Advanced Video Analysis & Imaging (5LSH0), Module 10B

Case Study 2
Content Analysis in Surveillance

Peter H.N. de With & Jungong Han
(p.h.n.de.with@tue.nl)

Introduction

* What is video surveillance?

System Overview

* Hardware becomes commodity over time
* Focus on video content analysis (software system)

Outline

* Background introduction and motivation
* System overview
* Problem definition and available techniques
* Our proposals
* Experimental results
* Conclusions and references
Problem Statement – (1)

* Problems
  - System level
    • Find good balance between complexity and accuracy
  - Component level
    • Segmentation is not robust against noise, poor lighting condition and so on
    • Tracking is more difficult when occlusion occurs

* Existing solutions
  - GMM-based background subtraction (Gauss. Mixed Model)
  - tracking: mean-shift and particle filter

Problem Statement – (2)

* GMM-based segmentation
  - General idea: every pixel has 3 color channels (R G B).
    For each channel, we use a GMM to model it
  - This can be used to deal with small lighting changes

Problem Statement – (3)

* Mean-shift vs Particle filter

Proposed System Architecture
**Markov Random Fields – (1)**

* Neighbors and cliques
  - Define
    - $S$ – set of lattice points
    - $s$ - a lattice point
    - $X_s$ - the value of $X$ at $s$
    - $\partial s$ - the neighboring points of $s$
  - **Neighborhood** system $\partial s$ must have symmetry
    - $r \in \partial s \Rightarrow s \in \partial r$ also $s \not\in \partial s$
  - **Clique** is a set of points $c$, which are all neighbors of each other
    - $\forall s, r \in c, r \in \partial s$

**Markov Random Fields – (2)**

- The **set** of all cliques of $S$ is denoted by $\Omega$
- **Example of neighborhood**
  - $c=1$
  - $c=2$
  - **Examples of cliques**
    - One-point clique
    - Two-point cliques
    - Three-point cliques
    - Four-point clique

**Markov Random Fields – (3)**

* Random Fields
  - Consider the random vector $X = (X_s)_{s \in S}$ on $S$ is called a **Random Field** and assumed to have density $p(X)$
  - Images seen as Random Fields: If vector $X$ represents intensity values of an image, then its component $X_s$ is the intensity value at location $s=(i, j)$

**Markov Random Fields – (4)**

* **Markov Random fields**
  - **Definition**: If $p(X)$ of a random field fulfills the so-called Markov condition with respect to a neighborhood system, it is called a **Markov Random Field**
    - $P(X_s, \forall t \neq s) = P(X_s, \forall t \in \partial s)$
  - **Advantages of MRF**
    - only local dependencies
    - isotropic behavior
    - can be used to model image
  - **Problem**
    - probability (local dependencies) is still difficult to compute
Markov Random Fields – (5)

\[
\begin{align*}
X \text{ is a MRF} & \iff P(X) \text{ is a Gibbs distribution} \\
& \quad \text{and} \quad P(X) > 0
\end{align*}
\]

* Beneficial if we specify MRF in terms of Gibbs distribution (a model gives structure)
  - Local characteristic can be modelled and computed
  - Probability can be computed by energy function
  - Joint and conditional probability could be calculated

MRF and Gibbs Distribution

* Definition of Gibbs distribution

\[
P(X) = Z^{-1} \exp(-U(X)/T)
\]

- \(Z\) is a normalizing constant
- \(U\) is the energy function. Low-energy configurations have high probabilities, and vice versa
- \(T\) is the temperature. It is a scale parameter for the distribution

MRF Object Segmentation – (1)

* Object segmentation
  - Let \(q(x,y)\) be the state of the pixel at \((x,y)\), where \(q(x,y) = u\) is labelled as background pixel and \(q(x,y) = c\) is foreground pixel. \(d(x,y)\) is difference image obtained by original image and background model
  - What we want is to know which state (background or foreground) is more probable for \((x,y)\), given \(d(x,y)\)

\[
P(q=u|d) >_{\text{BG}} P(q=c|d)
\]

- Using Bayes rule

\[
P(d|q=u) \times P(q=u) >_{\text{BG}} P(d|q=c) \times P(q=c)
\]

- \(P(d|q)\): Gaussian; \(P(q)\): MRF

MRF Object Segmentation – (2)

* MRF-based segmentation algorithm
  - MRF transforms segmentation into energy-function minimization over a region adjacency graph
  - The energy function to be minimized is defined as

\[
U(f|d) = \sum_{i} V_i(f_i, d) + \sum_{(i,j) \in E_i} V_{ij}(f_i, f_j)
\]

- Implementation of a MRF algorithm requires
  - Energy function definition
  - Energy function optimization: ICM, Graph-cut, Belief Propagation algorithm
MRF Object Segmentation – (3)

* Energy function definition
  - Spatial MRF model
  - Spatial-temporal MRF model
    • only consider backward frame
    • consider both forward and backward frame

* Disadvantages
  • object boundary is less- or over-estimated
  • optimization speed is too low

Texture-MRF Object Segmentation – (1)

* Concept
  - edge information may indicate the boundary of the object
  - we can decide texture-rich areas by texture analysis
  - we should consider texture information in MRF model

* Implementation
  - use Canny operator to detect edge
  - use structure tensor to detect texture-rich area
  - remove foreground pixels if they are in the texture-rich area
  - if the pixel is labeled as the edge pixel
    • edge energy will be reduced
    • we expect the boundary pixel more relies on its node energy

Texture-MRF Object Segmentation – (2)

* Dynamic graph-cut
  - energy functions of continuous frames have no big difference
  - the optimization result of the previous frame can be used to optimize the current frame

Occlusion Handling – (1)

* Basic ideas
  - project the human silhouette into the x-axis
  - count the projecting histogram
  - we should observe a peak in the vicinity of a head
  - this is true during the complete and partial occlusion
  - it will not be influenced by the view angle
  - finding peak equals to locate person during occlusion
Occlusion Handling – (2)

* Problems
  – segmentation results are not very good
  – some body parts may lead to fake “peak” point

* Solution
  – Gaussian curve-based regression

Occlusion Handling – (3)

* Two-step algorithm
  – locate the person
    • projection histograms of human silhouette
    • generate curves using regression model
    \[
    \chi^2 = \sum_{i=1}^{m} (y_i - F(c, x_i))^2, \quad F(c, x) = c_1 \exp\left(-\frac{(x-c_2)^2}{2c_3^2}\right) + c_4 \exp\left(-\frac{(x-c_5)^2}{2c_6^2}\right).
    \]
    – find peak on a smooth curve
  – recognize the person
    • histogram-based template matching
    \[
    C_j = \arg\max_n \rho(W_j(u_i), T_n(u_i)).
    \]

HMM-Based Event Detection – (1)

1. \(N\), number of states in the model
   • states, \(S = \{s_1, s_2, \ldots, s_N\}\)
   • state at time \(t\), \(q_t \in S\)

2. \(M\), number of distinct observations
   • observation symbols, \(V = \{v_1, v_2, \ldots, v_M\}\)
   • observation at time \(t\), \(O_t \in V\)

3. State transition probability distribution, \(A = \{a_{ij}\}\),
   \[a_{ij} = P(q_{t+1} = s_j | q_t = s_i), \quad 1 \leq i, j \leq N\]

4. Observation symbol probability distribution in state \(j\)
   \[b_j(k) = P(v_k | q_t = s_j), 1 \leq j \leq N, 1 \leq k \leq M\]

5. Initial state distribution, \(\pi = (\pi_1)\)
   \[\pi_1 = P(q_1 = s_1), 1 \leq j \leq N\]

HMM-Based Event Detection – (2)

* In our behavior evaluation project
  – we want to detect two pre-defined events

  – observation vector: \([x_1 y_1 x_2 y_2]\)
  – state definition can use the location of the important objects in the scene, like bed in this case
Experimental Results and Demos – (1)

* T-MRF Segmentation

Experimental Results and Demos – (2)

* Texture-MRF segmentation demos

Experimental Results and Demos – (3)

* Occlusion handling

Experimental Results and Demos – (4)

* Occlusion handling demo
Experimental Results and Demos – (5)

- Apply to the behavior evaluation project

Conclusions Case 2 on Surveillance

- A texture-involved MRF model for segmentation
- A novel occlusion handling algorithm
- A real-time system with good balance between accuracy and complexity
- References
  - J. Han, M. Feng, P.H.N. de With. “A real-time video surveillance system with human occlusion handling using nonlinear regression” IEEE ICME, 2008.