

Performance Evaluation of Frequent Events Detection Systems

X. Desurmont^a, R. Sebbe^b, F. Martin^a, C. Machy^a and J-F. Delaigle^a

^a Multitel A.S.B.L., Mons, Belgium

^b Faculté Polytechniques de Mons, Mons, Belgium.

PETS 2006

New York, USA - 18 June 2006

In Conjunction with IEEE CVPR 2006

- ✓ High demand on video analysis applications
 - Visio surveillance, traffic monitoring, marketing

- ✓ Need of evaluation
 - To prove the objective quality of a system for industrials.
 - To highlight improvement during research and thus to understand better how the system works to improve it adequately.

- ✓ Aim of the presentation
 - To describe an evaluation for detection systems.
 - To introduce an automatic re-alignment between results and ground truth.
 - To show an example with “ counting of people” system.

- ✓ Related Work
- ✓ Dynamic re-alignment
- ✓ Counting method overview
- ✓ Results and interpretation

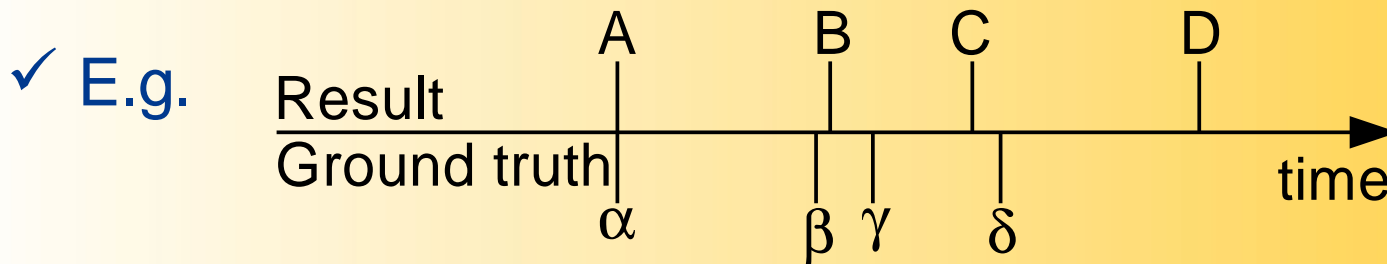
- ✓ Definition: A video event detection system is a method that triggers events when something specific happens in the scene

- ✓ The minimum metadata related to the event is the timestamp. Other information could be relevant: location, recognition of objects, behaviors, ...

- ✓ Examples
 - A leave of a luggage
 - A fight between two people
 - A car that goes too fast
 - Someone that crosses a virtual line
 - When the density of a crowd is too much

- ✓ Matching methods between ground truth (GT) and result (RS)
 - Previous approach of Desurmont and al [7] was simply to count the number of events in the video file and compute the detection rate (DR) and false alarm rate (FAR).
 - Bruneaut et al [8] propose a metric for challenge (CREDS) in 2005.
 - It matches events of GT and RS with handling temporal shift.
 - Drawback when the events are so frequent that the possibility of anticipated and delayed events entails the overlapping of several events.
 - Lopresti [9] introduces the use of dynamic re-alignment for performance evaluation of text processing.
 - In object tracking, Brown et al [10] propose a tracking evaluation with the determination of a match between RS tracks and GT tracks. However, this match is not necessarily the best one.
 - Needham et al [11] propose, for the matching, many methods to compare trajectories that can differ by spatial or temporal bias.

- ✓ Hypothesis of deviations of the event detection system: false positives, false negatives, delays and anticipations



➤ Two possible interpretations

- Assuming no big delay: A- α , B- β , C- δ , a false alarm (D) and a misdetection (γ).
- Assuming higher delay: A- α , B- β , C- γ , D- δ



Need to find the best interpretation

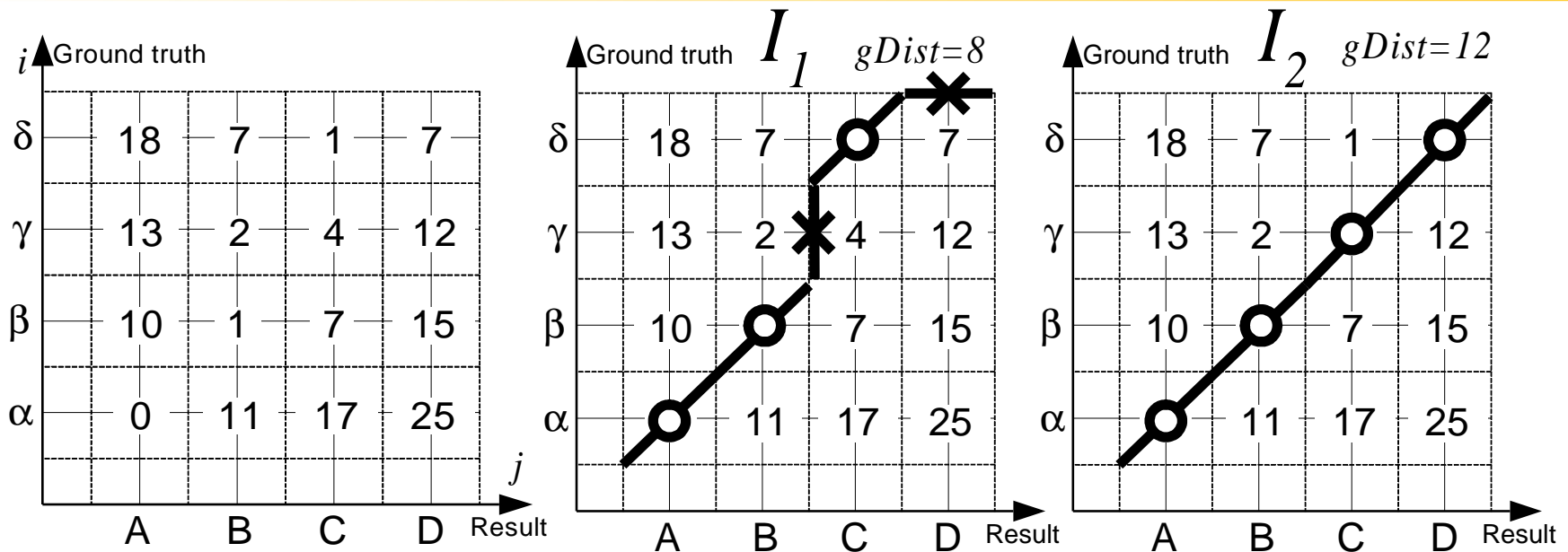
- ✓ Process the re-alignment in an objective and automatic way minimizing a global cost of matching events.
- ✓ We define a distance cost $dist_{i,j}$ between two events i,j
 $dist_{i,j}=|t_i-t_j|$; t_i and t_j are the timestamps in seconds.
- ✓ We also set a cost to false positive ($FPDist$) and false negative ($FNDist$) to decide the maximum delay and anticipation. E.g. We setup $FPDist= FNDist=3s$.
- ✓ The global cost is the sum of all costs.

Example

- ✓ Time of occurrence of events:

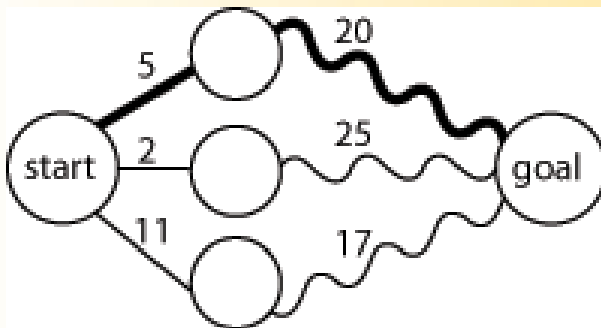
Event	α	A	β	B	γ	C	δ	D
Time	0	0	10	11	13	17	18	25

- ✓ Matrix with the distances between ground truth (GT) and system result events (RS). And then two interpretations.



Dynamic programming

- ✓ The number of interpretations is huge, we want to find the cheapest one in a reasonable time !
 - Use dynamic programming to solve this issue.
- ✓ Dynamic programming is a method for reducing the runtime of algorithms exhibiting the properties of optimal substructure.



- For example, the shortest path to a goal from a vertex in a graph can be found by first computing the shortest path to the goal from all adjacent vertices, and then using this to pick the best overall path, as shown in Figure .

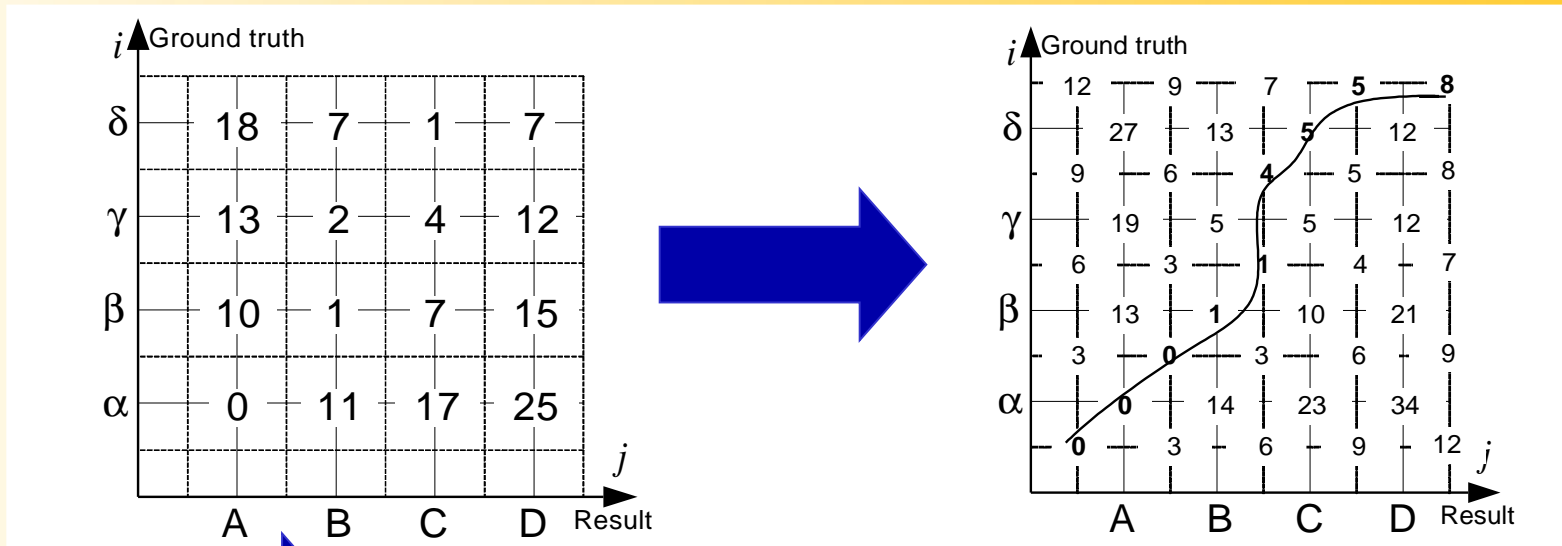
- ✓ The initial condition and the recurrence of the DP:

$$gDist_{0,0} = 0 \quad (1)$$

$$gDist_{i,j} = \min \begin{cases} gDist_{i-1,j} + FNDist & \text{if } i > 0 \\ gDist_{i,j-1} + FPDist & \text{if } j > 0 \\ gDistM_{i-1,j-1} & \text{if } (i > 0 \ \& \ j > 0) \end{cases} \quad (2)$$

$$gDistM_{i,j} = gDist_{i,j} + dist_{i,j} \quad (3)$$

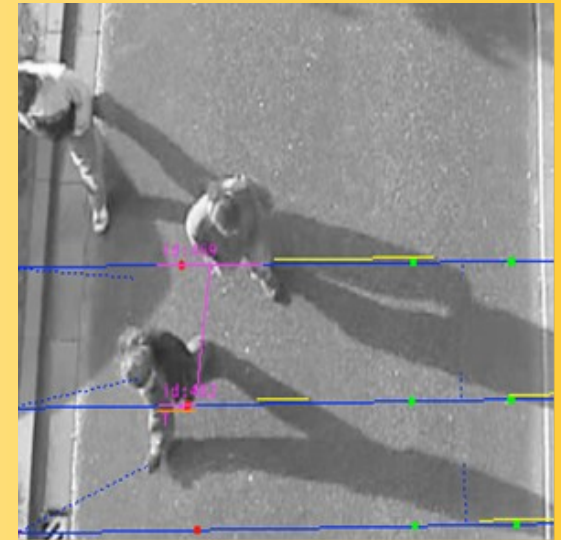
E.G



A- α , B- β , C- δ , a false alarm (D) and a misdetection (γ).

Counting method overview

- ✓ The goal of the counting method is to count people crossing a region, maintaining 2 counters, one for each direction.



- ✓ Processing steps

- 1) compute background estimation
- 2) compute foreground extraction
- 3) compute automatic thresholding + mathematical morphology operators
- 4) perform shadows removal
- 5) split/merge blobs according to size parameters to obtain candidates
- 6) evaluate the speed of candidates
- 7) perform tracking and collision detection from previous frames

- ✓ test it on 3 real sequences taken from a shopping center
 - Cumulated duration of 1 hour of video at different times in the day.
 - MPEG4 grey levels 25 fps from fixed PAL (768 x 576 pixels) cameras.



- ✓ Use the Boolean contingency table

System Observation	Ground truth	
	Positive	Negative
Positive	N_{tp} (true positives)	N_{fp} (false positives)
Negative	N_{fn} (false negatives)	N/A

Sequence 1: starting at 12:00:00, duration 15 min							
N1	GT	RS	Error	N2	GT	RS	Error
	71	74	+04%		55	66	+20%
N _{tp}		64		N _{tp}		53	
N _{fn}		7		N _{fn}		3	
N _{fp}		10		N _{fp}		12	
DR		90%		DR		96%	

Sequence 2: starting at 16:00:00, duration 15 min							
N1	GT	RS	Error	N2	GT	RS	Error
	72	76	+05%		81	99	+22%
N _{tp}		64		N _{tp}		75	
N _{fn}		8		N _{fn}		7	
N _{fp}		12		N _{fp}		23	
DR		89%		DR		93%	

Sequence 3: starting at 18:00:00, duration 15 min							
N1	GT	RS	Error	N2	GT	RS	Error
	216	225	+04%		204	257	+26%
N _{tp}		192		N _{tp}		196	
N _{fn}		24		N _{fn}		8	
N _{fp}		33		N _{fp}		61	
DR		88%		DR		96%	

✓ Interpretations:

- False alarms and false detections cancel each other when we just do counting
- One direction (N2) triggers more False Alarm => this is due to people behavior that stop to the flower shop when existing the shopping center

Display of the events editor



✓ Summary

- New robust metric for global performance evaluation of events detection.
- It is interesting to find out precisely which events are “ false” .
- Description of a counting method and evaluated it according to the new framework.

✓ Future work

- Extend the metric to the handling of substitution of events or aggregation of events (e.g. sometimes there are two detections for a single event).
- Take into account the implementation practical standardization of the metric (the definition of the XML/CEP format, etc.).

THANK YOU!

desurmout@multitel.be