Visuelle Perzeption für Mensch-Maschine Schnittstellen

Vorlesung, WS 2009

Prof. Dr. Rainer Stiefelhagen  
Dr. Edgar Seemann

Institut für Anthropomatik  
Universität Karlsruhe (TH)

http://cvhci.ira.uka.de  
rainer.stiefelhagen@kit.edu  
seemann@pedestrian-detection.com
Exam

- Oral exam
- 4 SWS
- Results from programming assignments will be taken into account

- All lectures are relevant for the exam
- Important for a good understanding are:
  - Big picture and inter connections
  - Details (e.g. equations)

- Don’t panic!
Overview

2009-10-19  Einführung
2009-10-23  Basics: Image Processing
2009-10-26  Basics: Image Transformation and 2D Structure
2009-10-30  Pattern recognition
2009-11-02  Computer Vision: Tasks, Challenges, Learning
2009-11-06  Face Detection 1
2009-11-09  Project 1: Intro + Programming Tips
2009-11-13  Face Detection 2
2009-11-16  Project 1: Questions
2009-11-20  Face Recognition 1
2009-11-23  Face Recognition 2
2009-11-27  Head Pose Estimation, Focus-of-Attention
2009-11-30  People Detection 1
2009-12-03  People Detection 2
2009-12-07  Project 1: Presentations, Project 2: Intro
2009-12-11  People Detection 3
2009-12-14  People Detection 4
2009-12-18  Context and 3D Structure
2009-12-21  Facial Feature Detection
2010-01-11  Facial Expression Recognition
2010-01-15  Gesture Recognition
2010-01-18  Tracking I
2010-01-22  Tracking II
2010-01-25  Activity Analysis I
2010-02-05  Activity Analysis II
2010-02-08  Audio-visual Speech Recognition
Goal

- What did you like/dislike?
- Which topics should be added/extended?
- Review the topics of this lecture
  - What are the applications?
  - Which methods can be used?
  - How is it done?
- Preparation for your exam
- Questions
- We'll proceed lecture by lecture
Basics

1. Image Transformations & 2D structure
   - pinhole model
   - Image formation as linear transformation
     - Homogenous coordinates
   - Camera calibration
     - Extrinsic vs. Intrinsic parameters
   - Lenses
     - Focus and focal length
     - Thin lens formula
     - Infinity focus, focus and distance
     - Lens systems and abbreviations
   - Alignment
     - Euclidian, affine, perspective transformation
     - Degrees of Freedom
Basics

2. Image Processing
   - Digital images, sampling artifacts
   - Filtering and convolution (continuous vs. discrete)
     - Kernels (blur, sharpen)
     - Properties of the convolution (linear system)
     - Gaussian averaging
     - Separability
   - Edges
     - Types of edges
     - 1st and 2nd derivatives
     - 1D vs. 2D equations
     - Gradient direction
     - Non-Maximum suppression
     - Thresholding, Hysteresis
     - Laplacian
Basics

- Color
  - Human perception
  - Reflectance model, spectra
  - Additive/subtractive color mixing
  - Color matching process
  - Grassman’s Law
  - Color spaces
    - Primary colors
    - HSV, RGB
    - Color space conversions
3. Pattern Recognition
   - Classification vs. Regression
   - Curse of dimensionality
   - PCA
     - Variance, Covariance-Matrix
     - Redundancy in data
     - Eigenvectors vs. Directions of variance
     - Basis change and projection
     - Assumptions for PCA
     - Algebraic tricks
     - Relation to LDA
   - Bayes Decision Theory
     - Prior/Posterior/conditional probability
     - Bayes Theorem
     - Loss functions
3. Pattern Recognition
   - Perceptron Algorithm
     - Linear separability
     - Convergence
   - Instance-based Learning
     - K-Nearest Neighbors
       - Misleading dimensions
       - Computation time grows with number of training samples
     - KD-Tree
       - Tree construction
       - Search and run-time
   - Ball Trees
Computer Vision Tasks

- Identification vs. Classification
- Segmentation and Localization
- Challenges
  - Occlusion
  - Scale, rotation
  - Robustness
- Sliding window technique

- Training, Validation, Testing
  - Generate Training data: labeling, adding perturbations
  - Cross-Validation
- Generative vs. Discriminative models
- Generalization, Overfitting
- Occam’s razor
Performance Measures

- TP, FP, TN, FN
- ROC
  - True positive rate
  - False positive rate
- RPC
  - Recall
  - Precision
- Other measures:
  - FPPW, FPPI
- Comparing bounding boxes
- Non-maximum suppression
Face Detection

- **Face detection I**
  - Motivation, difficulties, representation
  - Color-based approaches
    - How to model skin-color (parametric, non-parametric)
    - Histogram backprojection vs. Histogram matching
    - Gaussian densities, Mixture of Gaussians
    - Classifiers
    - Postprocessing: Morphological operators

- **Ellipsoid head model detection**
  - And combination with color-based detection
  - Basically a \( \rightarrow \) deformable template
Face Detection II

- Artificial Neural Networks for Face Detection
  - Short repetition of ANNs
  - Preprocessing: histogram equalization
  - Network topology
  - Training

- Viola & Jones approach
  - Haar features
  - Integral image
  - Variant of AdaBoost to select features and to build a strong classifier
  - Classifier cascade for fast processing
Face Recognition I

- Introduction, cognitive issues, history
- Face recognition tasks:
  - open set, closed set, authentication / verification
  - Related metrics
    - Mahalanobis distance
- Feature-based approach
- Eigenfaces
- Fisherfaces
  - Intra-class variation
- Applications
Face recognition II

- Local appearance based approaches
  - Modular Eigenspaces
  - Using DCT, zig-zag scan

- Face Recognition using a 3D morphable model
  - Blanz & Vetter

- Databases & Benchmarks
Head Pose Estimation

- Motivation, relation with focus of attention

- Model-based head pose estimation
  - Needs facial landmark tracking & model

- Head pose estimation with ANNs

- Modeling focus of attention from head pose
  - „who was looking at whom?“
  - Determining whether a robot was addressed or not
People Detection I

- Global vs. Part-Based Approaches
- Levels of supervision
- Contours vs. Colors
- Gradient Histograms
- Interpolation

- Global Approaches
  - HOG (feature computation)
People Detection II

- Silhouette/Chamfer Matching
  - Distance transform
- Advanced silhouette matching (edge orientation, spatio-temporal templates)
- Earth-Mover’s Distance

- Wavelets
  - Mother wavelet, scaling function
  - Basis and projection
  - Wavelet transformation with orthogonal spaces
People Detection III

- Pictorial Structures
  - Body tree decomposition
  - Parsing of body tree

- Local Features
  - Interest points
  - Hessian-/Harris-Points
  - Automatic scale selection
  - Feature Descriptors
People Detection IV

- Implicit Shape Model
  - Star-Model
  - Codebook Generation, Clustering
  - Model Training, Occurrence distributions
  - Hough Transform
  - Detection Loop
  - Figure-Ground segmentation
  - Articulations
    - 4D-ISM, silhouette verification
    - Cross-Articulation learning
  - Instance-Specific Models
Scene Context & Geometry

- The Role of Scene Context
  - Dynamics, Perspective, Ground-Plane assumption
  - Estimating horizon position
  - Surface Estimation

- Multi-view geometry
  - Disparity
  - Epiploar geometry
  - Essential and Fundamental matrix
  - Stereo calibration, rectification
  - Correspondence problem
Facial Expression Recognition

- Motivation, Problems, etc.

- Level of description
  - Facial expressions vs. Emotions
  - Six basic emotions
  - Action Units
  - Facial Action Coding System (FACS)

- Systems
  - CMU Facial Analysis System (Tian et al)
    - Model-based features, ANNs for classification
  - UCSD system for analysis of spontaneous facial behaviors
    - Appearance-based (Gabor-WL), SVM / HMM
Gesture Recognition

- Definition, applications, types of gestures
- HMMs
  - States, Observations
  - Problems: evaluation, decoding, learning
  - Forward, Viterbi algorithm
  - Baum-Welch (Forward-Backward)
- Systems
  - Sign Language Recognition (Starner et al)
  - Pointing Gesture Recognition (Nickel et al)
  - Combining Gestures and Speech
- Applications
Tracking I

- Definitions
- Features
  - Templates, color, background models

- Tracking Schemes
  - Mean Shift
  - Kalman Filter
    - Predict, Update
  - Particle Filter
  - Tracking by Detection

- Examples
  - Audio-Visual Tracking
Tracking II

- Multi-camera systems
  - Calibration, triangulation, stereo processing
- Multi-object tracking
- Tracking of heads and hands
  - Depth from stereo + color, …
- Articulated Body Tracking
  - Taxonomy
  - Models
  - Particle filter for tracking
  - Volume Carving / Voxels
- Metrics for Multi-Object Tracking
Activity Analysis

- Types of actions
- Types of activities
- Temporal Templates, Motion features
- Actions == space-time objects
  - Action features
Good Luck for Your Exam!!!