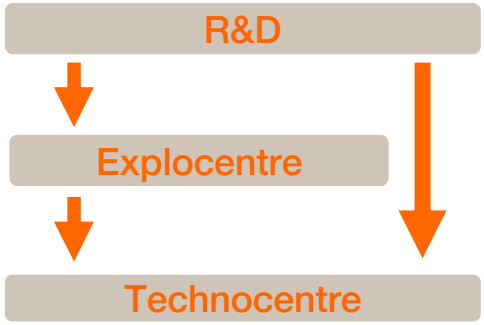
The Orange Labs VCD System
TRECVid 2008 Results

Nicolas Gengembre, Sid-Ahmed Berrani
Orange Labs – R&D/Technologies

November 17th, 2008.

Company – Lab Presentation

- **Orange Labs** is the worldwide innovation network of France Telecom.



- 5000 people
- 17 sites
- 8 countries
- international and multicultural teams

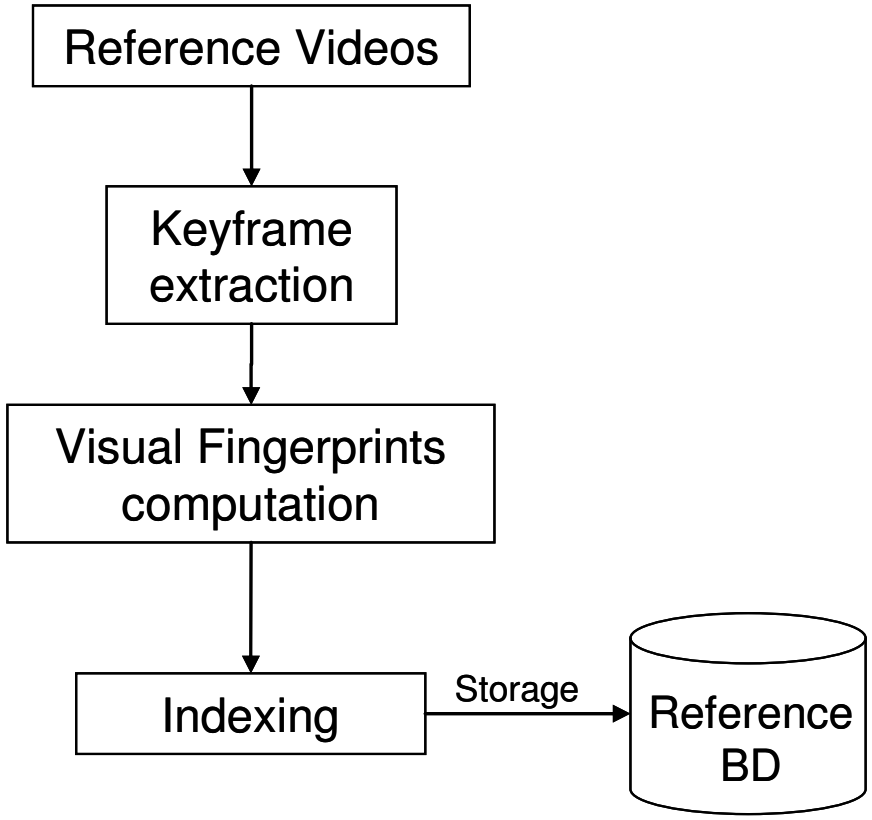
- International R&D
  - 15 locations
  - 3,800 researchers, engineers and scientists
  - More than 8,400 patents

Video Copy Detection: An Overview

- Objective:
  - Check if a video is a copy of another one even if the copy has been modified
- Constraints
  - Effectiveness (robustness): ability to be invariant to transformations the copy may undergo
  - Efficiency (rapidity): ability to handle large reference video DB and a huge number of queries
- Industrial applications:
  - Copyright protection
    - A better and secure use of videos $\Rightarrow$ Improvement of the audience
  - Duplicate detection
    - Video databases $\Rightarrow$ different copies of the same video are put together
      - A better organization
      - A more efficient database browsing
    - Video search engine $\Rightarrow$ video query answers are clustered
      - A better presentation of query results.

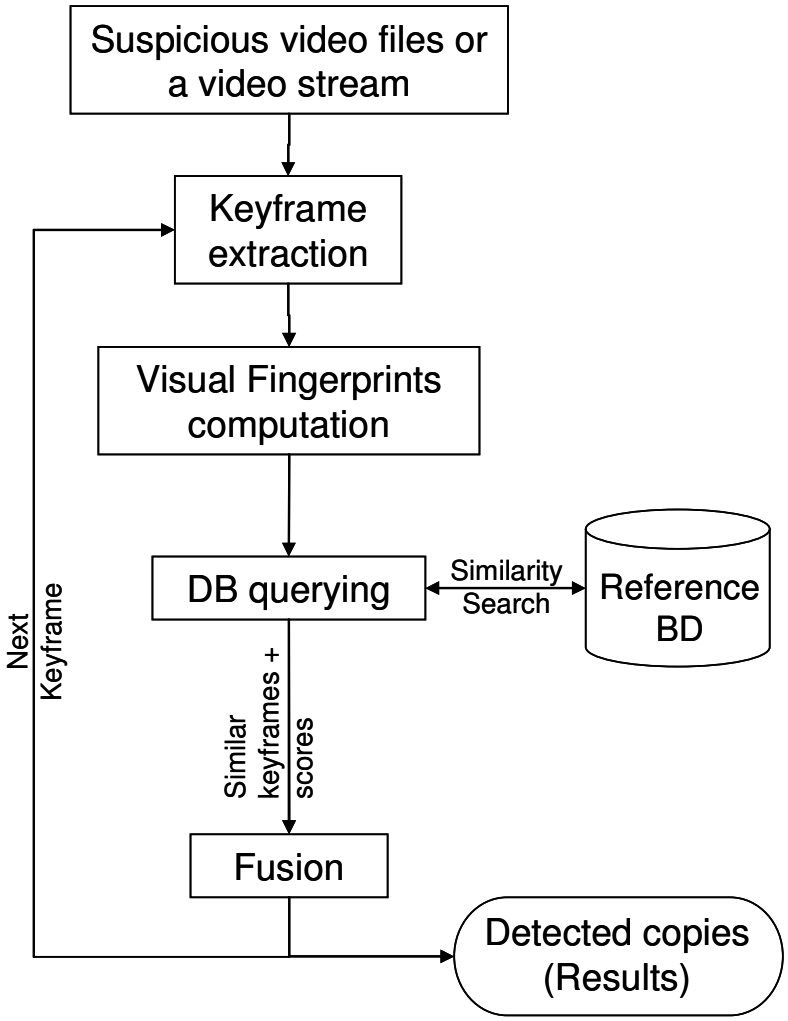
The Orange Labs System

- Off-line phase
  - Relies on visual fingerprints
  - Fingerprints are computed on keyframes
  - A local description scheme is used
  - Indexing relies on a hashing scheme



The Orange Labs System

- On-line phase
  - The same fingerprint computation scheme
  - The density of keyframes *may differ* from the off-line phase
  - Number of fingerprints per keyframes *may differ* from the off-line phase
  - Threshold computation is based on an *a contrario* approach
    [Gengembre et al., CBMI 2008]
  - Fusion relies on a Markovian framework
    [Gengembre and Berrani, CIVR 2008]



Experiments – Results

- Test Dataset Description
  - **201** query seq. × **10** trans. = **2010** queries (**1/3** are not from the reference DB).
  - Reference DB = **206** hours.
- Evaluation criteria
  - $NDCR = \frac{FN}{134} + \frac{FP}{21.5}$: false alarms are much more important than mis-detections.
  - Response time.
- Our three runs
  - Variation of the number of fingerprints per ***referenced*** keyframe.
  - Variation of the number of fingerprints per ***query*** keyframe.

Runs	Number of fingerprints per *referenced* keyframe	Number of fingerprints per *query* keyframe
Run1	150	100
Run2	200	200
Run3	200	300

Experiments – Results

- Robustness per transformation (MinNDCR)

	1	2	3	4	5	6	7	8	9	10	Mean
Run1	0.484	0.186	0.096	0.412	0.027	0.44	0.729	0.076	0.173	0.643	0.3266
Run2	0.48	0.188	0.111	0.433	0.061	0.435	0.754	0.077	0.184	0.643	0.3366
Run3	0.498	0.148	0.119	0.435	0.042	0.425	0.739	0.122	0.176	0.687	0.3391

- Efficiency

	Total processing time (s)	Acceleration factor w.r.t. real time
Run 1	41588	3.74
Run 2	50906	3.06
Run 3	57269	2.72

Experiments – Results

- Robustness per transformation (MinNDCR)

	1	2	3	4	5	6	7	8	9	10	Mean
Run1	0.484	0.186	0.096	0.412	0.027	0.44	0.729	0.076	0.173	0.643	0.3266
Run2	0.48	0.188	0.111	0.433	0.061	0.435	0.754	0.077	0.184	0.643	0.3366
Run3	0.498	0.148	0.119	0.435	0.042	0.425	0.739	0.122	0.176	0.687	0.3391

Camcording with ratio variations

Strong re-encoding

Combination of other transformations

- Efficiency

	Total processing time (s)	Acceleration factor w.r.t. real time
Run 1	41588	3.74
Run 2	50906	3.06
Run 3	57269	2.72

Experiments – Results

- Examples of copies correctly identified

Copy

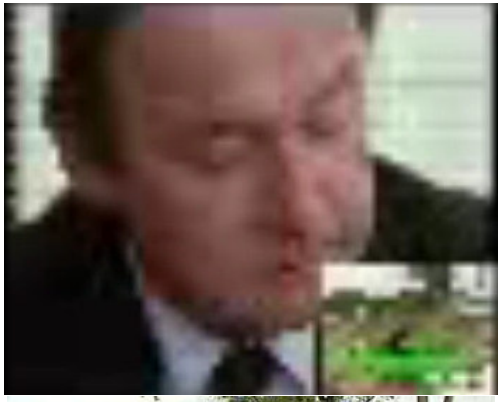
Original

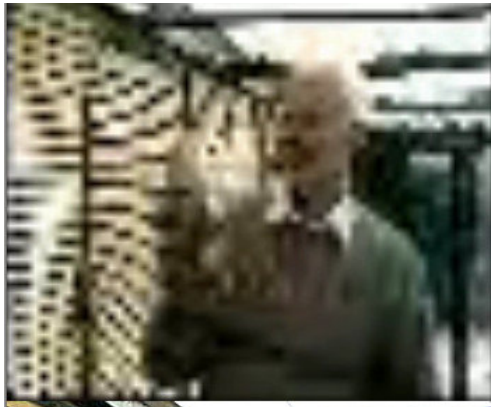
Experiments – Results

- Examples of copies not correctly identified

Copy

Original

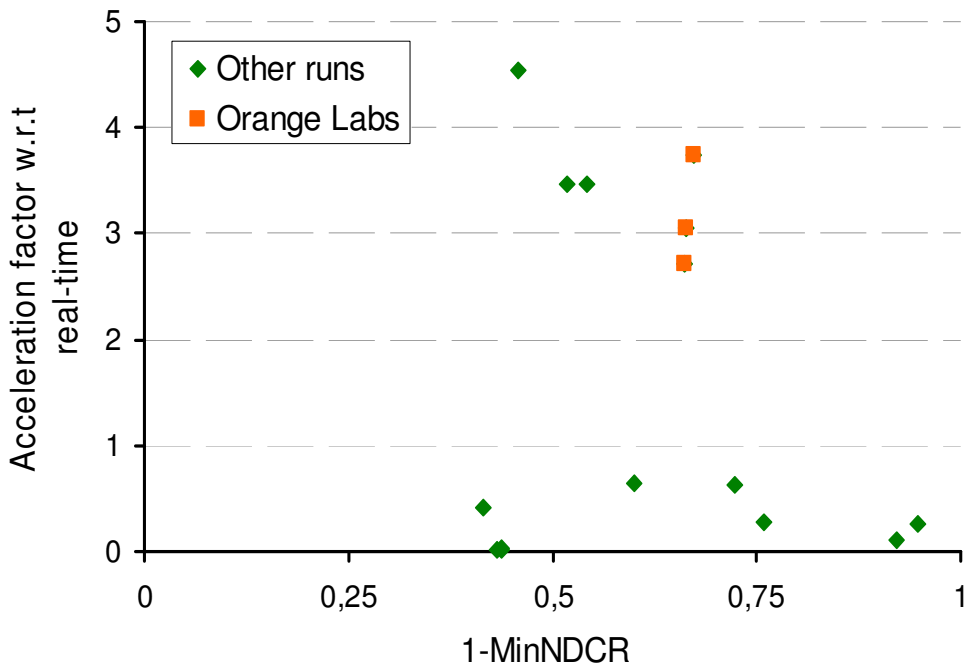




Experiments – Results

- Trade-off effectiveness/efficiency

- Our 3 runs are ranked 5th, 6th and 7th in terms of effectiveness (MinNDCR).
- They are between 5 and 30 times faster than the 4 most effective runs.



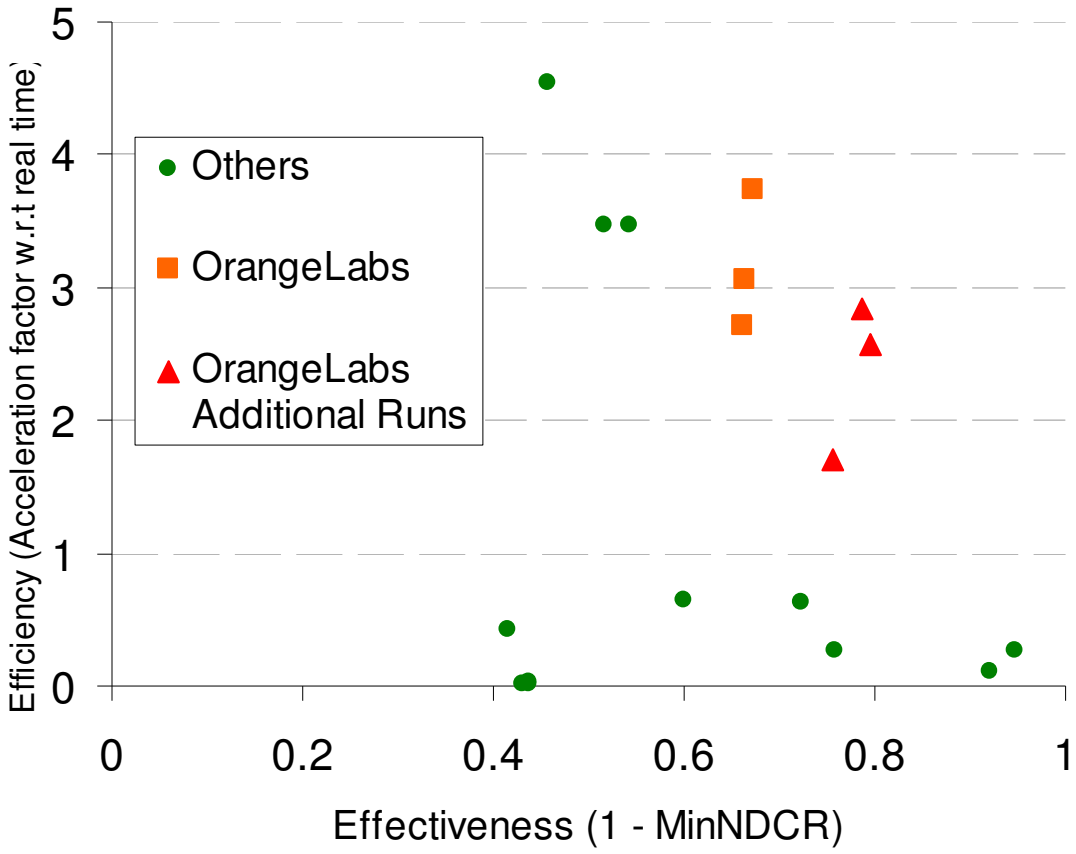
Efficiency vs. Effectiveness for the 15 most effective runs.

Experiments – Results

- Additional runs

- Use:
  - More keyframes and/or
  - Improved descriptors
- Still faster than real-time
- The 3rd most effective
- Robustness to trans. 4 improved (strong reencoding):

  MinNDCR: 0.41 → 0.2



Experiments – Results

- Detection threshold

Optimal threshold per transformation as computed for TrecVid 2008

	1	2	3	4	5	6	7	8	9	10	All
Run1	0.484	0.186	0.096	0.412	0.027	0.44	0.729	0.076	0.173	0.643	0.3266
Run2	0.48	0.188	0.111	0.433	0.061	0.435	0.754	0.077	0.184	0.643	0.3366
Run3	0.498	0.148	0.119	0.435	0.042	0.425	0.739	0.122	0.176	0.687	0.3391

A unique and a priori fixed threshold = 100

	1	2	3	4	5	6	7	8	9	10	All
Run1	0.486	0.232	0.076	0.444	0.015	0.420	0.699	0.087	0.188	0.680	0.3327
Run2	0.483	0.180	0.074	0.459	0.020	0.402	0.711	0.069	0.152	0.643	0.3193
Run3	0.490	0.132	0.078	0.430	0.020	0.383	0.687	0.081	0.156	0.620	0.3077

- Results almost identical
- It is not realistic to choose a different threshold for each transformation
- The ability to provide the optimal threshold is also an important evaluation criterion

Conclusions

- Trade-off effectiveness/efficiency:
  - The two criteria have to be considered together
- No need to be invariant to transformations that make the content useless
- Threshold selection = an important evaluation criterion for system evaluation
- It would be interesting to evaluate the scalability: 206 hours is a small database (~135 movies or ~9 days of TV broadcast)