Adaptation Techniques for Mobile Image Classification
Guidance Systems

• User-based
  • Human-readable labels, audio-guides

• Sensor-based
  • Short-range emitter: RFID, NFC, Infrared
  • Wide-range emitter: Bluetooth, WiFi, GPS (e.g. infSoft, Loopt)
  • Orientation: digital compass (e.g. Layar)

• Vision-based
  • Fiducials: Barcodes, AR markers (e.g. Semapedia)
  • Natural: object recognition (e.g. kooaba)

• Hybrid approaches
  • Sensor-based + vision-based + user-based (e.g. PhoneGuide)
PhoneGuide

Adaptive mobile museum guidance system

1. take foto of exhibit
2. recognize exhibit
3. display information

Erich Bruns, Oliver Bimber

WCI 2009
PhoneGuide - Motivation

- Enhanced museum guidance: hybrid approach
  - Rich multimedia presentations (images, video, audio, text, cg)
- Visitors use their own mobile devices
  - Little / no acquisition and maintenance cost for museums
- Why not labels (numbers, fiducials) [Rohs et al.'05]?
  - Distract overall appearance, more intuitive interaction
- Why not electronic tags (sensors) [Chang et al. ‘08]?
  - Not possible for densely located objects in showcases, maintenance.
- Why not client-server based [Ruf et al.'08]?
  - Advantages of on-device recognition
    - Decentralized, scalable
    - Classification requests processed in parallel, independent of the number of users
    - No sophisticated infrastructure → saves costs
    - Independent of network coverage (“Museum 411”)
PhoneGuide - Motivation

• Why adaptive?
  • Recently upcoming approaches port existing desktop algorithms to mobile devices without taking much into account the benefits of them [Bay et al.’06, Takacs et al.’08]
  • In contrast, we think that if the context of mobile phones are considered, the classification performance can be improved and outperform related traditional approaches.
• Advantages of mobile phones for object recognition applications
  • Location-based (e.g. information are retrieved at/for particular locations, [Paletta et al.’08])
  • User-centered (e.g. visitors can give feedback on recognized objects; visitors’ behavior/pathways give important information)
  • Multi-user system (e.g. multiple devices can interoperate)
Overview

image classification
- objects
- subobjects

adaptation
- pervasive tracking
- P2P communication
- pathway awareness
- P2P localization*

user interface

* in progress

Erich Bruns, Oliver Bimber
Image Classification

training

+ capture images
+ preprocessing
+ feature extraction
+ database storage

database

recognition

+ capture image
+ preprocessing
+ feature extraction

matching

Erich Bruns, Oliver Bimber

WCI 2009
1. Capture video of each exhibit
2. Separate each image into N patches
3. Compute for each patch 40 global color features
4. Train N neural networks for classification

| 0.2442 | 0.2442 | 0.2442 | 0.2442 |
| 0.9535 | 0.9535 | 0.9535 | 0.9535 |
| 0.5342 | 0.5342 | 0.5342 | 0.5342 |
| 0.2343 | 0.2343 | 0.2343 | 0.2343 |
| 0.6211 | 0.6211 | 0.6211 | 0.6211 |

| 0.2442 | 0.2442 | 0.2442 | 0.2442 |
| 0.9535 | 0.9535 | 0.9535 | 0.9535 |
| 0.5342 | 0.5342 | 0.5342 | 0.5342 |
| 0.2343 | 0.2343 | 0.2343 | 0.2343 |
| 0.6211 | 0.6211 | 0.6211 | 0.6211 |

| 0.2442 | 0.2442 | 0.2442 | 0.2442 |
| 0.9535 | 0.9535 | 0.9535 | 0.9535 |
| 0.5342 | 0.5342 | 0.5342 | 0.5342 |
| 0.2343 | 0.2343 | 0.2343 | 0.2343 |
| 0.6211 | 0.6211 | 0.6211 | 0.6211 |

Adaptive Training of Video Sets for Image Recognition on Mobile Phones
E. Bruns and O. Bimber
Journal of Personal and Ubiquitous Computing, 2008

Mobile Phone Enabled Museum Guidance with Adaptive Classification
E. Bruns, B. Brombach and O. Bimber
Subobject Detection through Spatial Relationships on Mobile Phones
Brombach, B., Bruns, E. and Bimber, O.
International Conference of Intelligent User Interfaces (IUI'09), 2009
Subobject Detection

1. Spirally shift search mask around the center of the image
2. For each location, perform classification
3. If subobject was detected, spatial relationships (angles, distances) define search areas for remaining subobjects

Subobject Detection through Spatial Relationships on Mobile Phones
Brombach, B., Bruns, E. and Bimber, O.
International Conference of Intelligent User Interfaces (IUI'09), 2009
1. Spirally shift search mask around the center of the image
2. For each location, perform classification
3. If subobject was detected, spatial relationships (angles, distances) define search areas for remaining subobjects
4. If new subobjects are detected, search area becomes more and more precise

Subobject Detection through Spatial Relationships on Mobile Phones
Brombach, B., Bruns, E. and Bimber, O.
International Conference of Intelligent User Interfaces (IUI'09), 2009
Subobject Detection
Overview

- Image classification
  - Objects
  - Subobjects

- Adaptation
  - Pervasive tracking
  - P2P communication
  - Pathway awareness
  - P2P localization* (in progress)

- User interface

* in progress

Erich Bruns, Oliver Bimber
WCI 2009
User Interface

• After an image was classified, a probability-sorted objects list of result candidates is presented on the device
• User is able to select the correct object through the user interface if the classification has failed
• Mapping of photographed image and real object ID is stored on mobile phone and transferred to the server
• Server retrains neural networks with new images
• Wrong assignments are rejected through clustering

➔ Continuous adaption (supervised learning)
  ➔ Users' preferences
  ➔ Photographed perspectives
➔ Opens opportunities for further adaptation techniques
Overview

image classification
- objects
- subobjects

adaptation
- pervasive tracking
- P2P communication
- pathway awareness
- P2P localization*

user interface
* in progress

Erich Bruns, Oliver Bimber
WCI 2009
1. Pervasive Tracking

- **Motivation**
  - Object recognition rate decreases with an increasing number of objects

- **Idea**
  - Distribute BT beacons for rough location estimation of museum visitors
  - For each location cell, train one neural network
2. Phone-to-Phone Communication

- **Motivation**
  - Global color features of object recognition are variant to illumination
  - Current adaptation process is not in real-time

- **Problem**
  - Retraining neural networks is too time consuming

- **Idea**
  - Exchange classification data through phone-to-phone communication for real-time adaptation (cp. car-to-car communication)
2. Phone-to-Phone Communication

classification successful

P2P communication
(image features)

Erich Bruns, Oliver Bimber

WCI 2009
3. Pathway Awareness

• Motivation
  • Bluetooth tracking gives only a rough estimation of the visitors’ location. High energy consumption
  • The smaller the number of potential result candidates / feature space, the higher might be the classification rate

• Idea
  • Determine the pathways of the museum visitors to derive transition probabilities between pairs of exhibits and combine them with image classifiers
3. Pathway Awareness

• Field experiment for determining pathways (132 visitors)
3. Pathway Awareness

• Pathway: threshold ≥ 2 visitors
3. Pathway Awareness

- Pathway: threshold $\geq 4$ visitors
3. Pathway Awareness

- Pathway: threshold $\geq 6$ visitors
3. Pathway Awareness

• Pathway: threshold $\geq 8$ visitors
3. Segmentation through Pathways

unconsidered exhibits

considered exhibits
4. P2P Localization

• Motivation
  • Localization of museum visitors with Bluetooth beacons is not optimal
    • Maintenance costs
    • Signal range inconsistent (depends on battery life)
  • Pathway classification can be improved through localization

• Idea
  • Localization through ad-hoc networks between mobile phones and landmarks
4. P2P Localization

• Current research

\[ R = \text{Bluetooth signal range (\sim 10m)} \]
\[ + \text{distance covered of remote visitor since last classification} \]
\[ + \text{distance covered of current visitor since last synchronization} \]
Evaluation

Classification Rate (%)

Object Recognition
Subobject Recognition*

*Classification rate of subobject recognition without object recognition: 85.9%

Adaptation Approach

Basic Image Classification (perspective invariance) +
Phone-to-Phone Communication (illumination invariance) +
Pathway Segmentation (localization) +
Pathway Prediction (localization) +
Bluetooth Localization (localization) +

50.67%  43.53%  52.05%  61.26%  71.31%  78.82%  67.71%
Thank you.

www.uni-weimar.de/ar/phoneguide

Thanks to:

Naturkundemuseum Erfurt

Stadtmuseum Weimar

Nokia