Neurogroup

Towards the Adaptive Identification of Walkers: Automated Feature Selection Using Distiction-Sensitive LVQ

Jaakko Suutala and Juha Röning

e-mail: {jaakko.suutala, juha.roning}@ee.oulu.fi

Intelligent Systems Group

University of Oulu, Finland

Infotech Oulu

Neurogroup

Outline

- Introduction
- EMFi Floor
- Footstep Data
- DSLVQ
- Experimental Results
- Conclusions

Infotech Oulu

Neurogroup

Introduction

- What we have done?
  - Recognizing walkers from the footstep profiles achieved with a pressure sensitive floor
    - A 100 square meter pressure sensitive floor was used
    - Test classifications included footsteps from eleven walkers
    - An automatic feature selection method applied to identification
- Identification and feature selection
  - A distinction-sensitive Learning Vector Quantization
    - Based on the standard LVQ classifier
    - A weighted distance metric and single feature relevance measurement added
    - Non-informative features get small weight values and are discarded in distance calculation

UNIVERSITY of OULU

Neurogroup

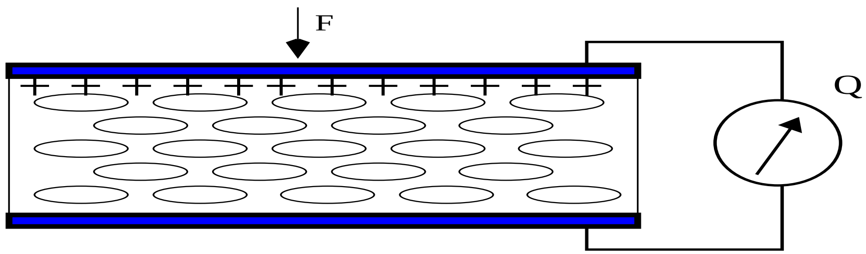
Introduction(2)

- A part of research on intelligent environments: to learn and to react to the behaviour of occupants
  - Hidden sensory system provides a natural and non-disturbing way to use a personal profile
- Applications
  - Monitoring hazardous situations
  - Surveillance systems
  - Helping child care
- Adaptive online identification system
  - To automatically correct small changes
  - To detect new unknown persons and learn their behaviour
- HOW TO AUTOMATICALLY DETECT THE MOST IMPORTANT FEATURES AVAILABLE?

Infotech Oulu

EMFi Floor

- ElectroMechanical Film (EMFi)
  - A thin, flexible, low-price electret material
  - Consists of cellular biaxially oriented polypropylene film coated with metal electrodes
  - It is possible to store a large permanent charge in the film by corona method using electric fields
  - An external force affecting on the EMFi's surface causes a change in the films thickness resulting a charge between the conductive metal layers
    - This charge can be detected as a voltage, which describes the changes in the pressure affecting the floor



EMFi Floor(2)

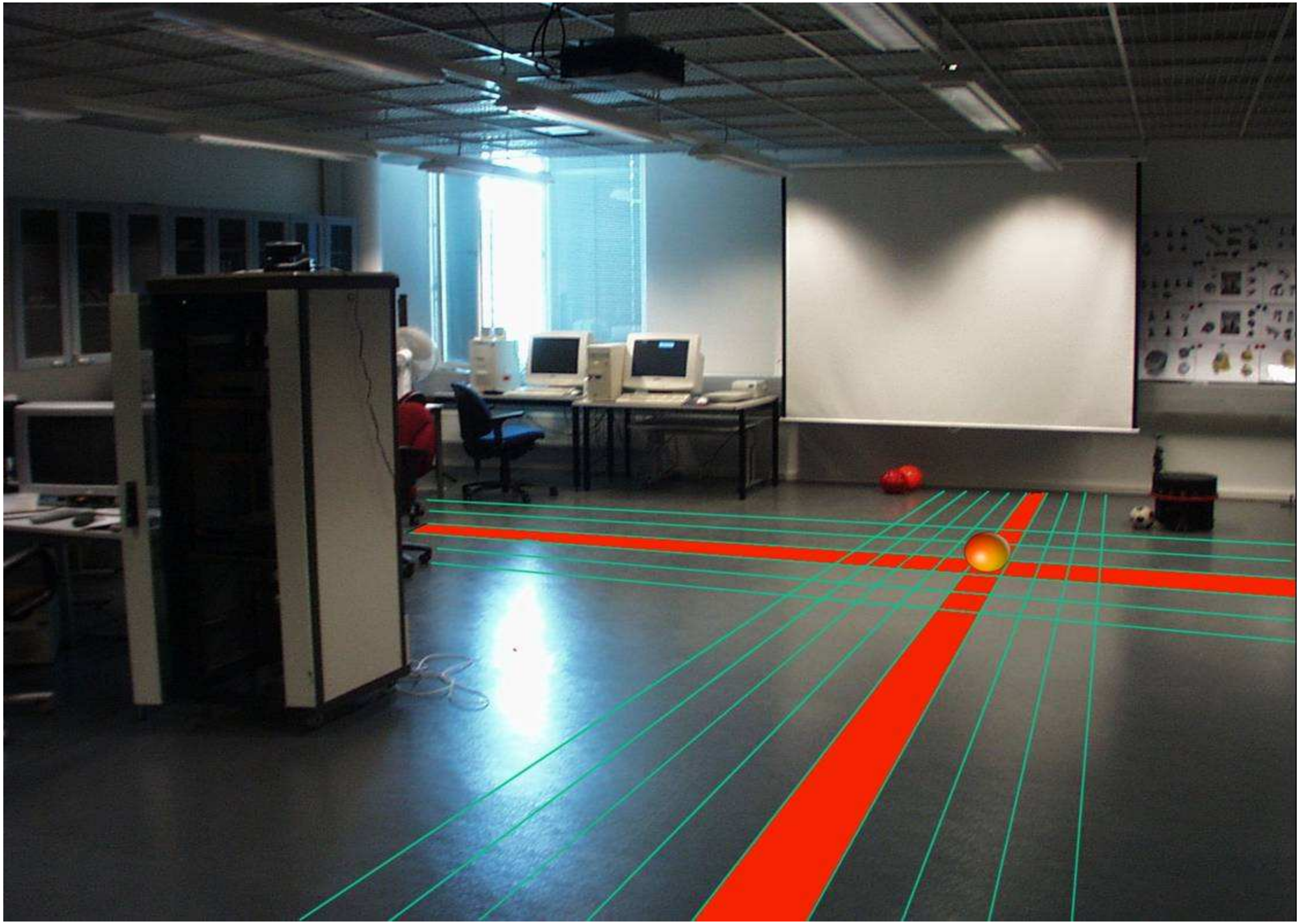
- Floor setting
  - In our research laboratory EMFi material is placed under the normal flooring
  - Consists 30 vertical and 34 horizontal EMFi sensor stripes, 30 cm wide each
  - Advantages
    - Number of wires
    - Number of channels to process
  - Disadvantages
    - Tracking multiple persons
    - To get “good quality” footsteps for identification

Neurogroup

Infotech Oulu

Neurogroup

EMFi Floor(3)



Infotech Oulu

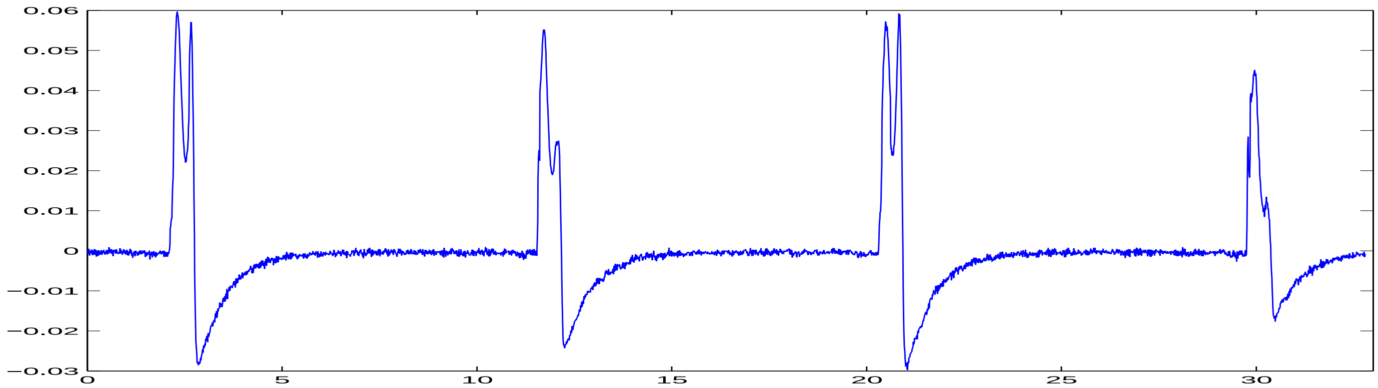
EMFi Floor(4)

- EMFi Data
  - Each 64 stripes produces continuous signal
  - Streamed into a PC from where the data can be analysed in order to detect and recognize the pressure events
  - The analogous signal is processed with National Instruments data acquisition -card (PCI-6033E), sampling rate can be chosen between 0.1 - 1.54 kHz
    - 100 Hz sampling rate is used in these experiments

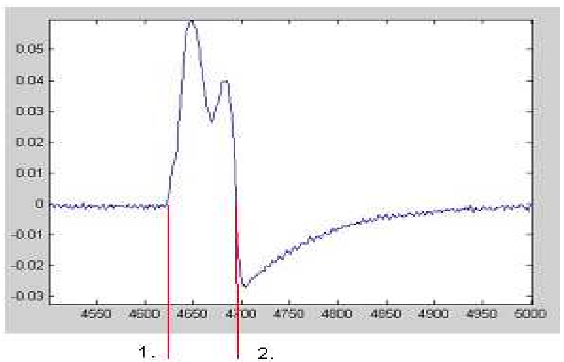
Neurogroup

EMFi Floor(5)

- Raw data



- Segmented footstep



Intelligent Systems Group
Department of Electrical and Information Engineering and InfoTech Oulu
www.ee.oulu.fi/research/neurogroup

Infotech Oulu

UNIVERSITY of OULU

Neurogroup

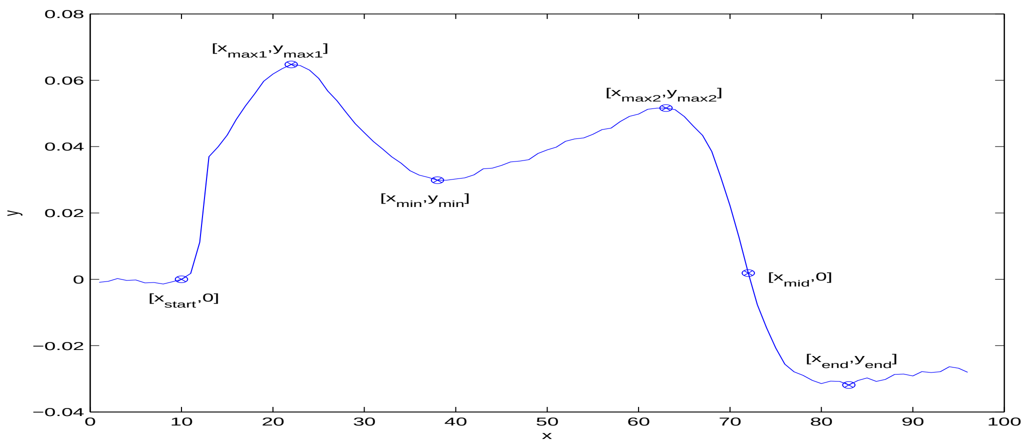
Footstep Data

- Collecting data
  - Footstep data was recorded from eleven different persons
  - The subjects stepped on one particular stripe
  - Collected data contained about 40 steps/person, including steps from both feet
- Pre-processing
  - Finding good-quality steps from noisy data
    - A raw segmentation was made with edge detection using FIR median hybrid filter, convolution, and thresholding
    - Footstep parts from adjacent channels were summed

Infotech Oulu

Footstep Data(2)

- Features
  - Each step was divided in two sections: The heel strike and the toe-off peak
  - Several features was calculated from both spatial and frequency domain
  - Totally 31 features were extracted



DSLVQ

- Learning Vector Quantization (LVQ)
  - A well known statistical distance based classification method
  - Based on piecewise linear class boundaries, which are determined by supervised learning
- LVQ classification
  - Classification is made with a codebook, which contains prototype vectors labeled for each classes
  - Learning algorithm iteratively minimizes the rate of misclassification error by updating the codebook vectors
  - The unknown sample is classified to the closest codebook vector using Euclidean distance

DSLVQ(2)

- LVQ3 training algorithm
  - Two closest codebook vectors m_i and m_j belonging to different and same class as x:

$$(1) \qquad m_i(t+1) = m_i(t) - \alpha(t)[x(t) - m_i(t)],$$
$$(2) \qquad m_j(t+1) = m_j(t) + \alpha(t)[x(t) - m_j(t)],$$

  - Two closest codebook vectors m_i and m_j belonging to same class as x:

$$(3) \qquad m_k(t+1) = m_k(t) - \epsilon\alpha(t)[x(t) - m_k(t)],$$

UNIVERSITY of OULU

Neurogroup

DSLVQ(3)

- Problem of standard LVQ: All features are treated equally in distance metric
- Solution: Weighting of features based on relevance of the single features
  - The non-informative features are assigned with small weight values
  - The informative features are assigned with larger weight values
- Using a weighted distance metric, the non-informative features are discarded !
- Advantages:
  - Feature selection is automatic
  - Feature selection can be moved from the pre-processing phase to the training phase of classifier

Infotech Oulu

DSLVQ(4)

- The weighted Euclidean distance

$$w_{dist}(\mathbf{w},\mathbf{x},\mathbf{m}) = \sqrt{\sum_{n=1}^{N}[\max(0,w_n)(x_n - m_n)]^2}, \tag{4}$$

- Weight adaptations:

$$\mathbf{w}(t+1) = norm(\mathbf{w}(t) + \alpha(t)[\mathbf{nw}(t) - \mathbf{w}(t)]), \tag{5}$$

$$nw_n(t) = norm\Big(\frac{d_{i_n}(t) - d_{j_n}(t)}{\max(d_{i_n}(t), d_{j_n}(t))}\Big) \tag{6}$$

$$norm(\mathbf{y}) = \frac{\mathbf{y}}{\sum_{n=1}^{N}|y_n|} \tag{7}$$

UNIVERSITY of OULU

Neurogroup

Experimental Results

- Feature selection
  - kNN classifier
    - 2/3 of data set was chosen for classification model, 1/3 for test the model
    - Finding best subset of features using forward-backward-search, minimizing the classification error in the test set
    - 13 features were chosen from the spatial domain
    - For example: x_{max1}, y_{max1}, x_{min}, y_{min}, x_{max2}, y_{max2}

Infotech Oulu

Neurogroup

Experimental Results(2)

- DSLVQ
  - LVQ codebook size: 18 prototype vectors/class
  - Learning algorithms: OLVQ1 (initialization), LVQ1, LVQ3, DSLVQ (fine tuning)
  - The data sets of 13 and 31 features were used
  - Features were normalized between 0 and 1 at the beginning of training

N	w_{ave}	w_{max}	w_{min}	nof $> w_{ave}$
13	0.0769	0.1407	0.0379	7
31	0.0323	0.0487	0.0056	17

Infotech Oulu

Experimental Results(3)

Neurogroup

- The identification results
  - The overall recognition accuracies of 11 walkers
  - Results are averages of 5-fold cross-validation (standard deviation in parentheses)
  - DSLVQ shows best results in both data sets

N	LVQ1	LVQ3	DSLVQ
13	66.8% ($\pm$5.4)	67.4% ($\pm$5.4)	70.2% ($\pm$5.7)
31	65.8% ($\pm$5.0)	66.5% ($\pm$4.7)	69.4% ($\pm$6.4)

Infotech Oulu

Neurogroup

Conclusions

- Experiments on applying automated feature selection to footstep identification were reported
- Method increases adaptiveness: the best subset of features can be chosen automatically during the training of a classifier
- Future plans
  - More analyses: For example, how do person wearing different shoes, backpacks etc. affect the selection of relevant features?
  - How well DSLVQ can adapt to a new situations? For example, when new person is detected
  - Adaptive real-time learning and a recognition application for tracking and identification
  - Detecting mobile robot movements from the floor and the co-operation with occupants

Infotech Oulu