Zorana Banković

Using Self-organizing Maps for Intelligent Camera-based User Interface

ETSI TELECOMUNICACIÓN B-105
LST Laboratorio de Sistemas Integrados

Departamento de Ingeniería Electrónica
Universidad Politécnica de Madrid

Objective

- Problem:
  - Human-Machine Interface
    - Growing fast
    - High importance in all technological systems
    - Most of the solutions based on complex methods
      - Require high-resource devices
      - Not viable for embedded systems
- Solution:
  - Low-cost
  - Camera-based gesture interface

Outline

- **Introduction**
- Motivation
- Implementation Details
- Evaluation
- Conclusions and future work

Introduction

- Human-machine Interface
  - Transparent
    - The user does not have to adapt to special conventions or rules;
    - The environment should be the one to adapt to the natural way of user interaction.
  - Hand gestures: the most natural and comfortable way
    - Most of the solutions consume significant resources
    - Embedded systems (camera) exhibit limited resources

Introduction

- Camera-based gesture interface
  - Model gestures capturing their temporal properties
  - Significantly reduces storage requirements
    - Appropriate for implementation in embedded systems
  - Self-organizing maps for gesture classification

Agenda

- Introduction
- **Motivation**
- Implementation Details
- Evaluation
- Conclusions and future work

Motivation

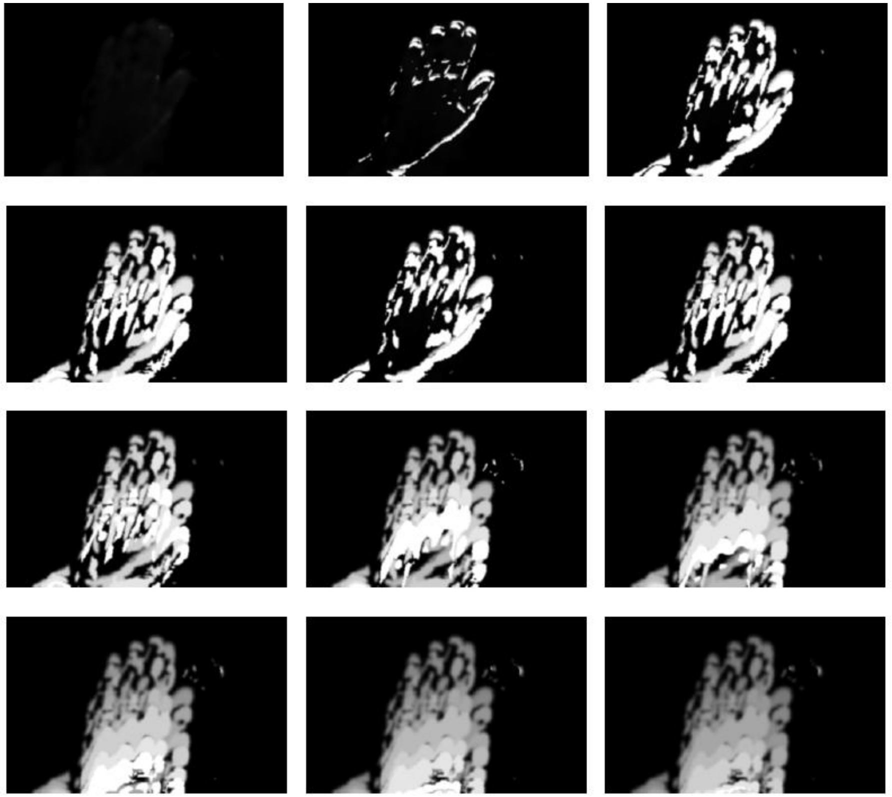
- Common SOM solutions
  - Two stages: the first stage captures the temporal properties → two learning algorithms
  - Standard characterization: trajectory of the hand, resultant direction of the movement, velocity of the movement…
    - Additional computational overhead

Agenda

- Introduction
- Motivation
- **Implementation Details**
- Evaluation
- Conclusions and future work

Implementation Details – Gesture Characterization

- Gesture: set of frames of variable size

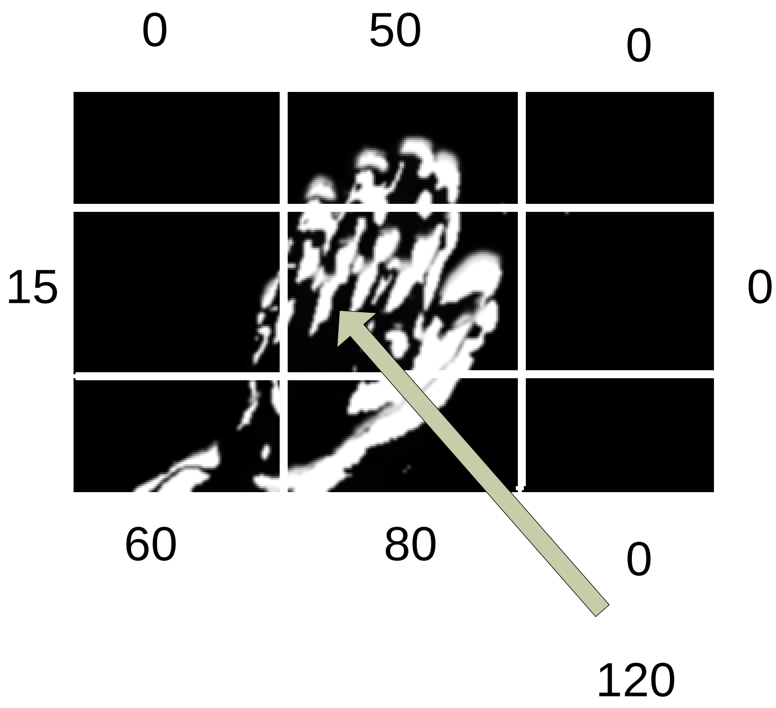


Implementation Details - Gesture Characterization

- Divide each frame into *n* x *n* parts
- Aassign to each part a value that corresponds to its luminosity (0 – 255)
- Characterize temporal evolution of each part
  - E.g. 0 0 20 40 50 60 70 50 10, sliding window size 3

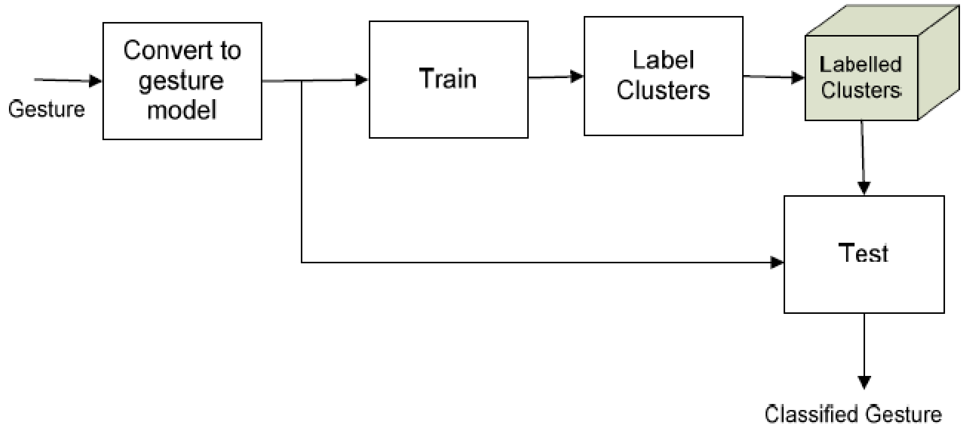
0 0 20	0.16
0 20 40	0.16
20 40 50	0.16
40 50 60	0.16
50 60 70	0.16
70 50 10	0.16

- Number of features not fixed → distance between sequences
  - Distance between two gestures the sum of absolute distances of the parts



Implementation Details - SOM

- Specific part: node update
  - If the node does not contain a feature from a certain input, we add it to the node with the value 0
  - Discard all the features that have at least 100 times smaller value of the maximal feature value of the node
- Label the nodes with the label of the gesture from the set of labeled gestures that is closest to the node according to the distance function



Advantages

- Simplicity
  - Gestures distinguished by clustering → no need to label all the gestures
  - The characterization significantly reduces the memory needed to store a gesture
    - 507kB reduced to 625B (5x5 division of the frame)
- Enables implementation on devices with limited resources

Agenda

- Introduction
- Motivation
- Implementation Details
- **Evaluation**
- Conclusions and future work

Training and Testing Dataset

- Five types of gesture:
  - left-right
  - right-left
  - up-down
  - down-up
  - random gestures labeled as unknown
- 12 persons – 760 gestures
  - 1.08GB of storage space → after characterization 3.11MB

Results

- Testing with both 3x3 and 5x5 frame partitions
- Gestures left-right and right-left → confused with each other

Gesture	Detection Rate (%)
Unknown	88
Down-up	100
Up-down	92
Left-right	13
Right-left	13
Overall	80

Agenda

- Introduction
- Motivation
- Implementation Details
- Evaluation
- **Conclusions and future work**

Conclutions and Future Work

- Low-cost algorithm for gesture classification
- Characterization of gestures that captures temporal properties of gesture
- Detection rate of up to 100% for certain gestures and overall detection of 80% at most
- Future work:
  - Add one more stage of SOM clustering in order to detect the users.

ETSI
TELECOMUNICACIÓN
B-105
LSI
Laboratorio de
Sistemas
Integrados