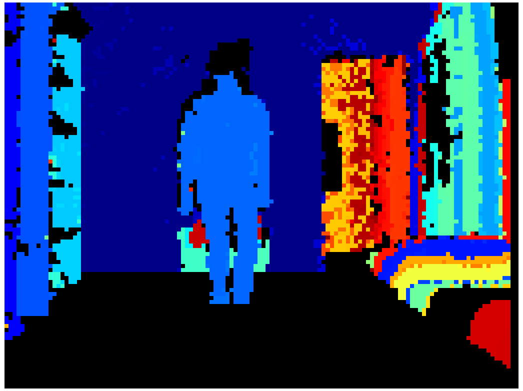
In some images the person is represented by several components:



Images should be refined by connected component operation.

Connected components

Connected component algorithm aims to connect at low computational cost the neighboring depth segments possessing similar depth.

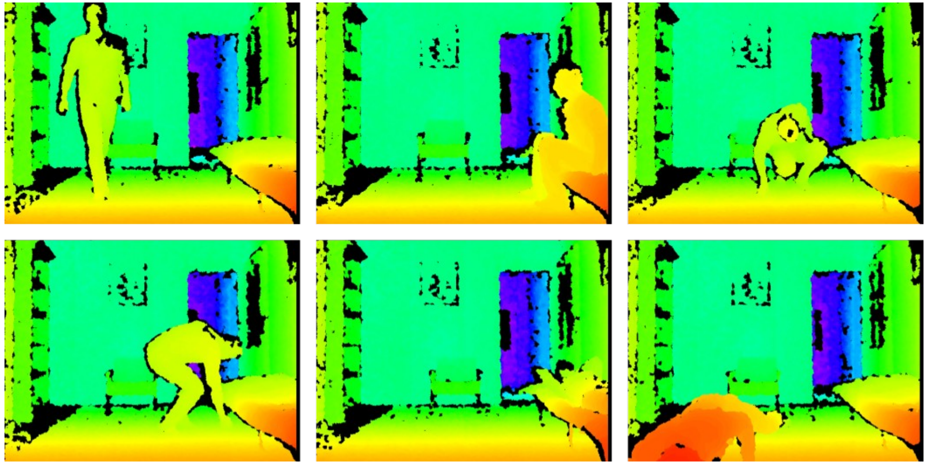


Experiments

- Five volunteers with age over 26 years attended in evaluation of our developed algorithm and the system.
- Each individual performed three types of falls, namely forward, backward and lateral at least three times.
- Each individual performed ADLs: walking, sitting, leaning down, crouching down, picking up objects, lying on a bed.

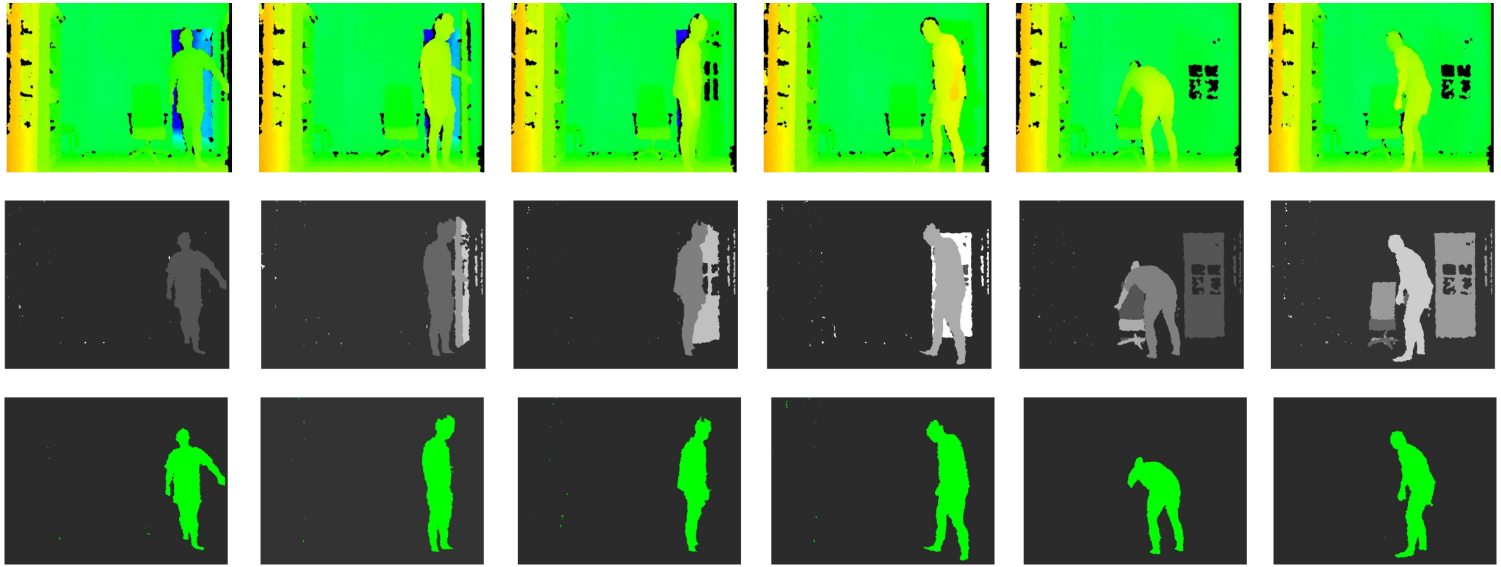
Experiments

Falls and non-fall events (ADLs) during system evaluation:



Experiments

Sequence of images with change of furniture setting:



Results

All intentional falls performed in home towards the carpet were detected correctly.

The system correctly detected seventeen falls of the eighteen falls in the gym towards the mattress.

fall	sitting down	crouching down	walking	lying in a bed	picking up objects
27/28	23/25	23/25	25/25	12/12	25/25

Summary

- In this paper we demonstrated how to achieve reliable fall detection using Kinect.
- A depth connected component algorithm is used to extract the person in sequence of images.
- Systems permits unobtrusive fall detection and preserves privacy of the user.