1

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)

2

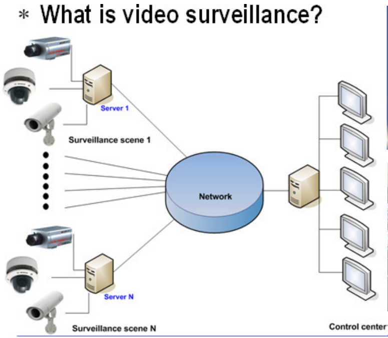
Outline

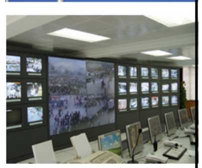
* Background introduction and motivation
* System overview
* Problem definition and available techniques
* Our proposals
* Experimental results
* Conclusions and references

3

Introduction

* What is video surveillance?

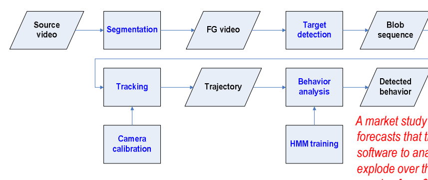




4

System Overview

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



A market study from IMS research forecasts that the world market for software to analyze video content will explode over the next five years, growing from $67.7 million in 2004 to $ 839.2 million in 2009, at a CAGR of 65.5%

5

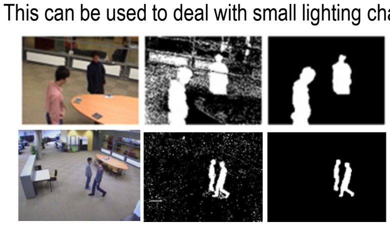
Problem Statement – (1)

* **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
  - tracking: mean-shift and particle filter

6

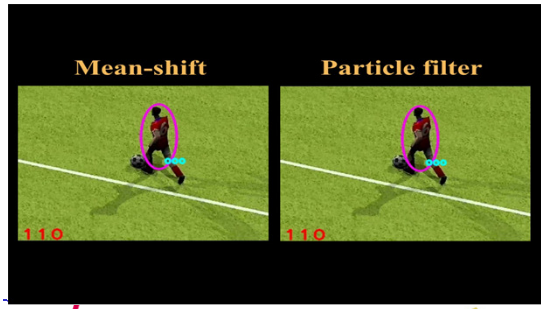
Problem Statement – (2)

* **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

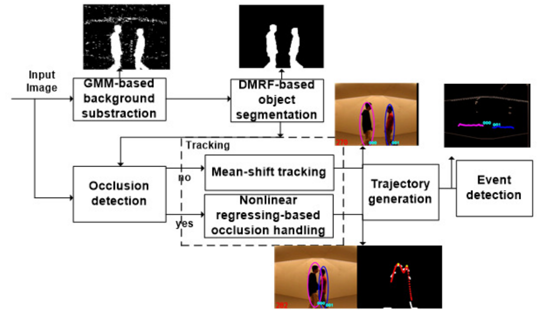


Problem Statement – (3)

* 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
    - ∂s - the neighboring points of s
  - Neighborhood system ∂s must be symmetry

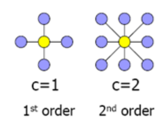
$$r \in \partial s \Rightarrow s \in \partial r \quad also \quad s \notin \partial s$$

  - Clique is a set of points, c, which are all neighbors of each other

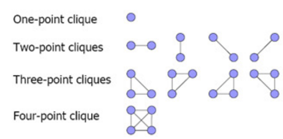
$$\forall s, r \in c, r \in \partial s$$

Markov Random Fields – (2)

- the set of all cliques of S is denoted by Ω
- example of neighborhood

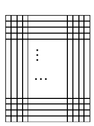


- example of cliques



Markov Random Fields – (3)

* Random fields
  - the random vector $X = (X_s)_{s \in S}$ on S is called a random field and assumed to have density $p(X)$
  - images 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)

s = 

X= 

640x480

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

$$P(X_s \mid X_t, \forall t \neq s) = P(X_s \mid X_t, \forall t \in \partial s)$$

  - 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)

$$\begin{pmatrix} X \quad is \quad a \quad MRF \\ \& \\ P(X) > 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} P(X) \quad is \quad a \\ Gibbs \quad distribution \end{pmatrix}$$

- **Benefits if we specify MRF in terms of Gibbs distribution**
  - 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 verse
  - T is the temperature. It is a scale parameter for the distribution

TU/e PdW-SZ-EB / 2016 Fac. EE SPS-VCA Adv. Video Analysis & Imaging. / 5LSH0 / Mod 10B Case 2: Surveillance VCA

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 $d(x,y)$

$$P(q=u\,|\,d) \underset{<\; FG}{\overset{>\; BG}{}} P(q=c\,|\,d)$$

  - using Bayes rule

$$P(d\,|\,q=u)\times P(q=u) \underset{<FG}{\overset{>BG}{}} P(d\,|\,q=c)\times P(q=c)$$

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

TU/e PdW-SZ-EB / 2016 Fac. EE SPS-VCA Adv. Video Analysis & Imaging. / 5LSH0 / Mod 10B Case 2: Surveillance VCA

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

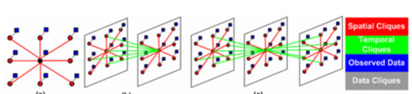
$$U(f|d) = \underbrace{\sum_i V_i(f,d)}_{\text{unaffected by label of j}} + \underbrace{\sum_{(i,j)\in C_i} V_{i,j}(f,d)}_{\text{depending on label of j}} .$$

  - implement a MRF algorithm
    - energy function definition
    - energy function optimization: ICM, Graph-cut, belief propagation

TU/e PdW-SZ-EB / 2016 Fac. EE SPS-VCA Adv. Video Analysis & Imaging. / 5LSH0 / Mod 10B Case 2: Surveillance VCA

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

TU/e PdW-SZ-EB / 2016 Fac. EE SPS-VCA Adv. Video Analysis & Imaging. / 5LSH0 / Mod 10B Case 2: Surveillance VCA

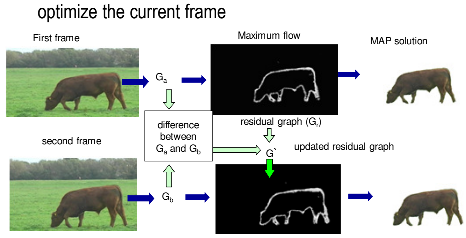
T-MRF Object Segmentation – (1)

- **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

TU/e PdW-SZ-EB / 2016 Fac. EE SPS-VCA Adv. Video Analysis & Imaging. / 5LSH0 / Mod 10B Case 2: Surveillance VCA

T-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

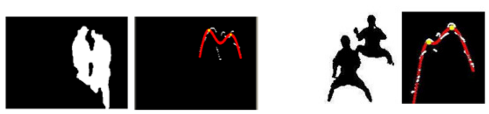


Occlusion Handling – (1)

* 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)

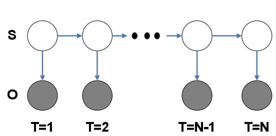
* Two-step algorithm
  - locate the person
    - projection histograms of human silhouette
    - generate curves using regression model

$$\chi^2 = \sum_{i=1}^{m} (y_i - F(\mathbf{c}, x_i))^2. \qquad F(\mathbf{c}, x) = c_1 exp(-\frac{(x-c_2)^2}{2c_3^2}) + c_4 exp(-\frac{(x-c_5)^2}{2c_6^2}).$$

    - find peak on a smooth curve
  - recognize the person
    - histogram-based template matching

$$C_j = \arg\max_n \rho(\hat{W}_j(u_i), \hat{T}_n(u_i)).$$

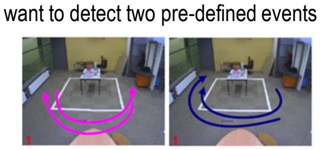
HMM-Based Event Detection – (1)



1. N, number of states in the model
   - states, $S = \{s_1, s_2, \dots s_N\}$
   - state at time t, $q_t \in S$
2. M, number of distinct observations
   - observation symbols, $V = \{v_1, v_2, \dots v_N\}$
   - observation at time t, $O_t \in V$
3. State transition probability distribution, $A = \{a_{i,j}\}$, $a_{i,j} = P(q_{t+1} = s_j | q_t = s_i), \quad 1 \le i, j \le N$
4. Observation symbol probability distribution in state j $B = \{b_j(k)\}$ $b_j(k) = P\left[v_k | q_t = s_j\right], 1 \le j \le N, 1 \le k \le M$
5. Initial state distribution, $\prod = \{\pi_i\}$ $\pi_i = P\left[q_1 = s_j\right], 1 \le j \le N$

HMM-Based Event Detection – (2)

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



  - observation vector: $[x_1\ y_1\ v_1\ x_2\ y_2\ v_2]$.
  - 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

VIDEO SEQUENCE CHARACTERISTIC DESCRIPTION

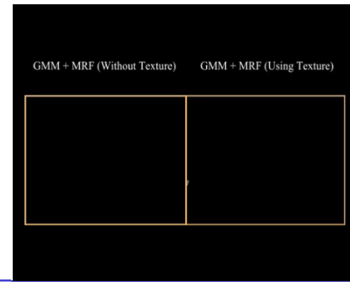
	Total frames	Environment	Difficulty	Characteristics
Video1	400	outdoor	*high*$^+$	large area with waving trees
Video2	150	outdoor	*high*$^-$	small area with waving trees
Video3	490	indoor	*middle*$^-$	sudden lighting condition changes
Video4	500	indoor	*low*	\
Video5	450	indoor	*middle*$^+$	object boundary's color ≈ background color

COMPARISON RESULTS OF CORRECT AND FALSE FRAMES FOR THE SEGMENTATION ALGORITHMS

Video	GMM model		SMRF model		STMRF model		STTMRF model	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
1	11.96%	85.74%	37.97%	84.43%	42.45%	85.78%	60.78%	89.14%
2	39.05%	90.88%	81.53%	87.88%	81.25%	89.61%	79.51%	94.19%
3	91.90%	94.98%	94.11%	96.94%	94.35%	96.98%	94.83%	97.42%
4	93.65%	95.30%	95.59%	97.49%	95.56%	97.51%	95.63%	99.57%
5	87.09%	88.21%	89.25%	92.78%	89.77%	93.61%	92.53%	95.09%

Experimental Results and Demos – (2)

* T-MRF segmentation demos

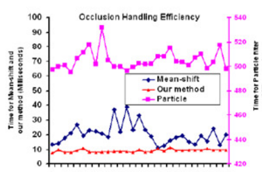


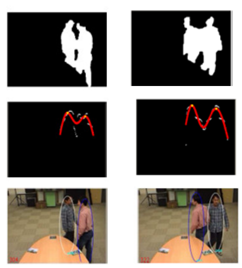
Experimental Results and Demos – (3)

* Occlusion handling

COMPARISON OF TRACKING ALGORITHMS.

	Type	MS	PF	Our method
Clip1	two people	80.6%	99.5%	99%
Clip2	two people	54.8%	99.2%	92.1%
Clip3	three people	90.1%	81.1%	90%
Clip4	three people	70%	85%	93%



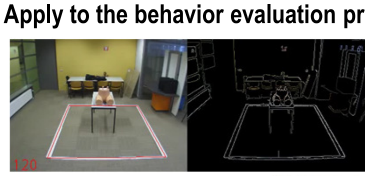


Experimental Results and Demos – (4)

* 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.
– T. Aach, A. Kaup, "Bayesian algorithms for adaptive change detection in image sequences using Markov random fields," Signal processing: image communication, 1995