Real-time motion and activity recognition

Seminar: Post-Desktop User Interfaces

Tim Hemig and Moritz Wittenhagen

November 9, 2006

Contents

Sensor overview

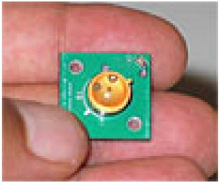
- Camera (lab conditions, absolute position with teamwork)
- Acceleration sensor (heading & orientation)
- Magnetic sensor (orientation)
- Gyroscope - angular rate sensor (orientation)
- Electromyogram (EMG) (force, frequency)

Pro & Contra

- Camera: High accuracy - but lab conditions and problems with light/markers
- Acceleration: Good heading information - but calibration & confusion with too fast movements
- Magnetic: absolute attitude without calibration - but sensible to (electro-)magnetic noise
- Gyroscope: Good attitude - but no information about position
- Electromyogram: good frequency information - but no accurate magnitude

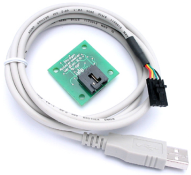
Camera

- High Speed Digital IP-Camera (500 FPS - 1280x1024)
- Up to 12m capture area with 6 cameras (stage)
- processing speed is enough for 2 actor realtime-performance (on a current gaming PC)
- cameras already prepare data for interpretation
- very small reflectors



Acceleration

- commercial field not limited to industrial products
- almost all scales available
- $+$-1.5g up to $+$-10g with around $\frac{600 steps}{1g}$ for miniaturised sensors
- Human peak acceleration: 12g



Magnetic

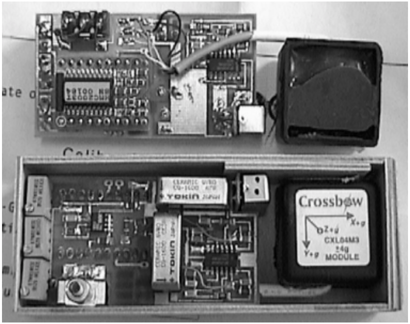
- HAL-sensor in general, very accurate
- determines absolute orientation related to Earth's Magnetic field
- from 120 micro-Gauss to 6 Gauss (Earth has 250 milli-Gauss)

Electromyogram

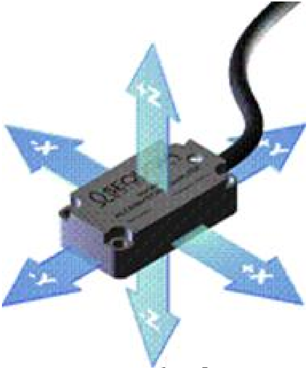
- Measures electrical current associated with muscular action. Does not measure movement directly.
- Not influenced by gravity.
- Contact resistance is a significant variable.
- Displacement, velocity and power cannot be obtained.
- Sensitive to remote muscle activity and interference.

Combined sensor types

- Magnetic Angular Rate Acceleration sensor by Bachmann/McGhee
- 3 magnetic sensors
- 3 angular rate
- 3 acceleration sensors
- all sensors are orthogonal arranged
- (size 10.1 x 5.5 x 2.5 cm)



Best choice



- Mostly acceleration data is used
- No movement: measures orientation
- Movement: measures heading and distance ($s = \frac{1}{2}at^2$)
- Most efficient sensor

Motion/Activity Recognition Comparison

Activity Recognition	Motion Recognition
series of movements	only one movement
few sensors needed	lots of sensors
classification	classification, comparison and transfer

Contents

What is Activity Recognition?

- Focused on the purpose of a movement and not the movement itself
- Usually done with accelerometer data
- Classification problem
- Purpose: Use the context information to build smart devices

The Approach to Activity Recognition

1. Gather data for activity templates
2. Train classifier from the collected data
3. Use the classifiers on "real" data

Training Data

Data source

- Mostly accelerometer data
- Several positioning possibilities
- Example: Sensor placement [Bao, Intille, PERVASIVE '04]
- Wireless data transfer

Collection of Training Data

- Laboratory training data
  - Tell subjects what to do and record the sensor data
  - Easy to get
  - Not necessarily good because subjects may behave different when being aware of what they are doing

Collection of Training Data

- Laboratory training data
  - Tell subjects what to do and record the sensor data
  - Easy to get
  - Not necessarily good because subjects may behave different when being aware of what they are doing
- Naturalistic training data
  - gathered by watching subjects performing their everyday activities
  - subjects are not aware of them being watched
  - very hard to achieve

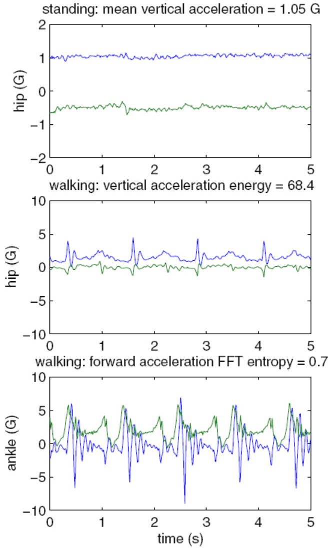
Collection of Training Data

- Laboratory training data
  - Tell subjects what to do and record the sensor data
  - Easy to get
  - Not necessarily good because subjects may behave different when being aware of what they are doing
- Naturalistic training data
  - gathered by watching subjects performing their everyday activities
  - subjects are not aware of them being watched
  - very hard to achieve
- Semi-naturalistic training data
  - Subjects get a list of activities to perform during the day
  - The collected data is self-described by the person
  - Much closer to naturalistic data then lab-gained data

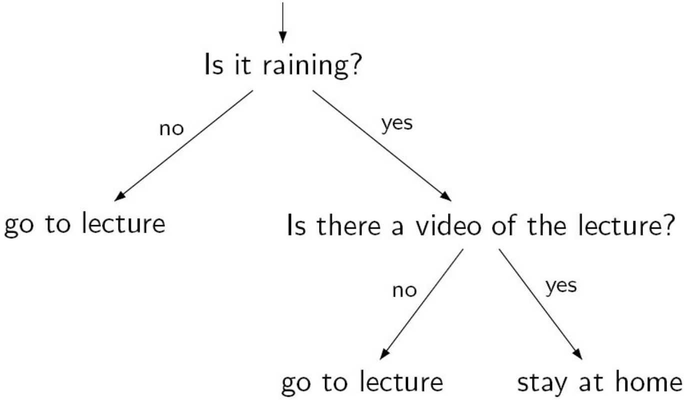
Classifiers

Some classifiers

- Base-level
  - Naive Bayes
  - k-nearest neighbor
  - Decision trees

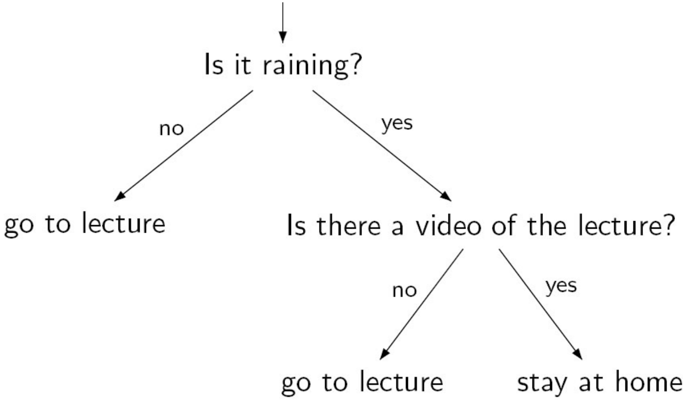


Decision Trees



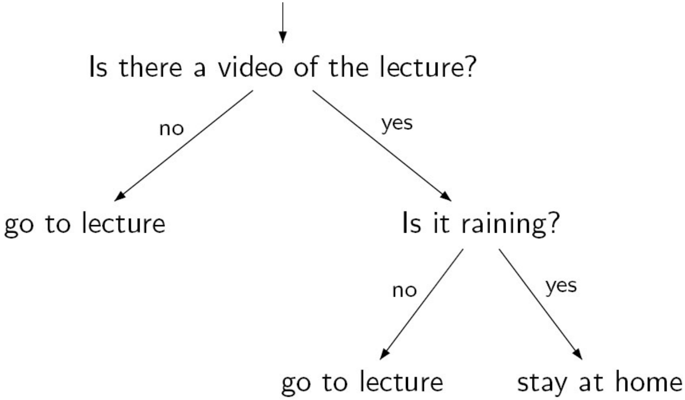
Decision Trees

Feature	Occurrence
Rain	20%
Videos	50%



Decision Trees

Feature	Occurrence
Rain	80%
Videos	50%



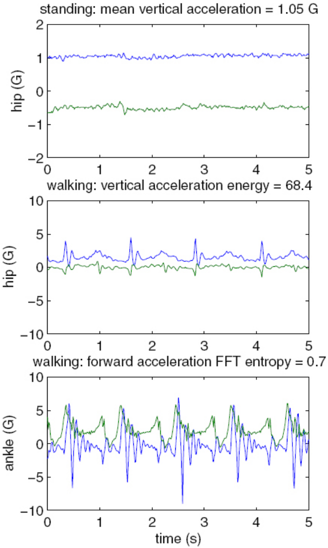
ID3

- ID3 is an algorithm used to train decision trees
- It computes the quality of the questions on a certain node and chooses the best one
- Quality is measured by entropy

$$H(X) = -\sum_{i=1}^{|Z|} p_i \cdot \log_2(p_i)$$

Some classifiers

- Base-level
  - Naive Bayes
  - k-nearest neighbor
  - Decision trees
- Meta-level
  - Boosting
  - Bagging
  - Plurality voting
  - Meta Decision Trees (MDTs)



Studies

Example Study

- Study by Ling Bao and Stephen Intille
- Activity Recognition from User-Annotated Acceleration Data
- Focus on:
  - Good data acquisition
  - What is the best classifier?

Example Study (Setup)

- 20 different activities were considered
- Data of 20 subjects was used
- The data came from 5 bi-axial accelerometers ($\pm$ 10g)
- Subjects were responsible for annotating the data themselves (semi-naturalistic)
- Classifiers were trained using two different protocols
  - Train the classifier with data from all subjects but one and test it on the remaining subject's testing data
  - Train the classifier with the subject's data and test it with the subject's own testing data

Example Study (Results)

Activity	Accuracy	Activity	Accuracy
Walking	89.71	Walking carrying items	82.10
Sitting & relaxing	94.78	Working on computer	97.49
Standing still	95.67	Eating or drinking	88.67
Watching TV	77.29	Reading	91.79
Running	87.68	Bicycling	96.29
Stretching	41.42	Strength-training	82.51
Scrubbing	81.09	Vacuuming	96.41
Folding laundry	95.14	Lying down & relaxing	94.96
Brushing teeth	85.27	Climbing stairs	85.61
Riding elevator	43.58	Riding escalator	70.56

- Decision trees were the most accurate classifier (84% accuracy rate)
- Some activities are easily confused with other activities
- The leave-one-out training policy yielded better results
- User-specific training may be required for some activities

Other Studies (Results)

- Real-Time Motion Classification for Wearable Computing Applications [DeVaul, Dunn, MIT '01]
  - You can rely on very short FFT windows for "easy" questions
  - More complicated questions can be answered by classifying the classification results
- Activity Recognition from Accelerometer Data [Ravi, Rutgers '05]
  - For some activities it is enough to use one triaxial accelerometer
  - Plurality voting is the best meta-classifier
- Predicting Human Interruptibility with Sensors [Hudson, Carnegie Mellon '03]
  - Speech detectors are a possibility to determine a person's interruptibility

Possible Future Studies

- How much improvement can be expected using user-specific training data?
- What is the best count and position for sensors?
- Are accelerometers the best data source?

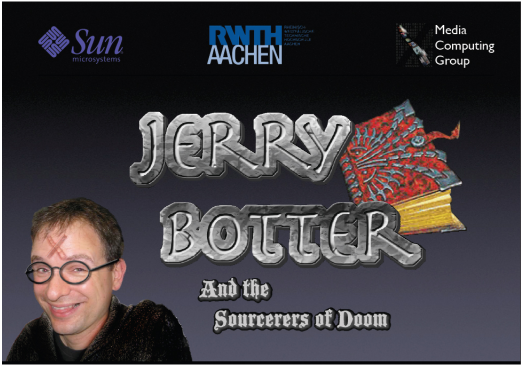
Conclusion

- Activity recognition is a classification problem
- The way of gathering training data is important
- Studies indicate that decision trees are the best "basic" classifiers
- Plurality voting is the best meta classifier
- Classification works fast and well on "easy" questions
- More complex questions might be answered after user-specific training

Contents

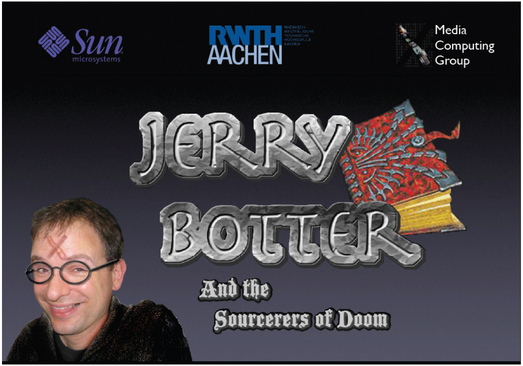
What does motion recognition do?

- Analyse sports movement or synchronous performance
- Transfer motion to an avatar
- User interface for ubiquitous computing

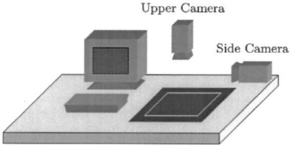


What does motion recognition do?

- Analyse sports movement or synchronous performance
- Transfer motion to an avatar
- User interface for ubiquitous computing
- And a lot more you did not think about yet...

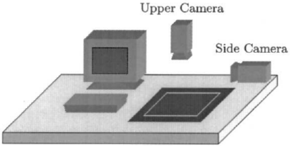


Approach to motion recognition



- Humans can do that with their eyes
- Optical sensors (cameras) would be the translation
- Cameras can track areas of special color or contrast
- Calibrated cameras can recognize exact movement in lab conditions

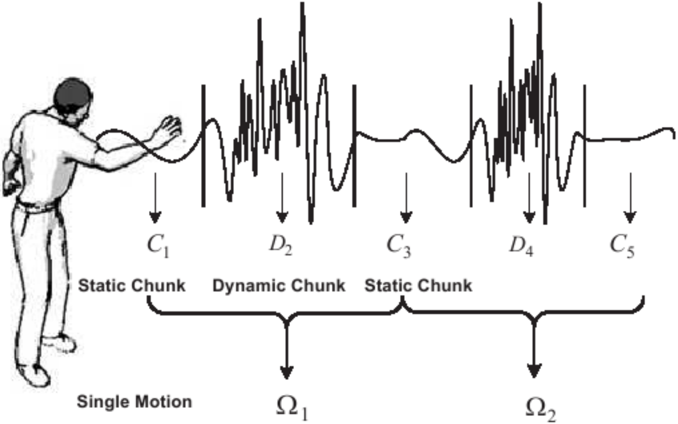
Approach to motion recognition



- Humans can do that with their eyes 
- Optical sensors (cameras) would be the translation
- Cameras can track areas of special color or contrast
- Calibrated cameras can recognize exact movement in lab conditions
- Under real conditions body centric sensors are necessary
- Sensors for magnetism, acceleration and angular rate

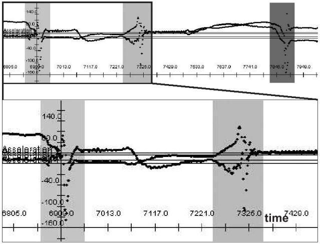
Combining all sensors

- sensor data can be transferred to absolute orientation of body parts
- all orientations on stable values are defining one posture
- dynamic values indicate a motion, a gesture



Detecting motion

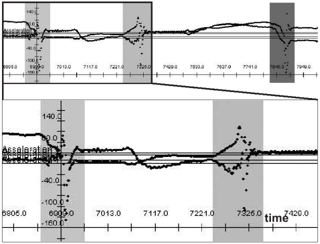
Motion detection by Kwon&Gross (ETHZ)



- Motion chunks are detected by high changes in values
- using Hidden Markov Model (Pattern recognition)

Detecting motion

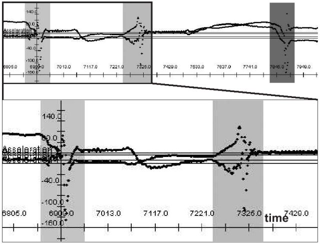
Motion detection by Kwon&Gross (ETHZ)



- Motion chunks are detected by high changes in values
- using Hidden Markov Model (Pattern recognition)
- Difference between starting value and ending value indicate a characteristic

Detecting motion

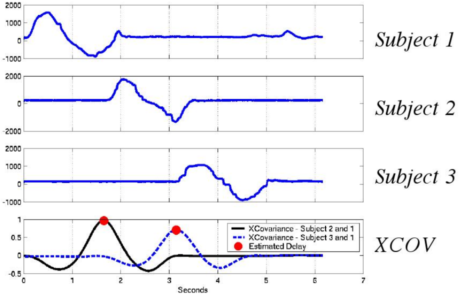
Motion detection by Kwon&Gross (ETHZ)



- Motion chunks are detected by high changes in values
- using Hidden Markov Model (Pattern recognition)
- Difference between starting value and ending value indicate a characteristic
- Stochastic process has to be trained

Comparing motion

- Comparing data for delay of similar movement (Aylward & Paradiso - NIME '06, pages 134-139)
- Cross-covariance used on small time windows of the signals
  $(f \star g)(x) \equiv \int f^*(t) g(x+t)\, dt$



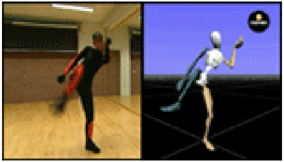
- Sensor data as signal data for review by expert.

Transferring motion

- Transformation does not need any interpretation, it is just rescaled and processed into an 3D Engine

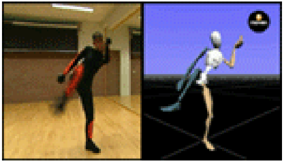
Transferring motion

- Transformation does not need any interpretation, it is just rescaled and processed into an 3D Engine
- Full body suits with acceleration sensors - (VIDEO)



Transferring motion

- Transformation does not need any interpretation, it is just rescaled and processed into an 3D Engine
- Full body suits with acceleration sensors - (VIDEO)



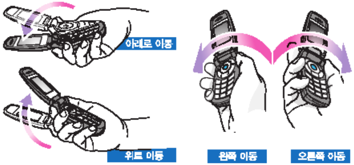
- Transferring motion to animated characters

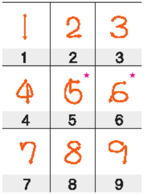
Detecting other Motion

- Not only human motion can be detected

Detecting other Motion

- Not only human motion can be detected
- Interaction with e.g. mobile phones by moving the phone
- User can controll almost all features





Interesting projects

- Training Avatar for martial arts

Fig. 1. Test subjects wearing the wired sensor boxes while performing (expert on top, amateur below)

Interesting projects

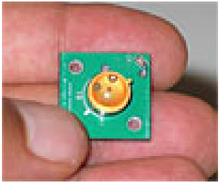
- Training Avatar for martial arts
- The "Moven" motion capture suit by Xsens (Enschede, NL)

Fig. 1. Test subjects wearing the wired sensor boxes while performing (expert on top, amateur below)

Interesting projects

Fig. 1. Test subjects wearing the wired sensor boxes while performing (expert on top, amateur below)

- Training Avatar for martial arts
- The "Moven" motion capture suit by Xsens (Enschede, NL)
- "Motion Captor" by Meta Motion



Interesting projects

- Mobile phones with motion detecting interface by Samsung (Korea)

Interesting projects

- Mobile phones with motion detecting interface by Samsung (Korea)
- "Sketch Furniture" by Front

Video "FRONT"

Future work in motion recognition

- Accuracy of all sensors can be improved
- In general sensors are small enough, but smaller is better
- More implementation work (Frameworks)

Conclusion

Motion recognition

- consists of transferring, comparing and detecting motion
- generally sensors meet necessary requirements
- is commonly used in commercial applications
- helps to create cool projects