

IRIM @TRECVID

**Collaboration between  
French labs for a joint  
participation to TRECVID  
video semantic indexing task**

DGA Workshop - July 6th, 2010

[www.IRIT.fr/~Herve.Bredin](http://www.IRIT.fr/~Herve.Bredin)

# Outline

- TREC Vid
- IRIM
- IRIM @ TREC Vid 2009

# TREC Vid

- TREC Video Retrieval Evaluation
  - Every year, since 2001
  - Organized by the **N**ational **I**nstitute for **S**tandards and **T**echnology

*“Encourage research in information retrieval by providing a large test collection, uniform scoring procedures, and a forum for organizations interested in comparing their results”*

<http://trecvid.nist.gov>

# TRECVID 2010

- **Semantic indexing**
- Known-item search
- Content-based copy detection
- Surveillance event detection
- Instance search (pilot)
- Event detection in Internet multimedia (pilot)

# Semantic Indexing @TRECVID

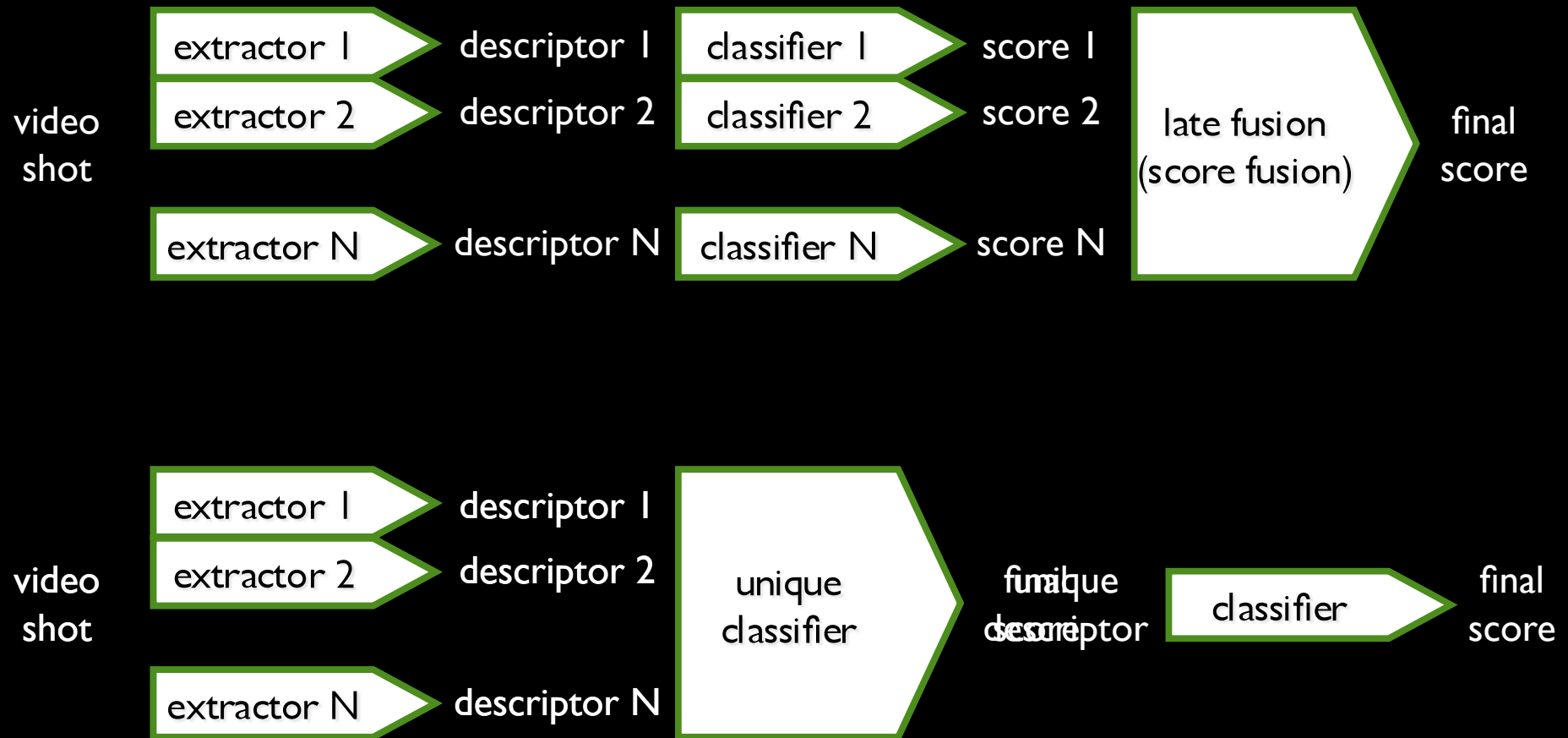
motorcycle natural disaster kitchen  
stadium dog prisoner chair crowd  
bridges tent vegetation  
ex laboratory  
soccer  
musical instrument US flag demonstration  
roadway junction tennis scientist  
singing meeting dancing waterscape

**Encourage design  
of generic approaches**

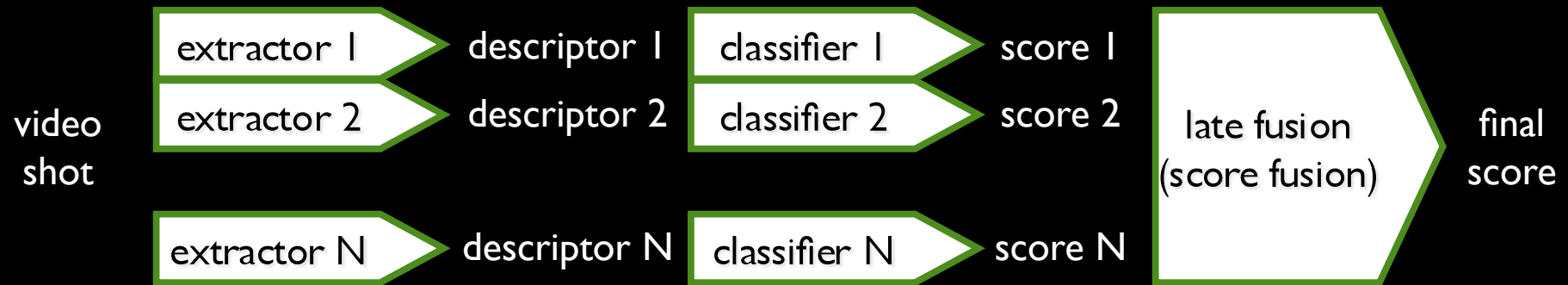
# Semantic Indexing @TRECVID

- **Input (provided by the organizers)**
  - Training/Development set
    - 200 hours video with shot boundaries
    - Annotations for 130 “*semantic concepts*” at shot level
  - Test set
    - 200 hours video with shot boundaries
- **Output (submitted by participants)**
  - 130 sorted lists of 2000 shots most likely to contain each concept

# Various Approaches to Semantic Indexing



# Various Approaches to Semantic Indexing



## Expertise required in multiple fields

### **video processing**

image

audio

text

### **machine learning**


noisy data

few positive samples

### **multimodal fusion**



# IRIM

- In the framework of GdR 
- Stands for  
“***I**ndexation & **R**echerche d’**I**nformation **M**ultimédia”*”
- Coordinated by J. Benois-Pineau (LABRI),  
G. Quénot (LIF), M. Cord (LIP6) & H. Bredin (IRIT)
- Help french labs participate to TRECVID
  - Collaborate and share expertise
  - Avoid starting from scratch for newcomers

# IRIM (2)



14 labs in 2009

2 workshops a year  
(pre/post TRECVID)

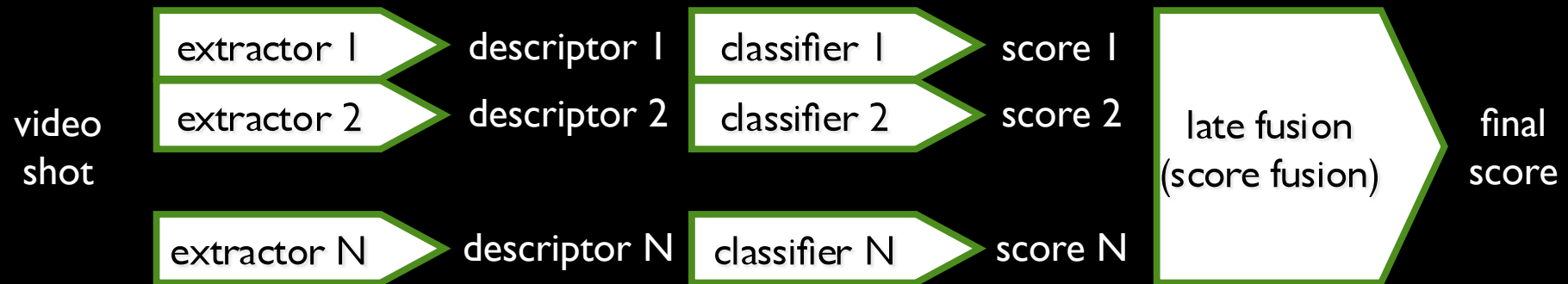


- sponsored  
thematic workshop

*"Scalable and cross-media indexing"*

[mrim.imag.fr/irim/](http://mrim.imag.fr/irim/)

# IRIM @TRECVID 2009



- **100+ descriptors**
- **Video**
  - Color: global and regional histograms
  - Texture: Gabor filters, wavelets, ...
  - Shape: edge orientations, local descriptors
  - Motion: optical flow
  - Face detection
- **Audio**
  - Energy
  - Spectral centroid
  - Mel-Frequency Cepstrum Coefficients
- **Multiple classifiers**
  - k-Nearest Neighbors
  - Support Vector Machine
- **Multiple fusion algorithms**
  - Genetic fusion
  - Weighted combination

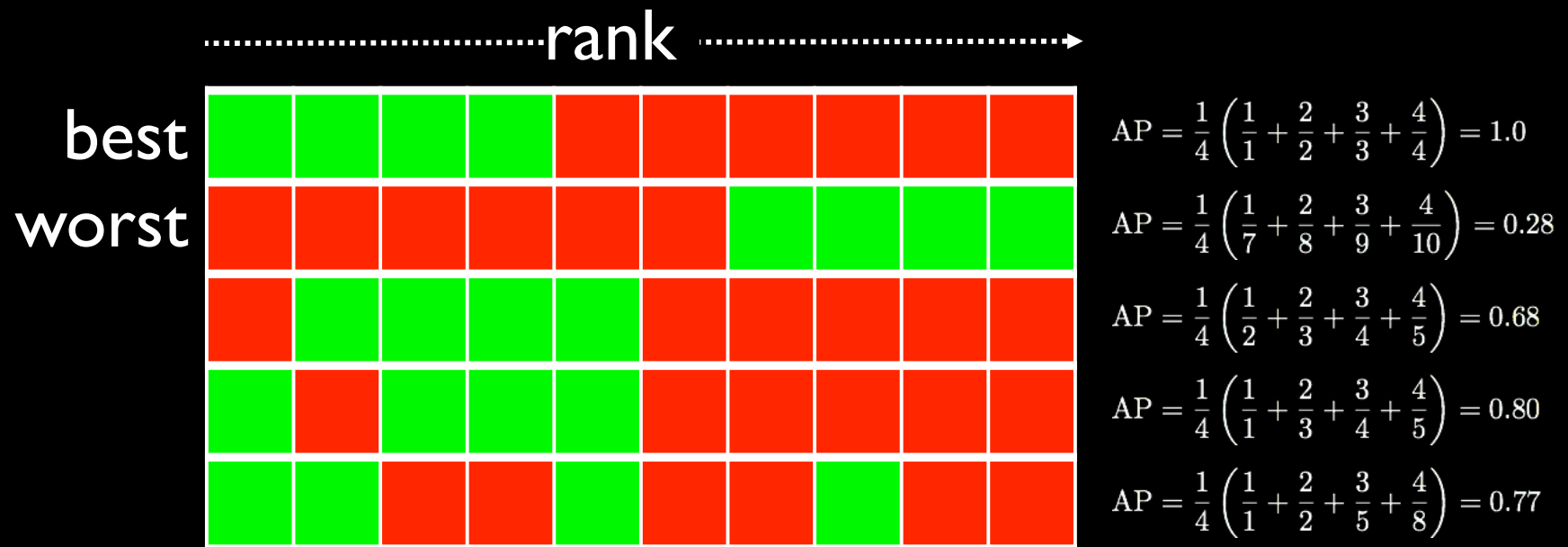
# Average Precision

$$AP = \frac{\sum_{r=1}^N \delta_r \cdot \text{precision@r}}{\sum_{r=1}^N \delta_r}$$

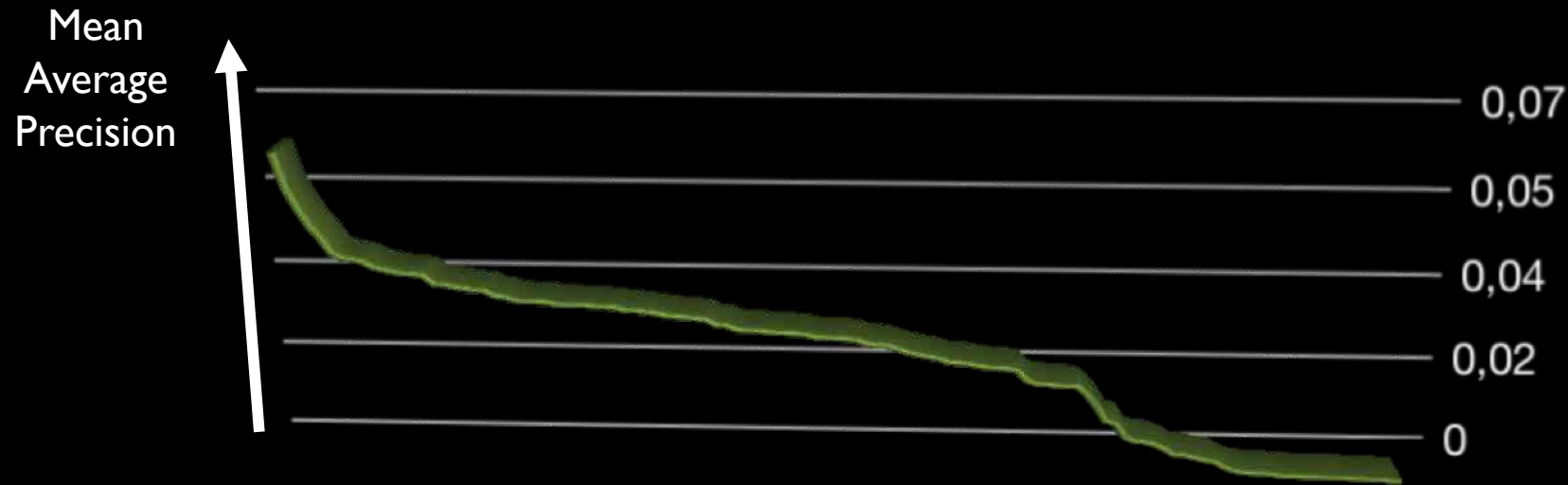
$N$  number of shots

$\delta_r$  1 if shot  $r$  is relevant, 0 otherwise

precision@ $r$  ratio of relevant samples in top  $r$  list



# Performance overview



- **Best descriptors are visual**  
Variant of Scale-Invariant Feature Transform (MAP = 0.0644)
- **Best audio descriptor**  
Mel-Frequency Cepstrum Coefficient (MAP = 0.019)
- **Late fusion**  
Depends on the combination of selected descriptors & classifiers  
**0.092 < MAP < 0.132**

# TRECVID 2009

## official results

- 222 runs
- MediaMill - U of Amsterdam  
**MAP = 0.228**
- IRIM submitted 18 runs  
**0.092 < MAP < 0.132**
- **#10** among 40 participants
- Average = 0.071 / Median = 0.051

# Lessons learned

- From our runs
  - **A+V** is better than **V-only**.  
*though, no **major** improvement (yet?)*
  - Even weak descriptors can help.
- From others
  - Use multiple frames per shot.
  - Re-rank.

# Agenda 2010

- 2010-08-09: **submission deadline**
- 2010-09-17: **results available**
- 2010-11-15: **workshop @ NIST**



# IRIM @ TREC Vid

[mrim.imag.fr/irim](http://mrim.imag.fr/irim)

[trecvid.nist.gov](http://trecvid.nist.gov)

DGA Workshop - July 6th, 2010

[www.IRIT.fr/~Herve.Bredin](http://www.IRIT.fr/~Herve.Bredin)