Integrating a discrete motion model into GMM based background subtraction

ICPR 2010

Christian Wolf and Jean-Michel Jolion

Université de Lyon, CNRS
INSA-Lyon, LIRIS, UMR5205, F-69622, France
http://liris.cnrs.fr/christian.wolf

August 23rd, 2010

Plan

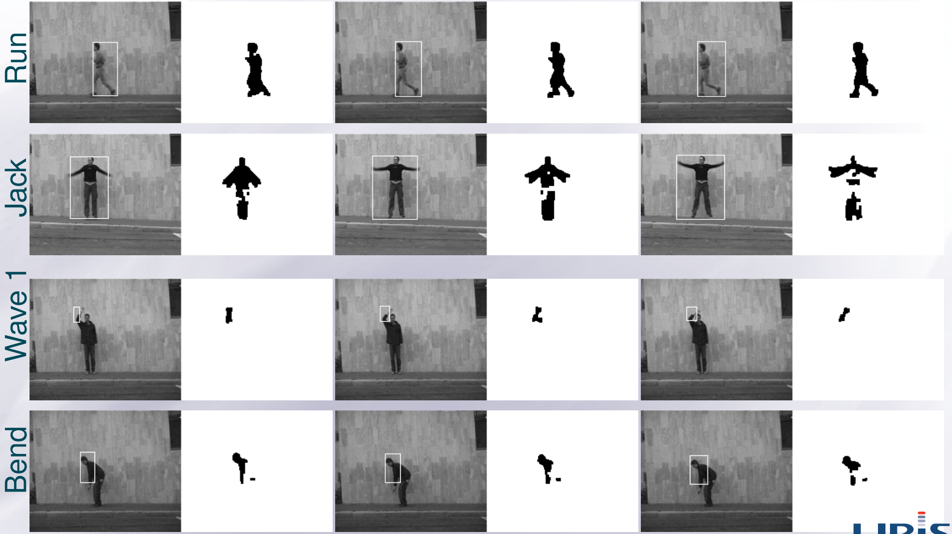
Introduction

Objectives

- Binary segmentation into Background/moving objects
- Objects which stop and temporarily do not move need to stay 'object'
- An explicite background model is created, mostly pixelwise

Background subtraction

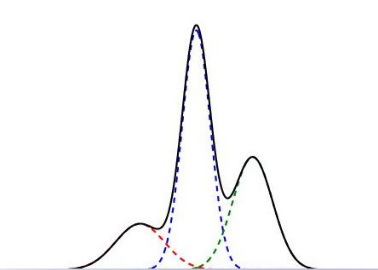
Problems of the frame differencing algorithms



GMM based background subtraction

State of the art [Stauffer and Grimson, 2000]

- K Gaussians are held and updated for each pixel
- Grayvalues y_i of pixels i are matched agains Gaussians
- A subset of the Gaussians is identified as BG (low variance, high matching frequency)



Component k of pixel i :

- Mean μ_i^k
- Variance σ_i^k
- weight w_i^k

$\Longrightarrow$ Several FG distributions ! (waving trees, flags, etc.)

Is this pixel foreground or background ?

- Identify the best matching BG distribution
- Is this distribution FG or BG ?
  - Order all distributions according to (low) std dev σ and (high) weight w :

$$\frac{w^k}{\sigma^k}$$

  - The first distributions occupying $T = 70\%$ weight (w) are considered as BG.

- the decision is crisp and may fail (often) !
- pixels are treated independently

Our contribution : global decision

Take the decision globally together for all pixels of a short spatio-temporal block :

$\Longrightarrow$

- Spatial regularization (neighbors tend to have similar labels)
- Temporal regularization (labels tend to remain for some time)

Markov random field (MRF)

Global minimization of a global energy function

$$E(x, y) = (...)$$

- $x = \{x_1, x_2, \ldots, x_N\}$ labels of a short space-time block, sequentially indexed ; BG=0, FG=1.
- $y = \{y_1, y_2, \ldots, y_N\}$ the grayvalues/colors.

$$\hat{x} = \arg\min_x E(x, y)$$

Trivial energy function : no regularization

- Identify the best matching background (! !) distribution for pixel i
- Measure the normalized deviation from the mean μ_i

$$\Delta_i = \frac{|y_i - \mu_i|}{\sigma_i}$$

Minimize over x :

$$E(x,y) = \sum_i E_d(\Delta_i, x_i)$$

$$E_d(\Delta_i, x_i) = \begin{cases} \Delta_i & \text{if } x_i = 0 \text{ (BG)} \\ 2D - \Delta_i & \text{if } x_i = 1 \text{ (FG)} \end{cases}$$

$\Longrightarrow$ equal to pixelwise thresholding with threshold D.

Spatio-temporal regularizer

Spatial AND temporal regularization :

$$\begin{aligned} E(x,y) &= \alpha_d \sum_i E_d(\Delta_i, x_i) \\ &+ \alpha_s \sum_{i \sim j} \delta(x_i, x_j) \\ &+ \alpha_t \sum_{i \wr j} \delta(x_i, x_j) \end{aligned}$$

Favors similar labels over spatial and temporal neighbors

$i \sim j$	i and j are spatial neighbors
$i \wr j$	i and j are temporal neighbors
$\delta(a,b)$	0 if $a{=}b$, 1 else (Kronecker delta)

Include a motion model

Problem : moving objects disturb the temporal regularization.

$\Longrightarrow$ include a dense motion vector field $\boldsymbol{u}_i$ with horizontal and vertical components $(u_{i,x}, u_{i,y})$:

$$\begin{aligned}
E(x,y,\boldsymbol{u}) &= \alpha_d \sum_i E_d(\Delta_i, x_i) \\
&+ \alpha_s \sum_{i \sim j} \delta(x_i, x_j) \\
&+ \alpha_t \sum_i \delta(x_i, x_{i \rightarrow \boldsymbol{u}_i}) \\
&+ \alpha_m \sum_{i \sim j} E_m(\boldsymbol{u}_i, \boldsymbol{u}_j) \\
&+ \alpha_m \sum_{i \wr j} E_m(\boldsymbol{u}_i, \boldsymbol{u}_j)
\end{aligned}$$

Problem : exact global optimization is difficult

Include a motion model

Problem : exact global optimization is difficult

Solution : separate optimization

- Step 1 : optimize over the FG/BG labels with fixed motion vectors
- Step 2 : estimate motion vectors with fixed labels
  $\Longrightarrow$ Optical flow problem

... and iterate until convergence/tired of it.

Step 1 : estimate labels

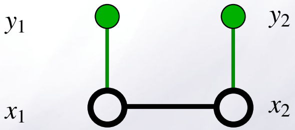
Optimize over the FG/BG labels x_i with fixed motion vectors $\boldsymbol{u}_i$:

$$\begin{aligned}\hat{x} = \arg\min_x \quad & \alpha_d \sum_i E_d(\Delta_i, x_i) \\ + \quad & \alpha_s \sum_{i \sim j} \delta(x_i, x_j) \\ + \quad & \alpha_t \sum_i \delta(x_i, x_{i \rightarrow \boldsymbol{u}_i})\end{aligned}$$

Discrete optimization problem

- All terms are binary labeled and submodular
- $\Longrightarrow$ exact minimization using min st-cut and Kolmogorv et. al's graph construction.

Graph cuts min. of binary submodular functions



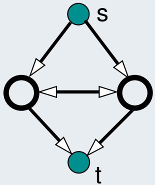
$$E(x,y) = E_1(x_1;y_1) + E_1(x_2;y_2) + E_2(x_1,x_2)$$

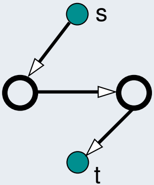
Submodular terms

$$E_2(0,0) + E_2(1,1) \le E_2(0,1) + E_2(1,0)$$

Graph construction methods

Boykev et al. (PAMI 2001)





Kolmogorov et al. (PAMI 2004)

Step 2 : approximate optical flow

Motion vectors $\boldsymbol{u}_i$ are calculated using approximate optical flow.

Trivial version : simple block matching with discretized vectors $\boldsymbol{u}$

$$\hat{\boldsymbol{u}} = \arg\min_{\boldsymbol{u}} \quad \alpha_t \sum_i \min_{\boldsymbol{u} \in [1,T]^2} |y_i - y_{i \to \boldsymbol{u}}|$$

$\Longrightarrow$ pixelwise optimization

Next step : additional performance gain through regularization of the motion vectors $\boldsymbol{u}$

Step 2 : approximate optical flow

Block matching with regularization of motion vectors $\boldsymbol{u}$:

$$\begin{aligned}\hat{\boldsymbol{u}} = \arg\min_{\boldsymbol{u}} \quad & \alpha_t \sum_i \min_{a\in[1,T]} \left|y_i - y_{i\to\left[\begin{smallmatrix}u_{i,x}\\ a\end{smallmatrix}\right]}\right| \\ + \quad & \alpha_t \sum_i \min_{a\in[1,T]} \left|y_i - y_{i\to\left[\begin{smallmatrix}a\\ u_{i,y}\end{smallmatrix}\right]}\right| \\ + \quad & \alpha_0 \sum_i |u_{i,x}| + \alpha_0 \sum_i |u_{i,y}| \\ + \quad & \alpha_m \sum_{i\sim j} |u_{i,x}-u_{j,x}| + |u_{i,y}-u_{j,y}|\end{aligned}$$

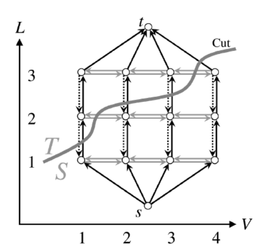
Discrete optimization problem

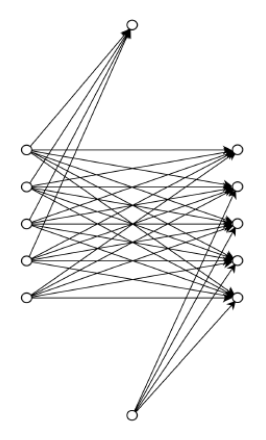
- Multiple Labels with linear order
- Terms are convex in label differences
- $\Longrightarrow$ exact minimization using min st-cut and Ishikawa's graph construction.

Graph cuts minimization of convex functions

Convexity in label differences

$$E_2(x_i, x_j) = g(x_i - x_j), \qquad g(\cdot) \text{ convex}$$



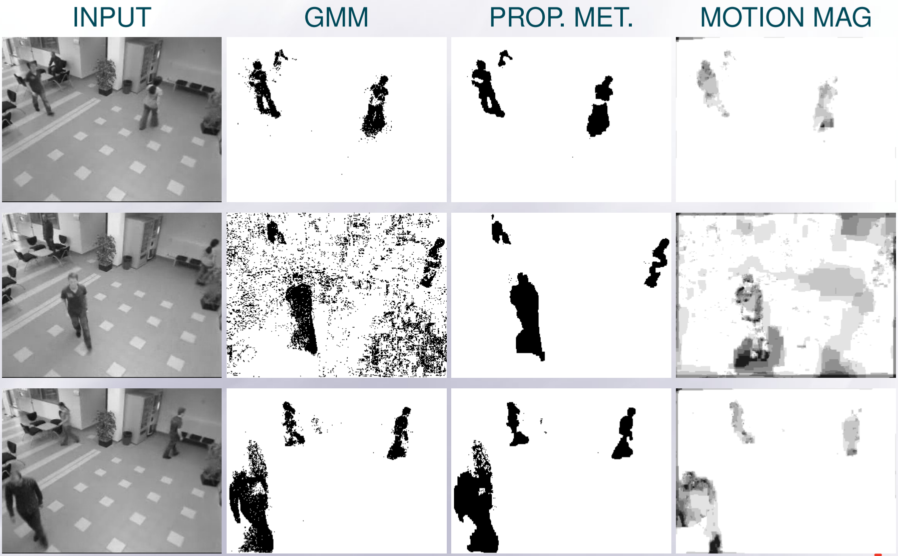


Ishikawa (PAMI, 2003)

Approximate optical flow : results

First frame Second frame Motion magnitude

BG subtraction : results



Conclusion

- New background subtraction algorithm which improves upon Grimson & Stauffer
- Spatio-temporal regularization across a whole section of the ST-cube
- A motion model compensates object movement
- Particularly powerful in case of global changes in the scene

References I

Stauffer, C. and Grimson, W. (2000).
Learning patterns of activity using real-time tracking.
IEEE Tr. on PAMI, 22(8) :747–757.