Interactive Search Using a Lexicon too Large to Remember

Marcel Worring, Bouke Huurnink,

Ork de Rooij, Cees Snoek, Andy Thean[1]

Intelligent Systems Lab Amsterdam,
University of Amsterdam, The Netherlands

[1]TNO, The Netherlands

MultimediaN

Interactive Retrieval

Text does not always describe the visual content

Information need

Query

Nice, but only if you have examples to begin with …

Our preferred choice

Textual?

Visual?

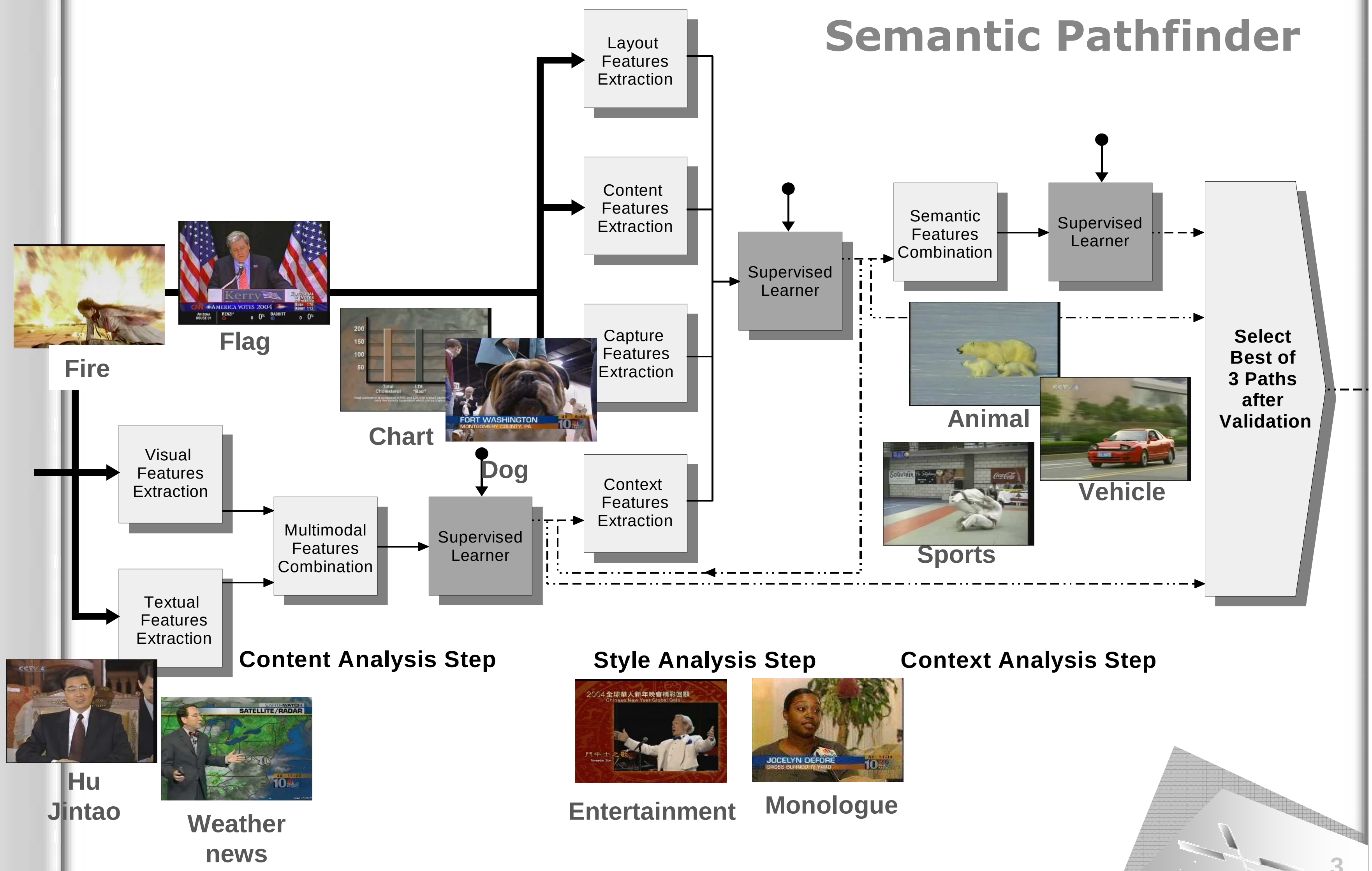
Semantic?

Saturday evening a hotel in Gaithersburg got on fire due to an Elvis look-a-like who played with fire on stage.

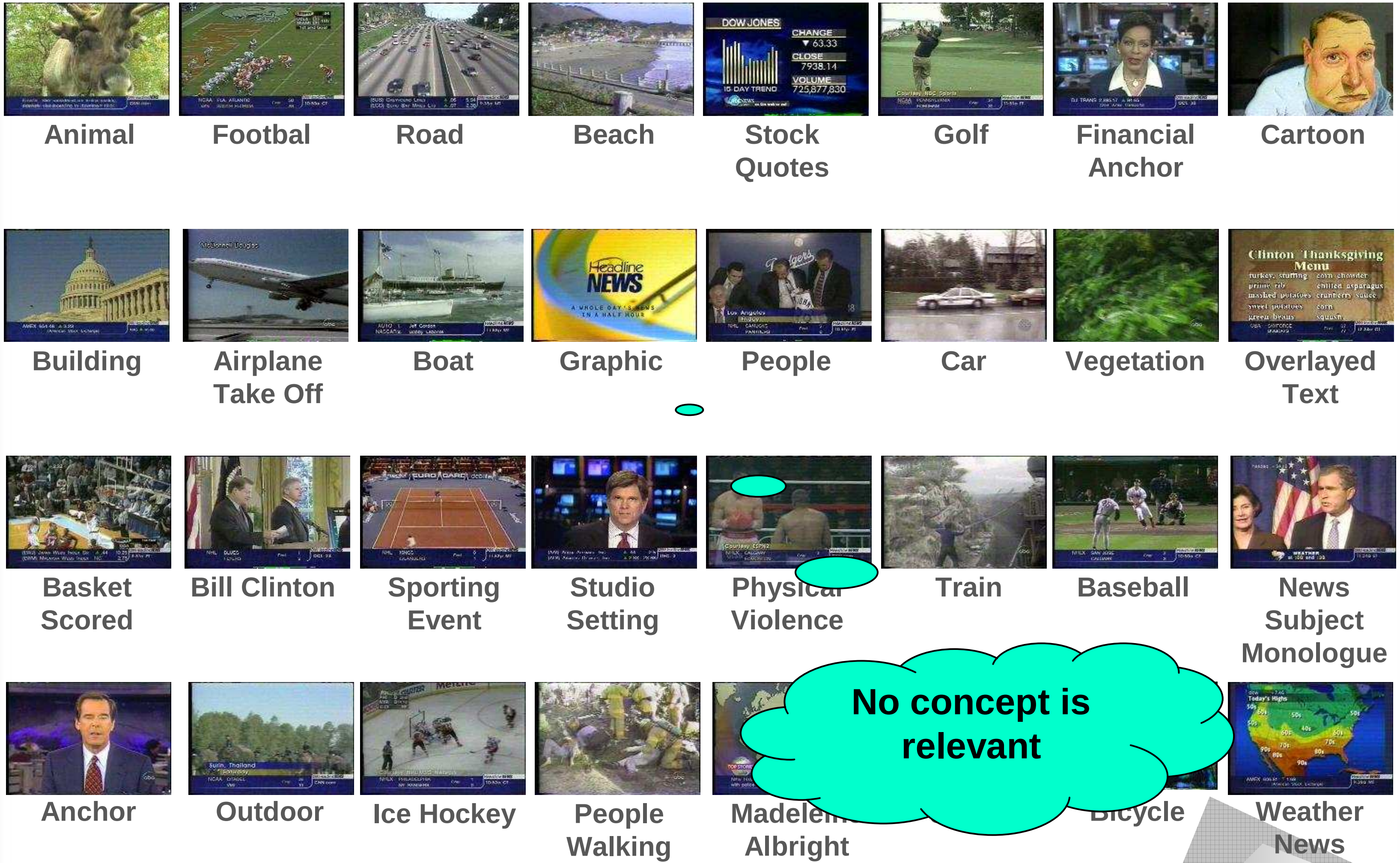
…………..

Find fragments such as this in the database

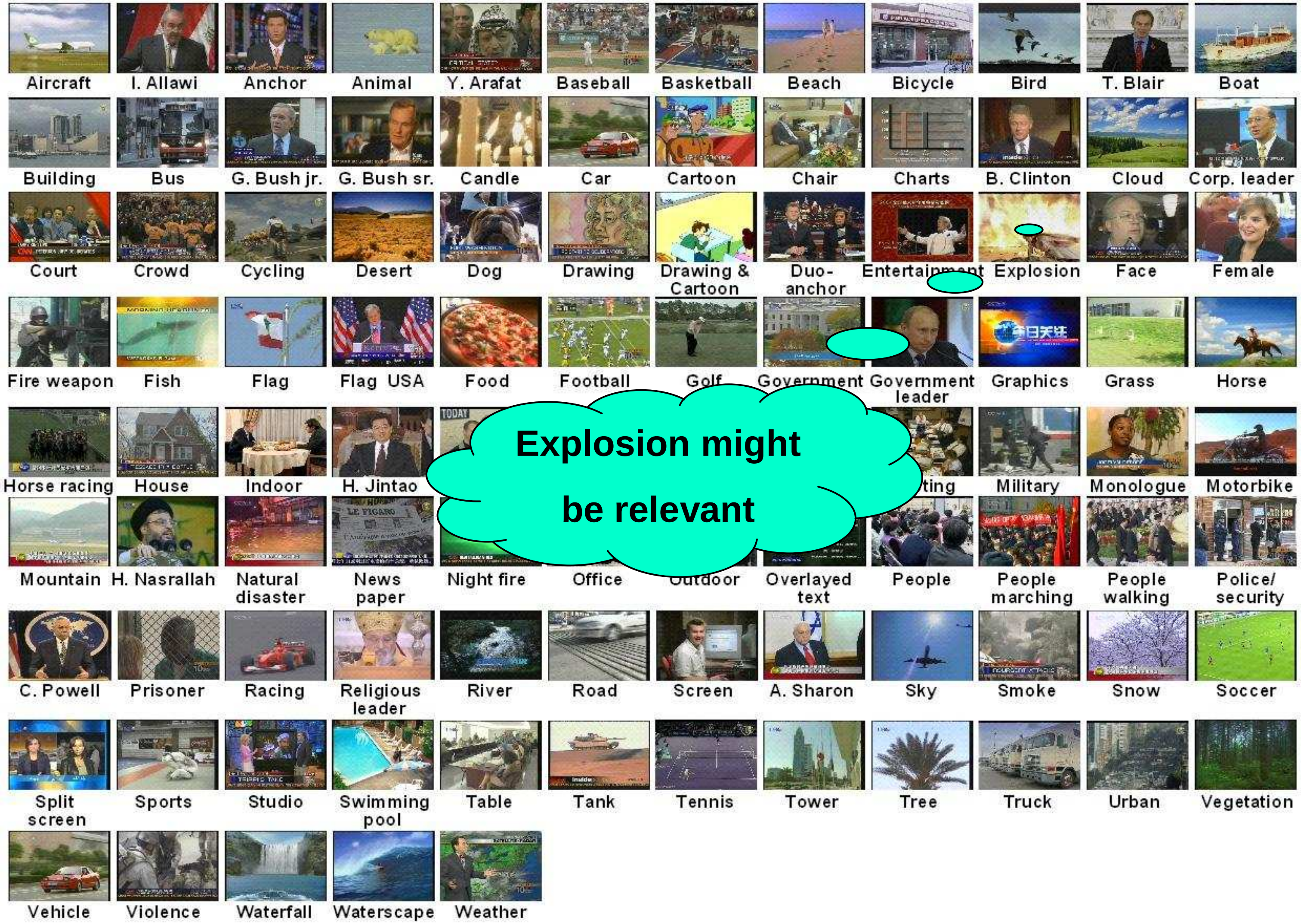
Generic Semantic Indexing



Learned lexicon of 32 concepts



Learned lexicon of 101 concepts



2006: 491 concepts

000_parade
001_exiting_car
003_running
004_airplane_crash
005_earthquake
006_demonstration_or_protest
007_people_crying
008_airplane_takeoff
009_airplane_landing
168_dredge_powershovel_dragline
010_helicopter_hovering
013_singing
019_speaking_to_camera
020_riot
022_tornado
025_flood
027_talking
028_dancing
029_car_crash
030_funeral
033_cheering
034_greeting
035_throwing
036_shooting
037_address_or_speech
038_bomber_bombing
039_celebration_or_party
040_airport
042_castle
043_college
044_courthouse
045_fire_station
046_gas_station
050_hospital
051_hotel
052_house_of_worship
053_police_station
054_power_plant
055_processing_plant
056_school
057_shopping_mall
058_stadium
059_supermarket
060_airport_or_airfield
061_aqueduct
063_river_bank
064_aircraft_cabin
065_canal
068_cityscape
070_conference_room
071_construction_site
072_graveyard
073_highway
074_hospital
075_industrial_setting
077_military_base
079_ruins
080_suburban
082_underwater
083_adobehouses
084_laboratory
086_tent
087_beach
089_parking_lot

090_ditch
091_golf_course
093_warehouse
094_airport_terminal
095_bazaar
097_embassy
098_foxhole
099_hill
100_marsh
101_urban_park
102_subway_station
105_civilian_person
106_sitting
107_standing
109_windows
110_female_anchor
111_female_reporter
112_first_lady
113_male_anchor
114_male_reporter
115_commercial_advertisement
116_armed_person
117_firefighter
118_judge
119_athlete
120_congressman
121_logos_full_screen
122_high_security_facility
123_emergency_medical_response_people
124_election_campaign
125_airplane_flying
126_female_news_subject
127_golf_player
128_politics
129_press_conference
130_celebrity_entertainment
131_swimming
132_bride
133_golf_caddy
134_construction_worker
135_toll_booth
136_guard
137_hunter
138_clock_tower
139_factory_worker
140_steeple
141_groom
142_ground_crew
143_election_campaign_convention
144_election_campaign_debate
145_election_campaign_greeting
147_security_checkpoint
149_actor
150_head_and_shoulder
151_street_battle
152_tractor_combine
153_landscape
154_alley
155_attached_body_parts
156_hand

157_indoor_sports_venue
158_bar_pub
159_emergency_room
162_forest
163_grassland
165_lawn
167_administrative_assistant
169_gym
172_kitchen
175_clearing
176_dining_room
178_fighter_combat
002_handshaking
180_individual
181_adult
182_agent
183_boy
184_girl
185_lawyer
186_mug
188_glass
189_security_checkpoint
191_commentator_or_studio_expert
192_dead_bodies
193_eyewitness
195_finance_busines
194_male_news_subject
196_science_technology
198_host
199_guest
200_ground_combat
205_walking_running
217_person
218_airplane
219_government_leader
225_news_studio
238_control_tower_-_airport
239_studio_with_anchorperson
240_animal_pens_and_cages
241_antenna
242_apartments
243_apartment_complex
244_armored_vehicles
245_artillery
246_asian_people
247_baby
248_backpackers
249_backpack
251_barge
254_beards
256_bicycles
258_dark-skinned_people
259_blank_frame
260_bridges
261_briefcases
262_business_people
263_cables
265_camera
266_canoe
267_capital
269_cart_path
270_cats
271_caucasians
272_cell_phones

273_child
274_cigar_boats
275_classroom
277_cloverleaf
280_communications_tower
281_computers
282_conference_buildings
283_construction_vehicles
285_cordless
286_cows
287_cruise_liner
288_cul-de-sac
289_flying_objects
290_daytime_outdoor
291_dirt_gravel_road
295_dresses
296_dresses_of_women
297_driver
302_factory
299_emergency_vehicles
300_empire_state
301_exploding_ordinance
303_farms
304_fields
307_flowers
308_pedestrian_zone
309_free_standing_structures
310_freighter
311_frigate
312_furniture
314_glasses
315_grandstands_bleachers
316_group
317_handguns
319_harbors
321_helicopters
324_hu_jintao
326_infants
331_islands
327_insurgents
329_interview_sequences
330_interview_on_location
333_john_edwards
335_lakes
337_landlines
339_body_parts
340_machine_guns
341_medical_personnel
342_microphones
345_mosques
348_muddy_scenes
349_cutter
350_muslims
351_newspapers
352_nighttime
354_non-uniformed_fighters
356_oceans
357_office_building
358_officers
359_old_people
360_outer_space
362_road_overpass
363_pavilions
364_peacekeepers
365_agricultural_people

366_telephones
368_photographers
369_pickup_truck
370_pipelines
371_pipes
372_police
374_porches
375_powerplants
376_coal_powerplants
378_nuclear_powerplants
380_power_transmission_line_tower
381_powerlines
382_protesters
384_radar
385_raft
386_railroad
387_rainy
390_reporters
391_residential_buildings
392_rifles
393_road_block
394_traffic
395_rocky_ground
396_rpg
398_room
399_rowboat
400_runway
401_rv_farm
402_sailboat
403_scene_text
405_ship
406_shipyards
407_single_family_homes
408_single_person_female
409_single_person_male
410_single_person
411_smoke_stack
412_still_image
413_soldiers
414_speaker
415_sports
416_stock_m
417_store
418_streets
419_striking_people
421_suits
422_sunglasses
423_sunny
424_swimmer
425_swimming_poo
426_tanks
429_text_label
430_text_o
und
432_ties
433_tony_bla
436_tropical_
437_tugboat
440_valleys
442_sidewalks
444_waterways
445_weapons
446_white_house
447_windy

463_airport
484_baker
487_banks
490_barracks
535_clocks
541_network_logo
609_ground_vehicles
677_military_buildings
685_moonlight
697_non-us_national_flags
702_observation_tower
aircraft
animal
boat
building
bus
car
charts
corporate_leader
court
crowd
desert
entertainment
explosion
face
flag_usa
government_leader
maps
meeting
military
mountain
natural_disaster
office
outdoor
people
people_marching
police_security
prisoner
road
screen
sky
snow
sports
studio
truck
urban
vegetation
clinton

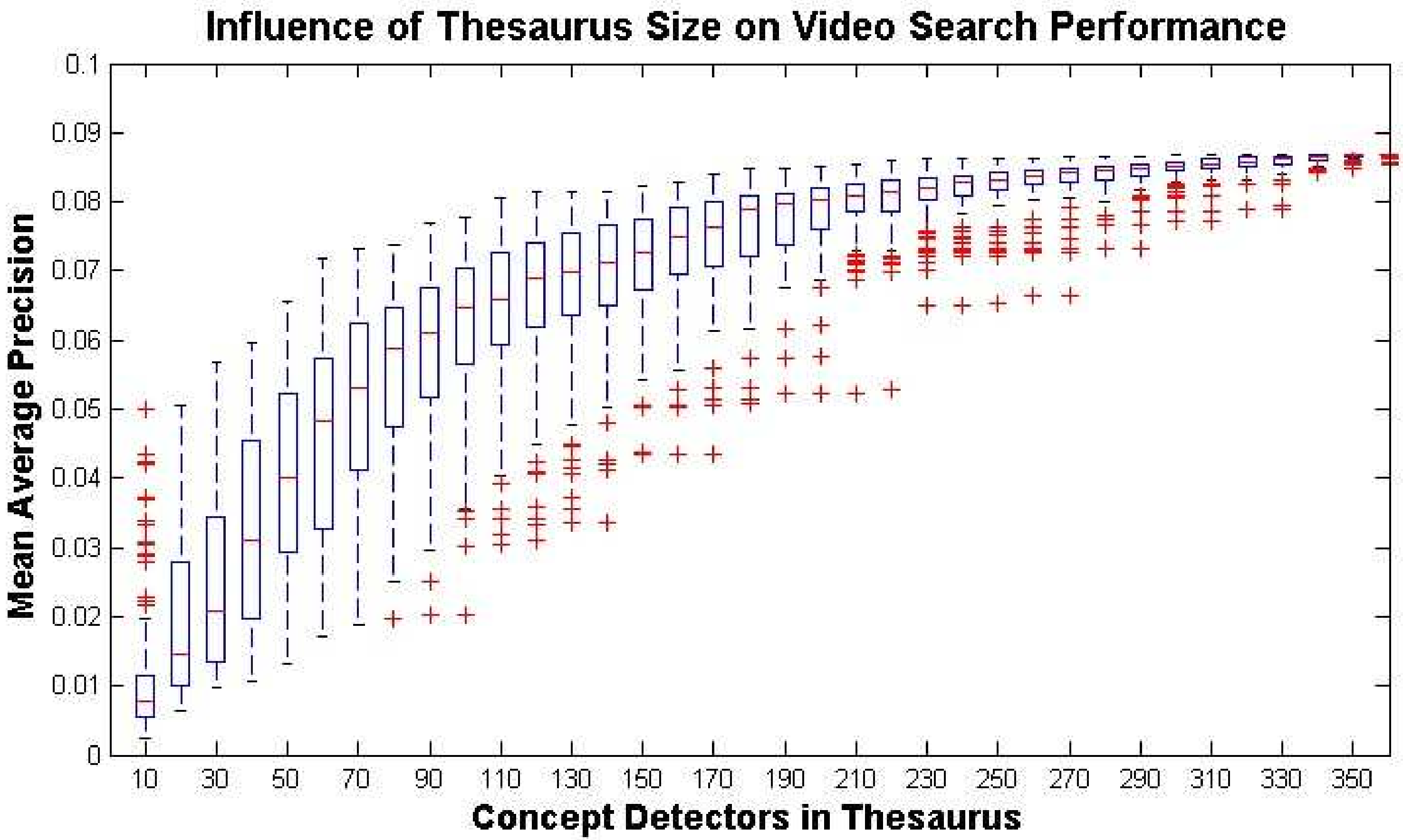
cloud
cycling
dog
drawing
drawing_cartoon
duo_anchor
female
fireweapon
fish
flag
food
football
golf
government_building
graphics
grass
hassan_nasrallah
horse
horse_racing
house
hu_jintao
indoor
kerry
lahoud
male
monologue
motorbike
newspaper
nightfire
overlayed_text
powell
racing
religious_leader
river
sharon
smoke
soccer
splitscreen
swimmingpool
table
tank
tennis
tony_blair
tower
tree
esert
33_boat_ship

257_birds
268_car_racing
276_clouds
292_dogs
306_flags
320_sketches
383_colin_powell
014_baseball
015_basketball
103_female_person
229_weather
231_military_personnel
232_truck
234_prisoner
235_vegetation
236_mountain
313_george_bush
011_golf
347_motorcycle
361_overlaid_text
012_walking
016_football
021_natural_disasters
085_office
104_male_person
164_house
179_head_of_state
204_maps
215_people_marching
220_meeting
224_outdoor
228_police_private_security_personnel
108_vehicle
161_smoke
206_road
207_sky
208_urban_scenes
209_waterscape_waterfront
346_religious_figures
434_tower
435_trees
017_soccer
018_tennis
024_snow
078_river
448_yasser_arafat

Too many
to remember

Should we bother?

TRECVID2005



Yes size does matter.

So we need to help the user in selecting the right concept

And how to browse for more?

Textual?

Time?

Saturday evening a hotel in Gaithersburg got on fire due to an Elvis look-a-like who played with fire on stage.

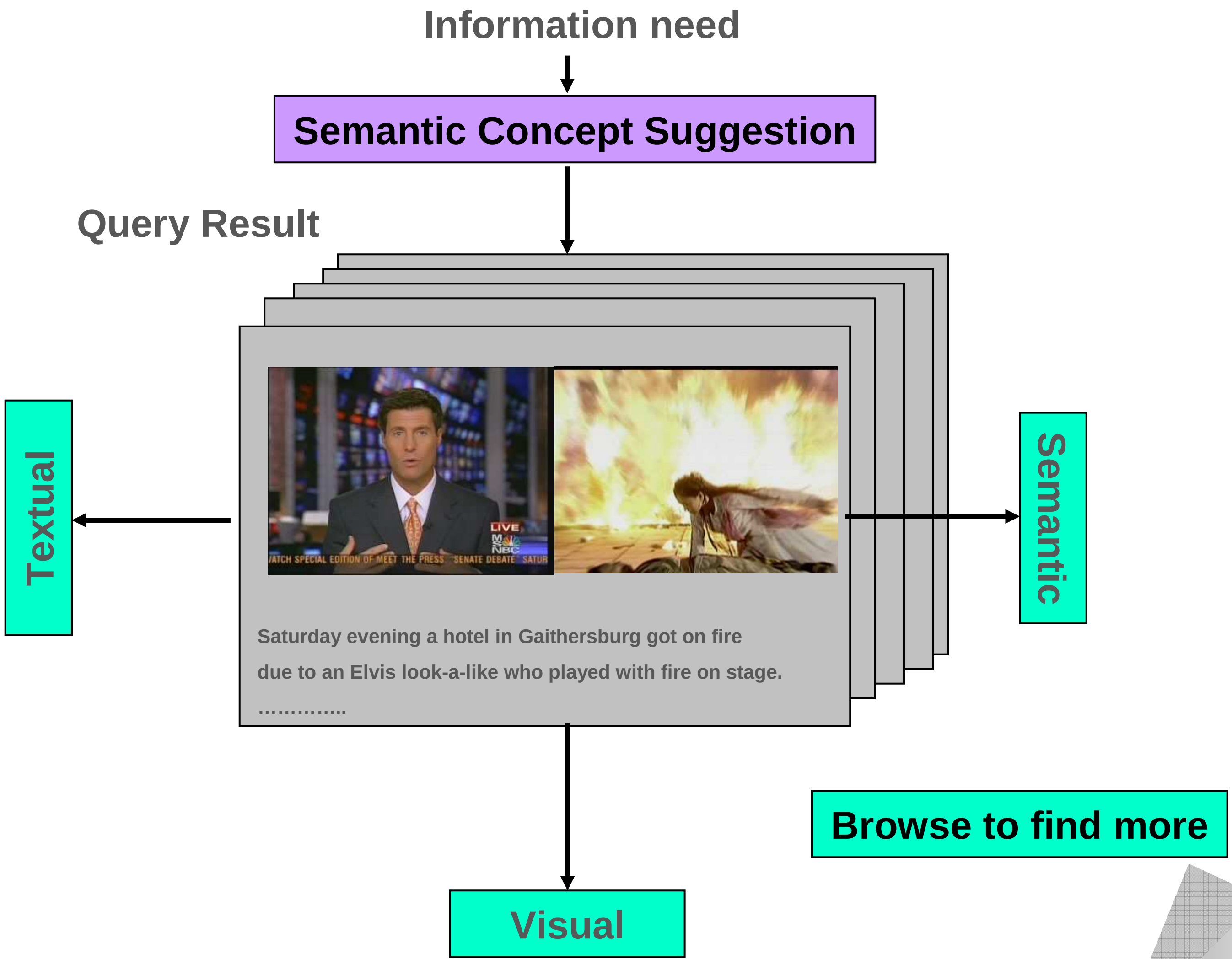
…………..

Semantic?

Visual?

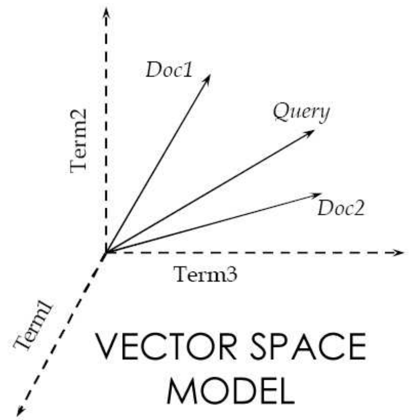
Depends on what you are looking for

Overview of the system

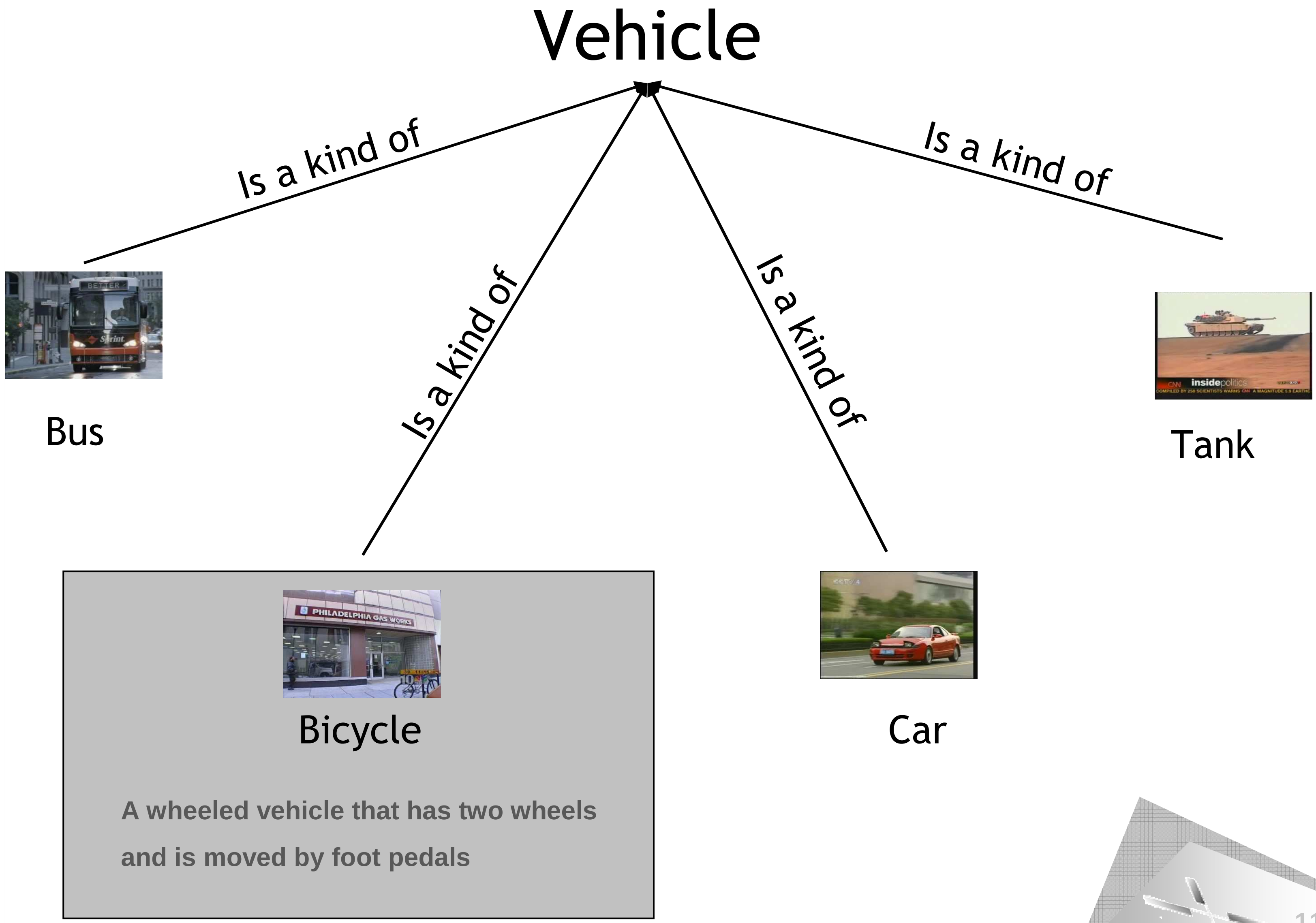


Suggestion based on text matching

- Index concept descriptions
  - Represent as term vector
  - Only 363, so rather small collection
- Need to increase recall?
  - Porter stemming algorithm
  - Character *n*-gramming, here sequences of 4 characters
- We use the vector space model to match queries to descriptions
  - Pick detector that maximizes query/document similarity
- Turns out that perfect match yields best performance



Next step: use Wordnet

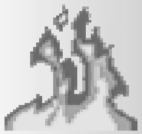


Suggestion using ontology querying

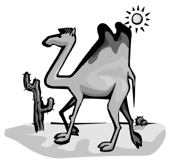
"Find a report from the desert showing a house or car on fire."

1. Identify objects in WordNet

car



fire



desert

house

References:
Bouke Huurnink (UvA) & Laura Hollink (VU)

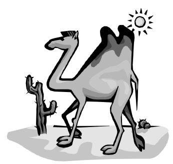
Ontology querying

"Find a report from the desert showing a house or car on fire."

2. Identify related concept detectors

car

fire



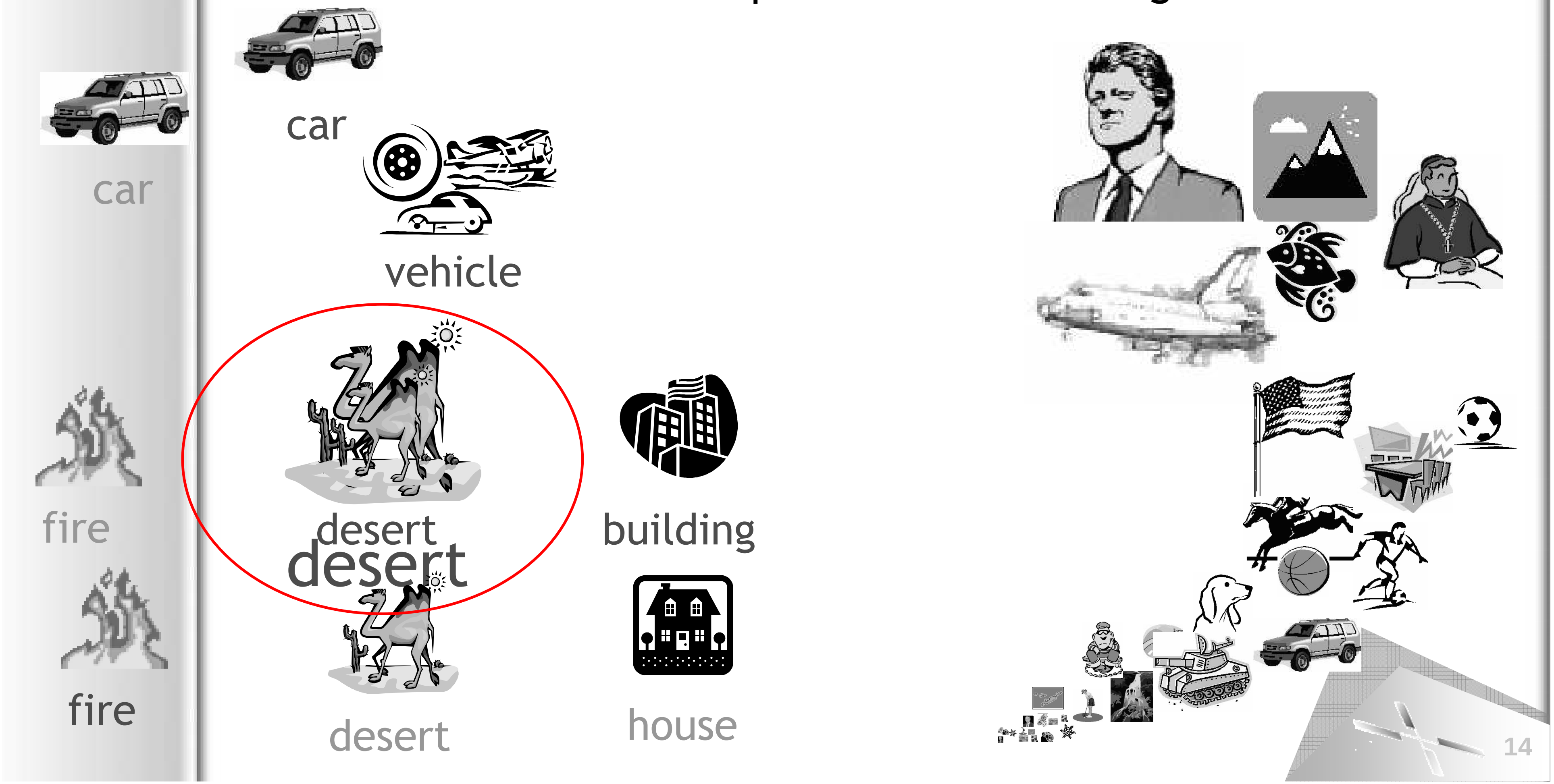
desert

house

Ontology querying

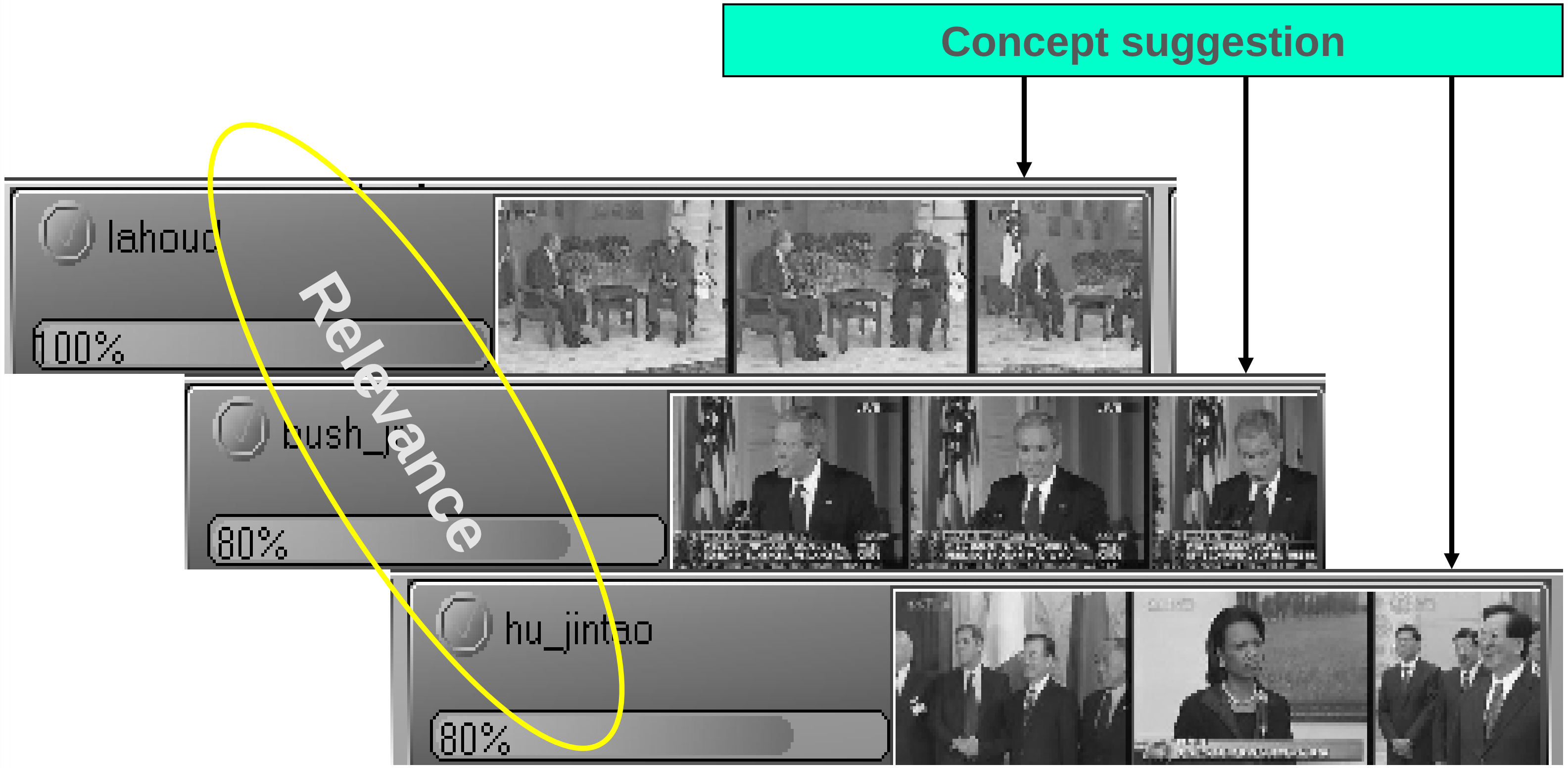
“Find a report from the desert showing a house or car on fire.”

3. Find most similar and specific detector using Resnik’s measure



Automatic concept suggestion

Query: Find shots of the current president of America



Overview

Information need

Semantic Concept Suggestion

Query Result

Saturday evening a hotel in Gaithersburg got on fire
due to an Elvis look-a-like who played with fire on stage.
…………..

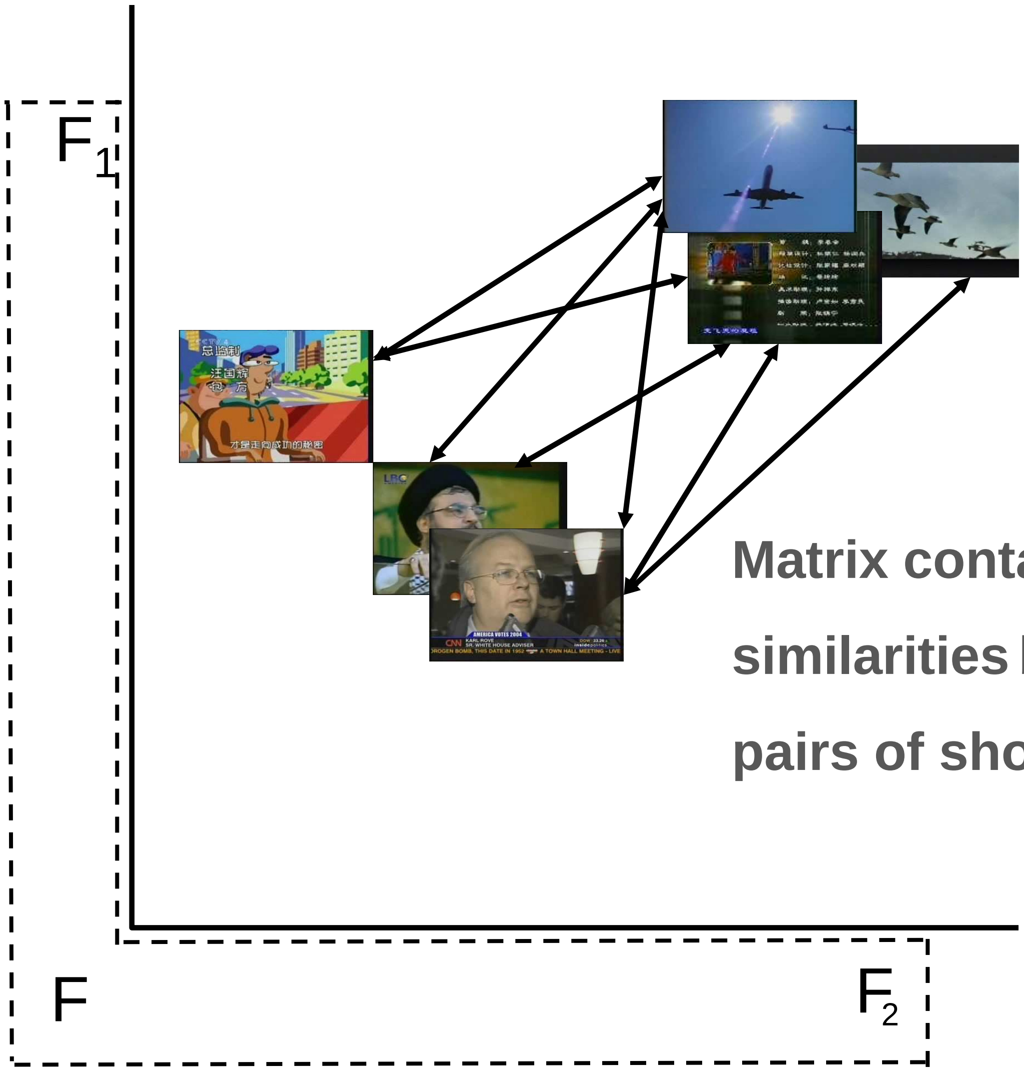
Time

Semantic

Visual

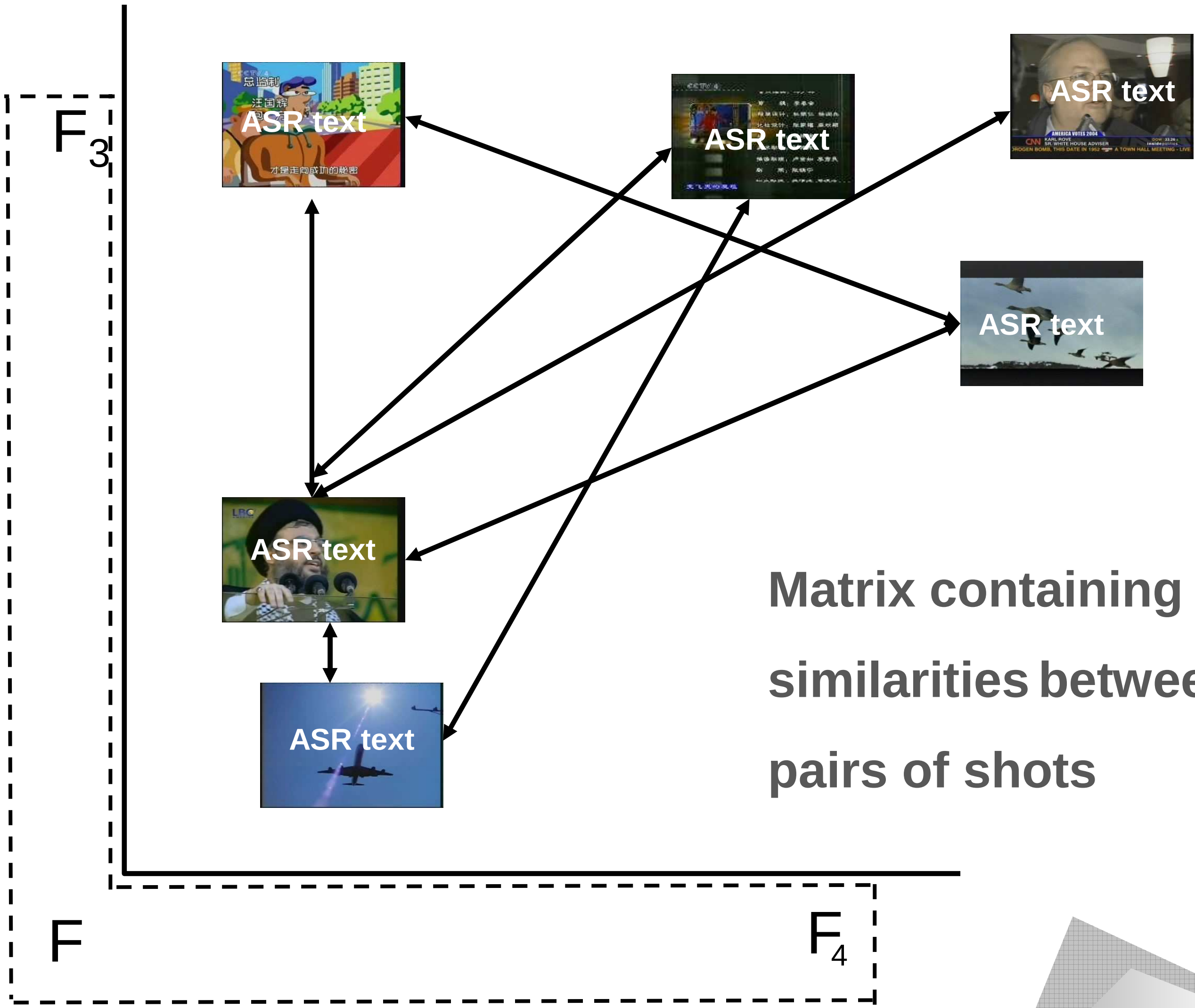
Browse to find more

Visual Similarity Space



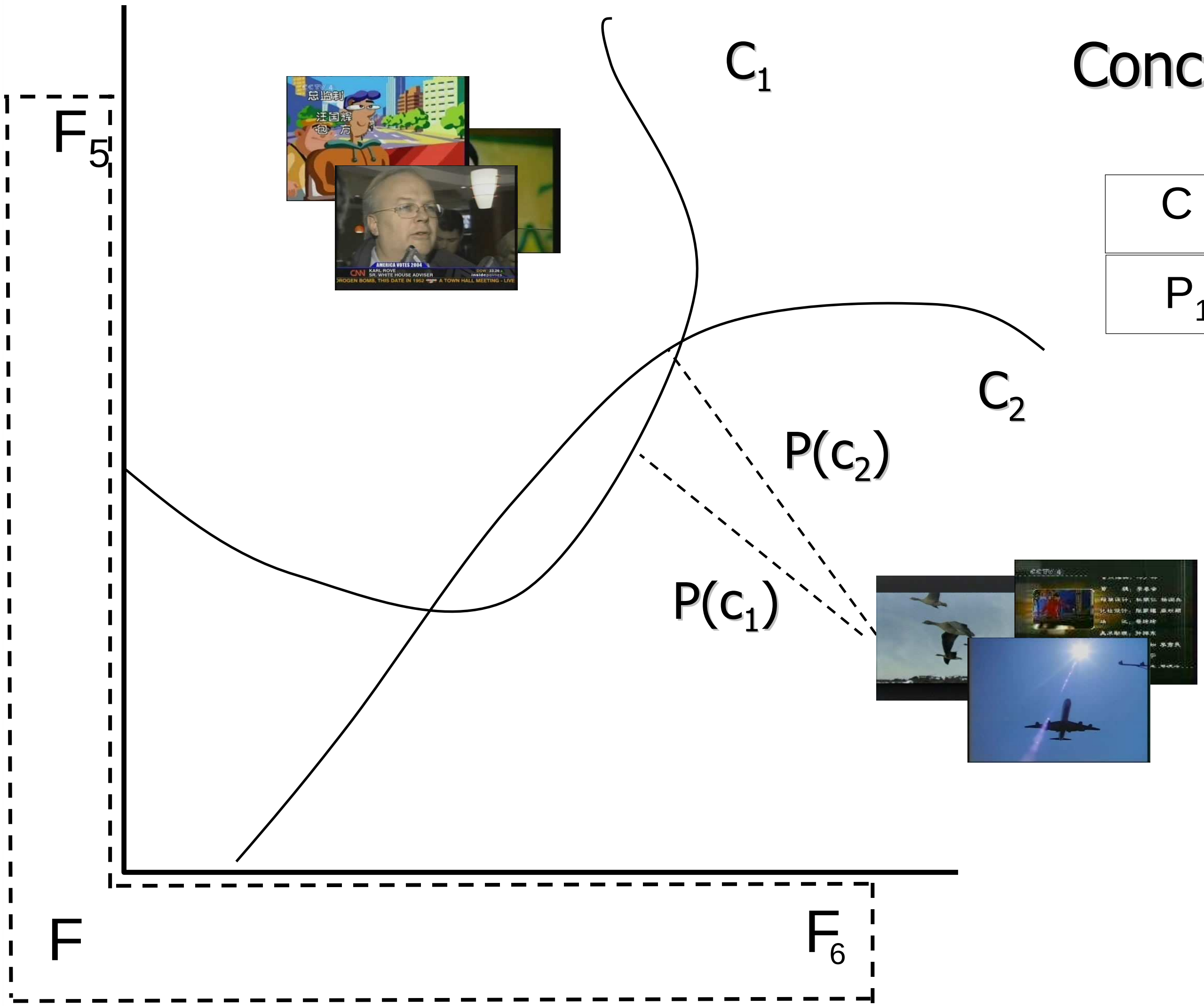
Matrix containing all similarities between pairs of shots

Textual Similarity Space



Matrix containing all similarities between pairs of shots

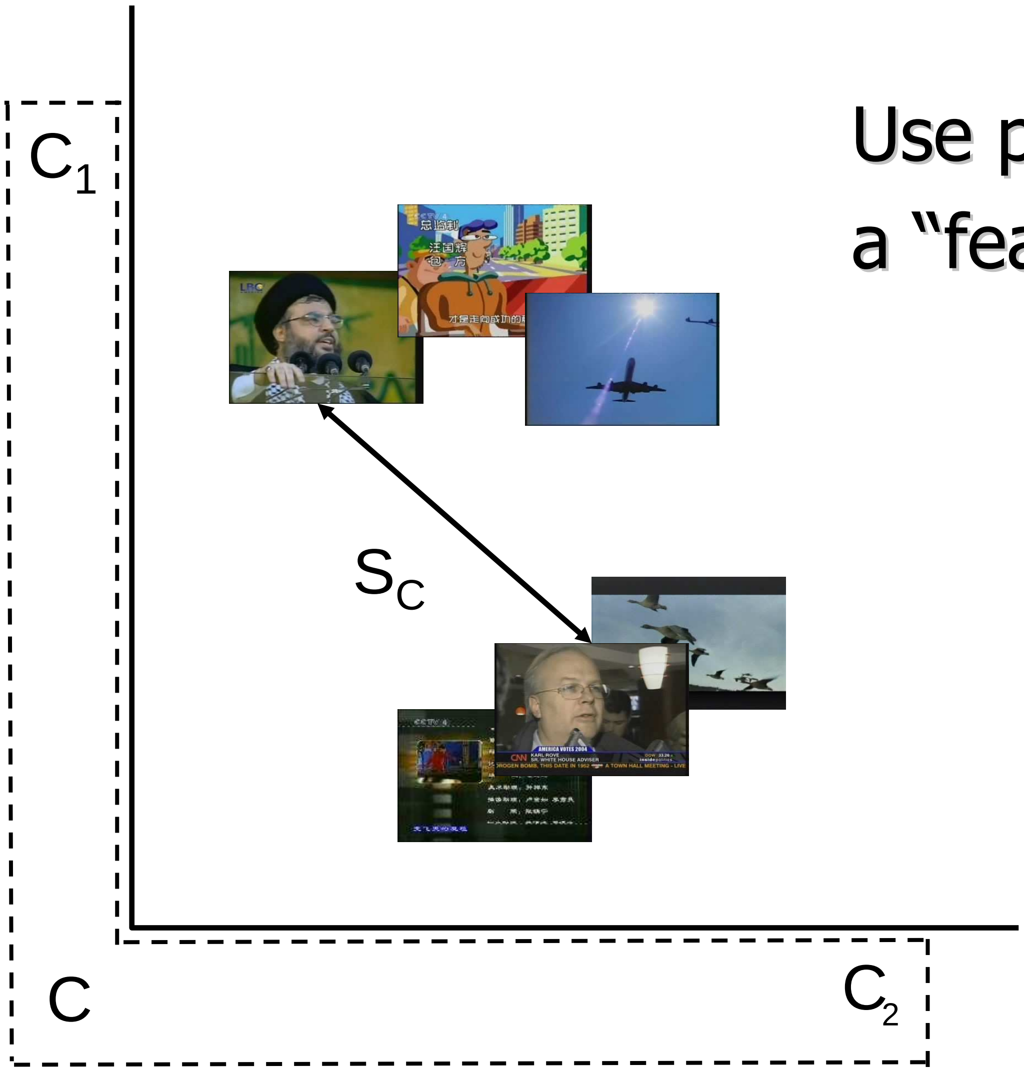
Semantic probabilities



Concept Lexicon

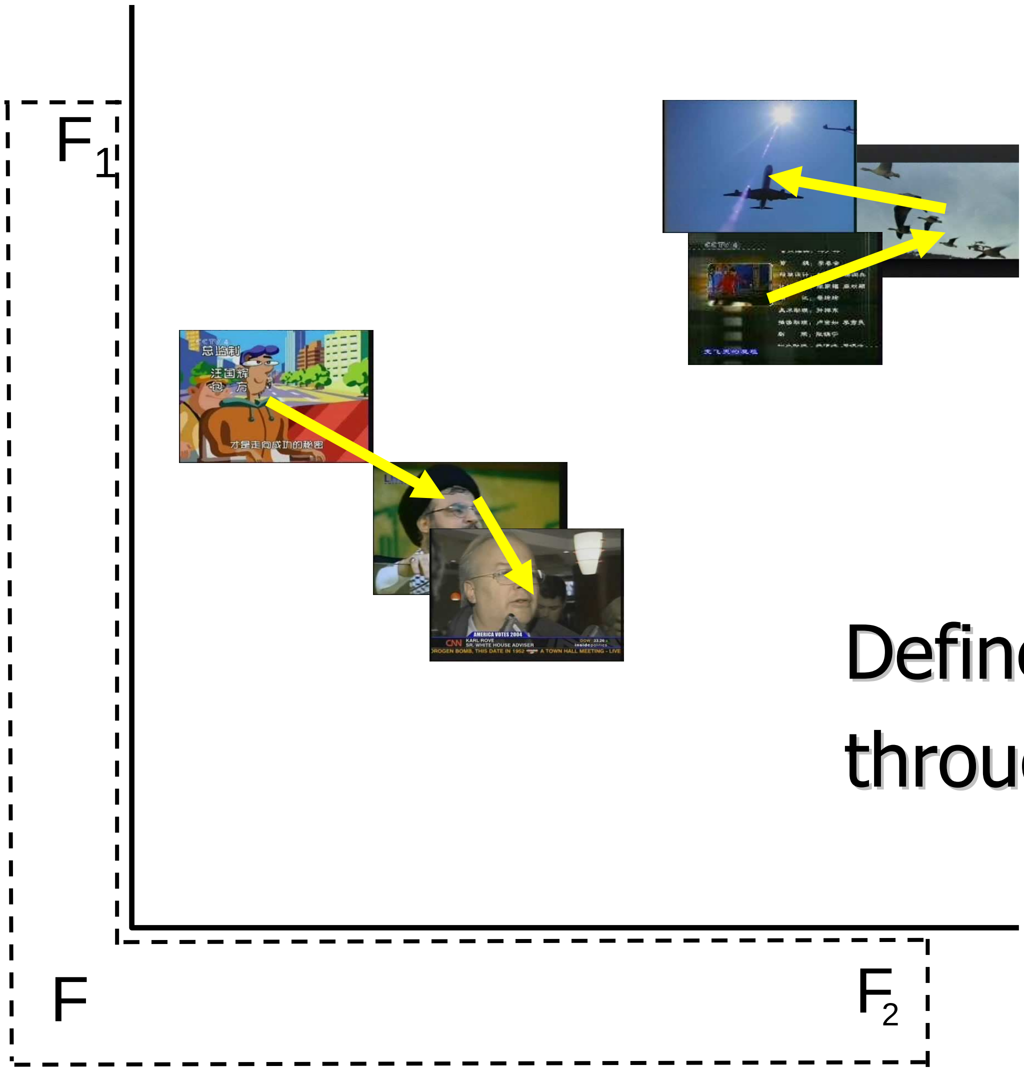
C_1	C_2
P_1	P_2

Semantic Similarity Space



Use probabilities as a “feature vector”

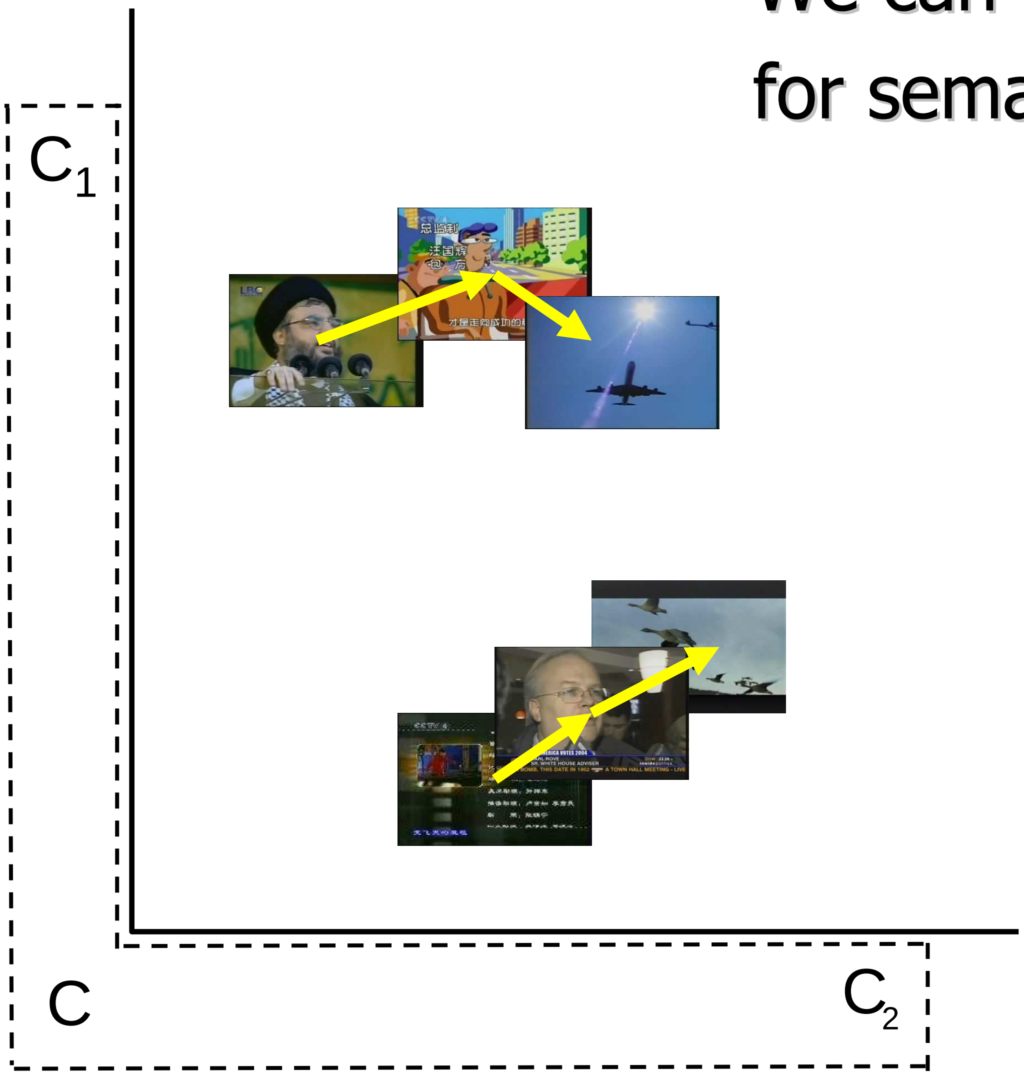
Visual Thread Space



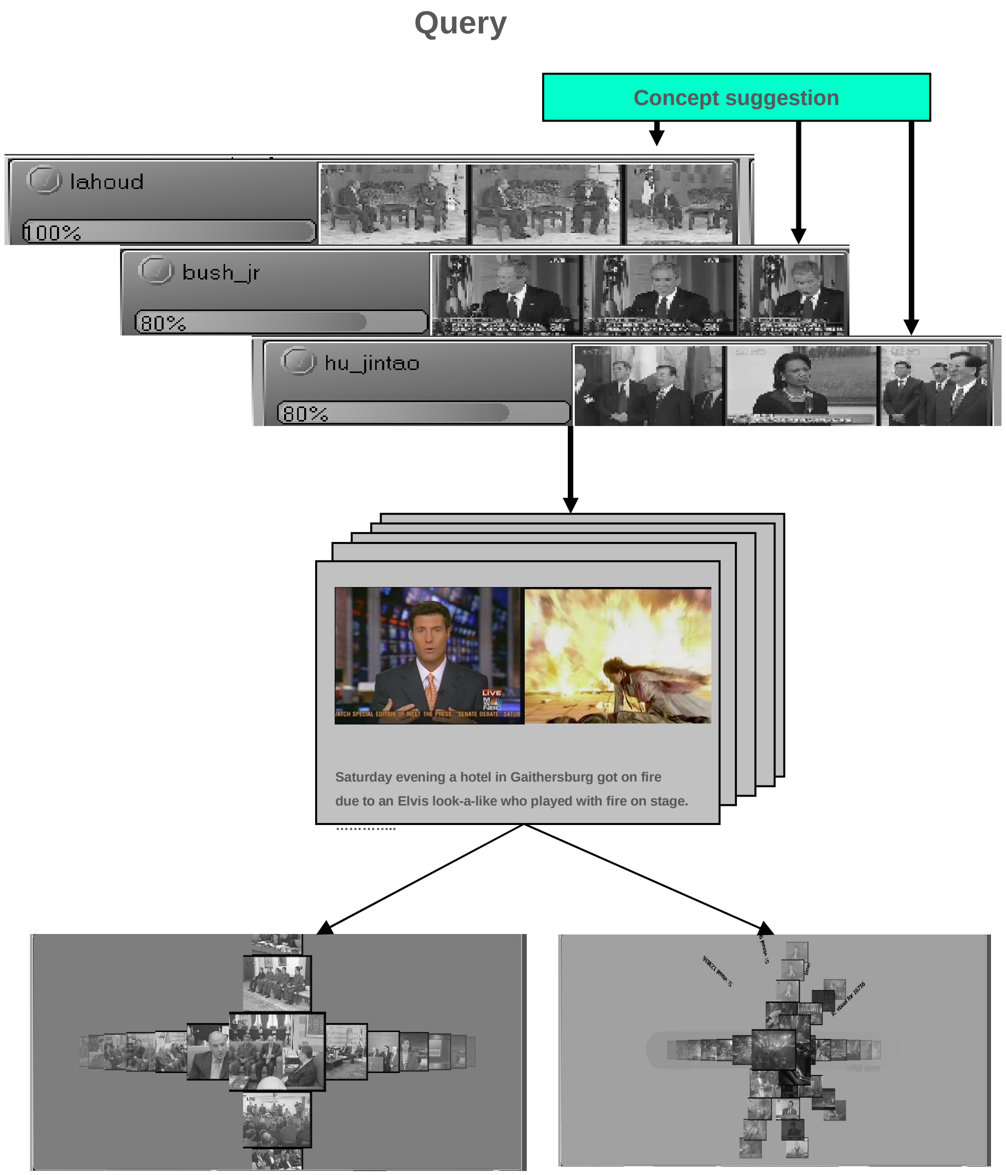
Define a shortest path through each cluster

Semantic Thread Space

We can do the same for semantics (and text)



Overview

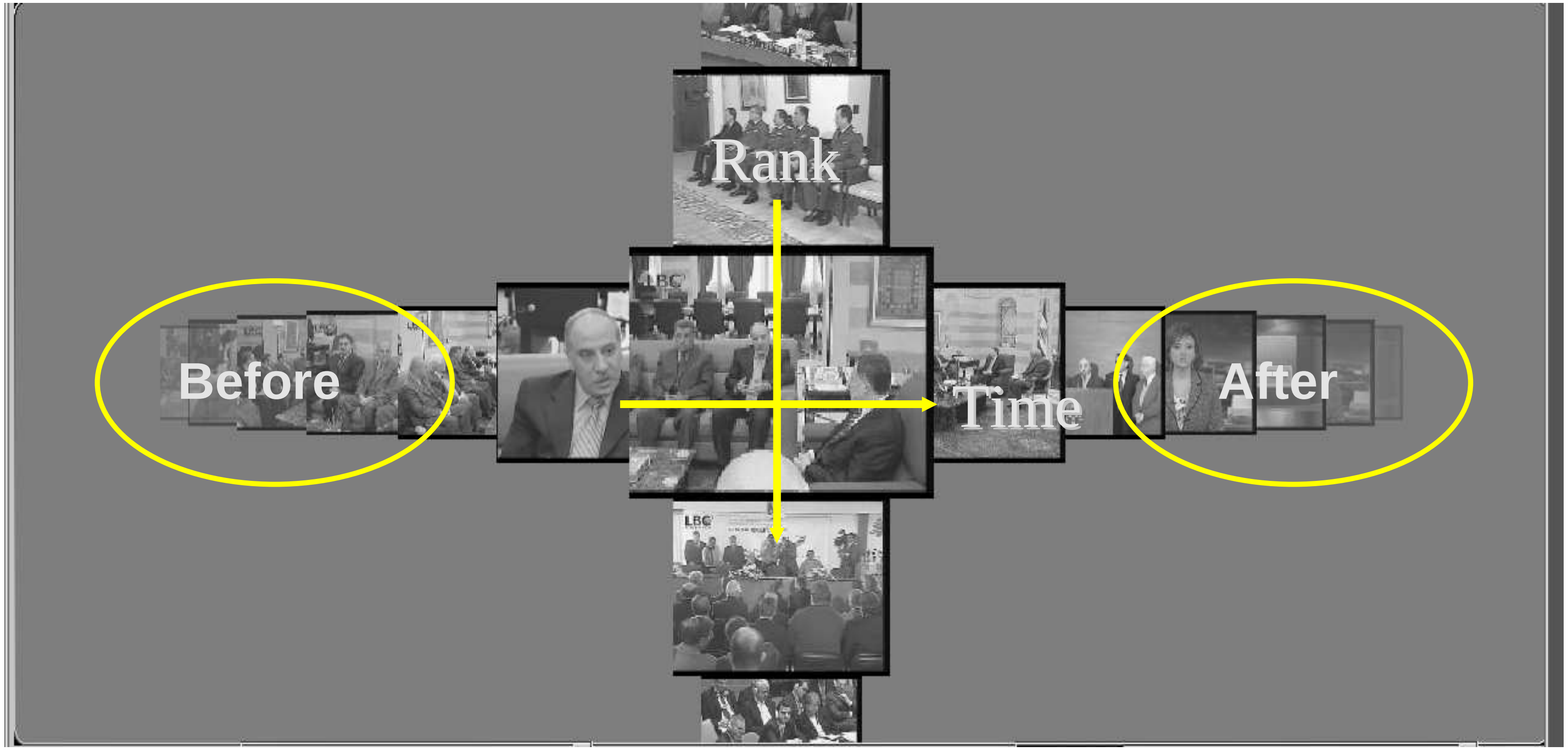


Semantic Concept Suggestion

Query result

Browse to find more

The CrossBrowser

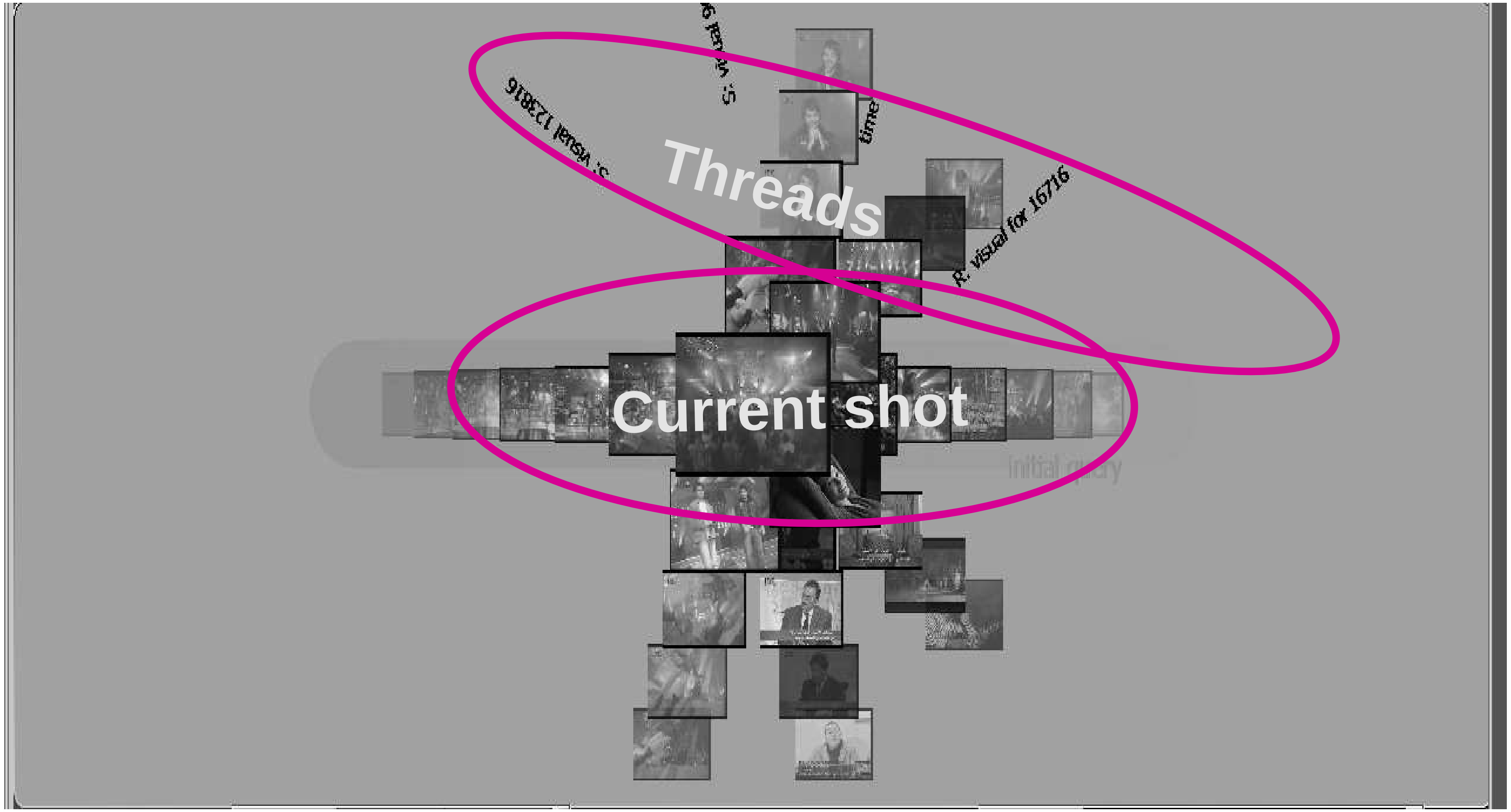


Other browsing dimensions

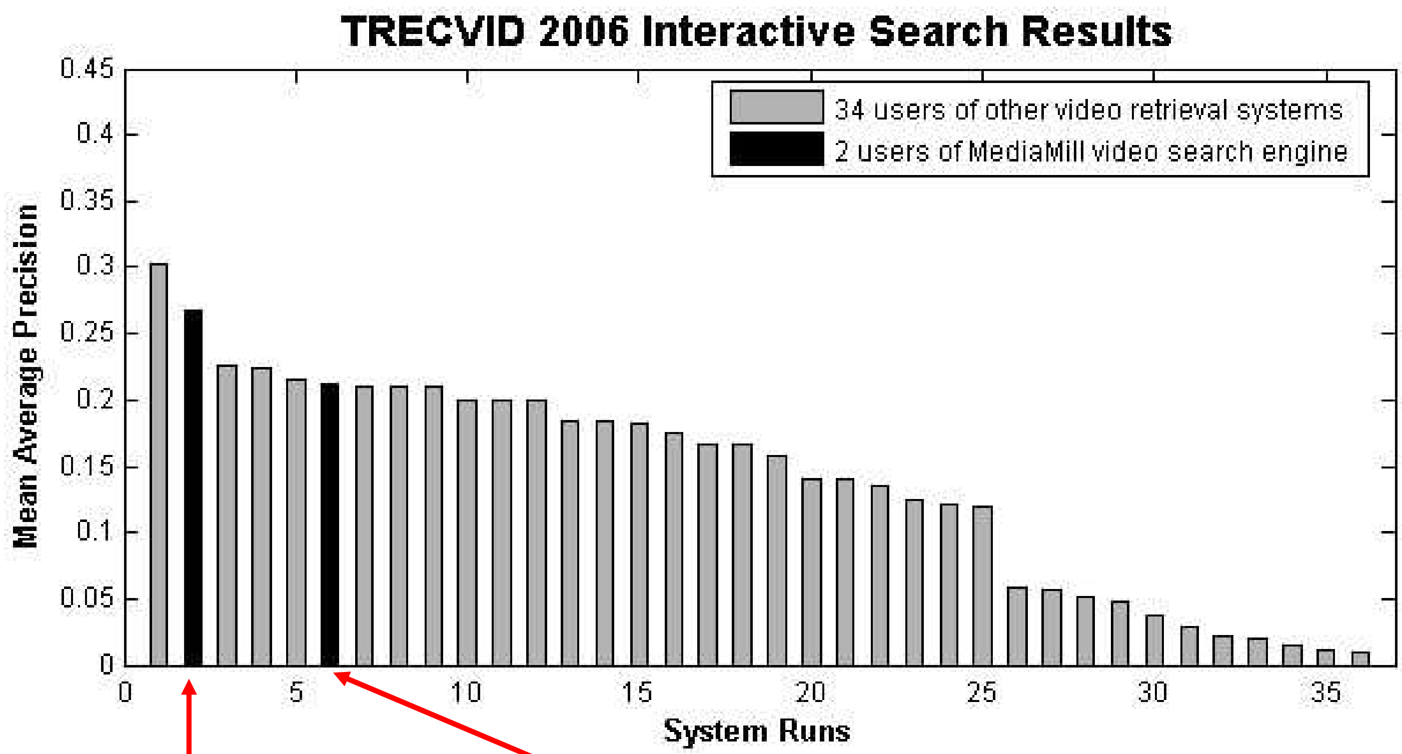
- Time
  - The timeline of the original video
- Keypoint based similarity
  - Requires interaction by the user, cannot be pre-computed

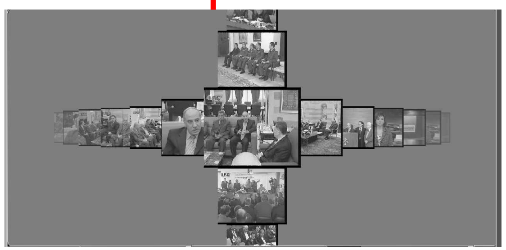
- The concept for which the shot receives a top-rank

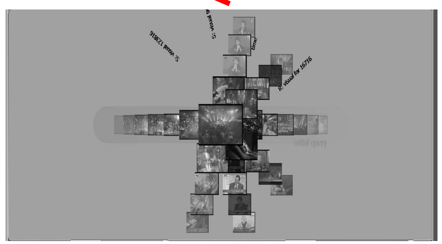
The RotorBrowser



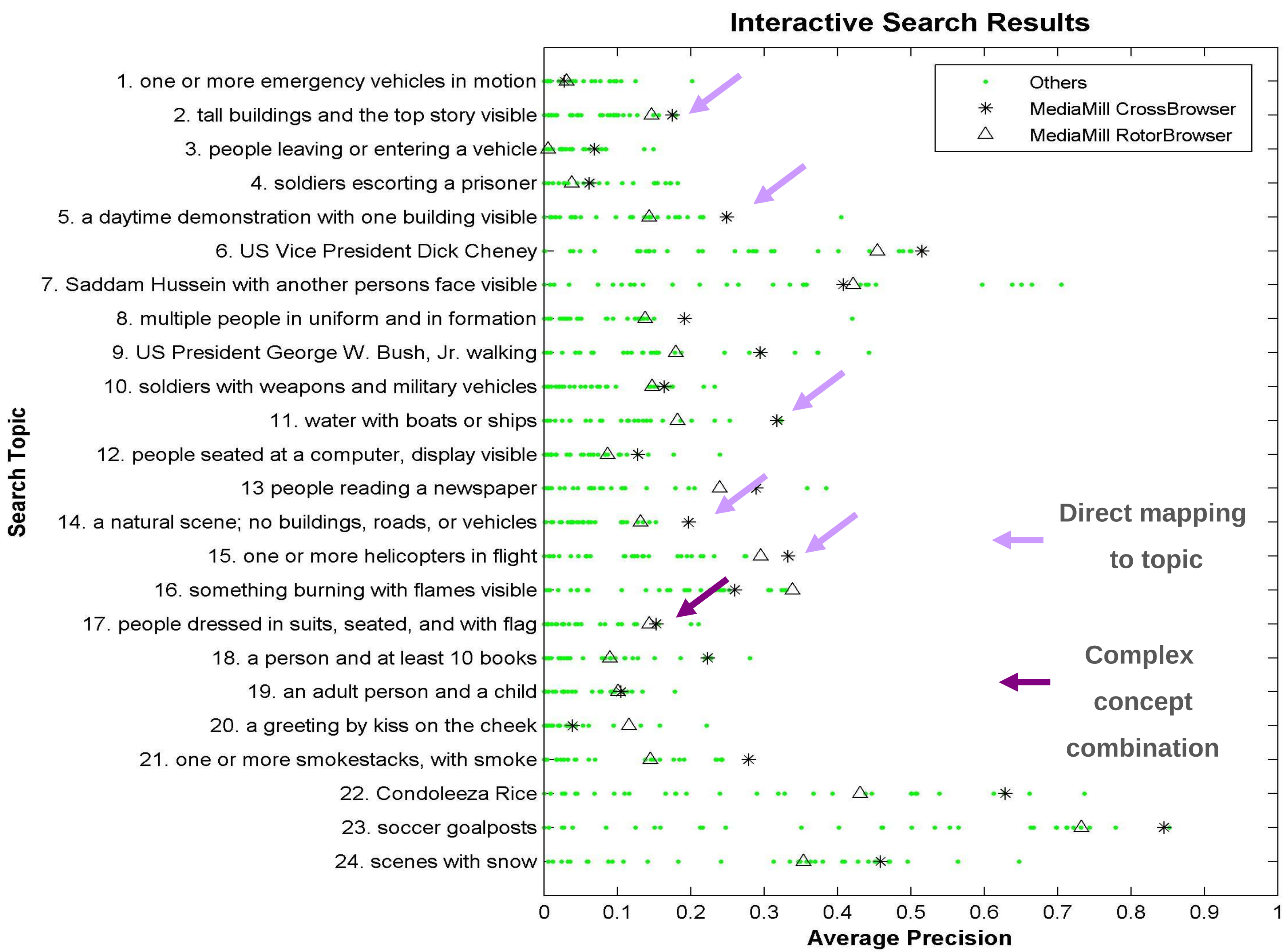
Results


TRECVID 2006 Interactive Search Results
34 users of other video retrieval systems
2 users of MediaMill video search engine
Mean Average Precision
System Runs
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
5
10
15
20
25
30
35





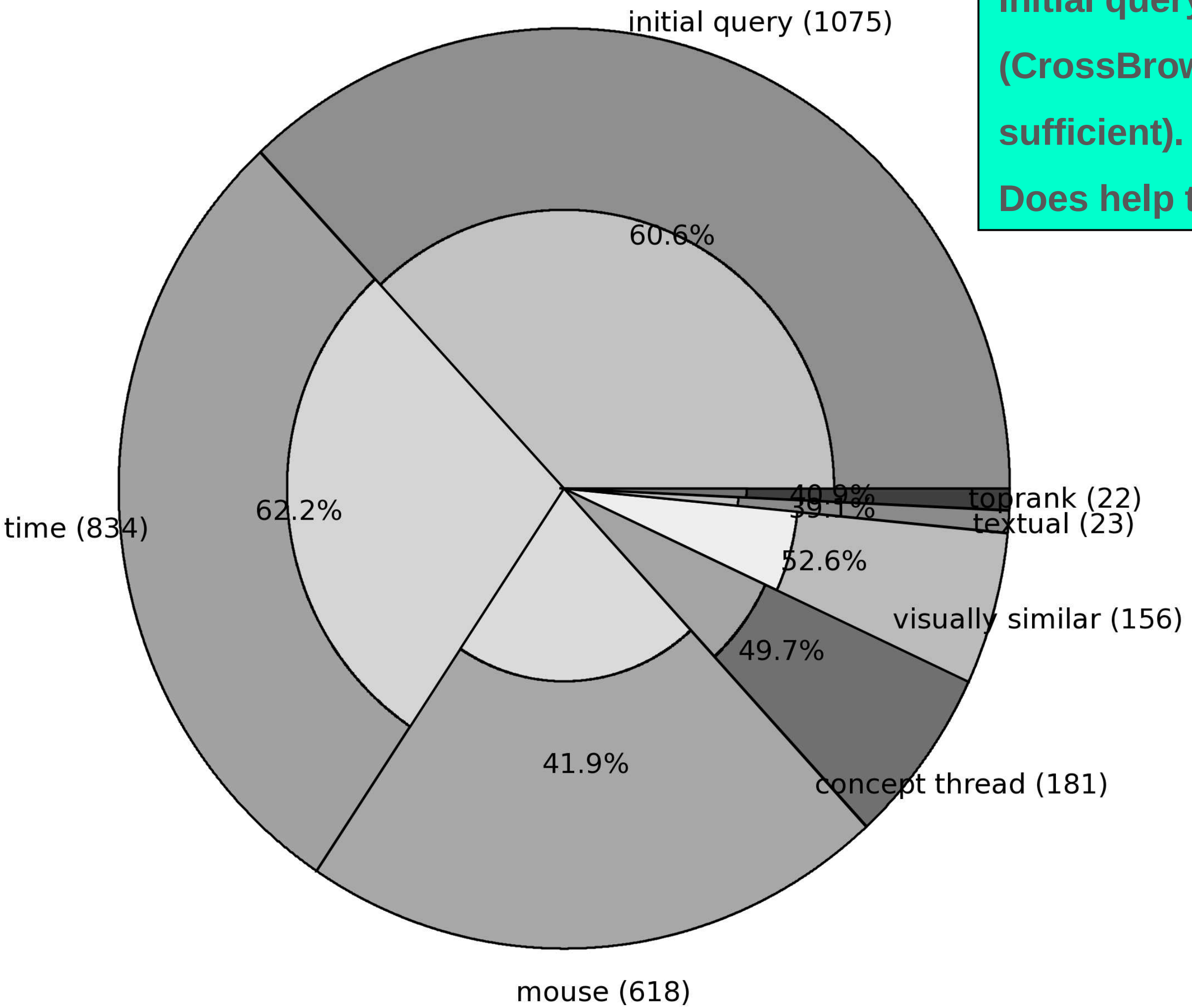
Results: CrossBrowser



Results: RotorBrowser

Our RotorBrowse run

Bookmarks and their judgement by NIST

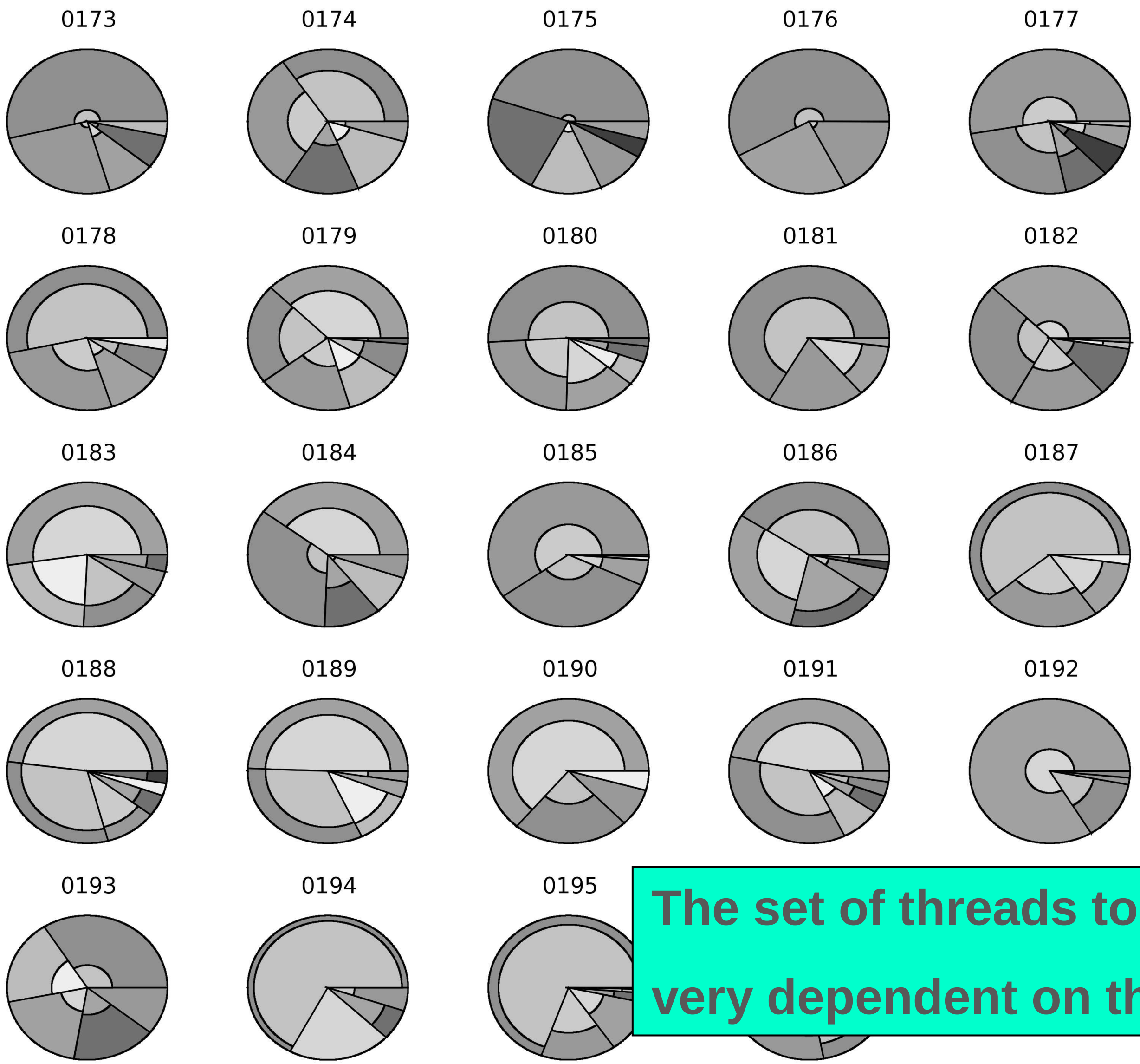


Initial query and time dominant.

(CrossBrowser would have been sufficient).

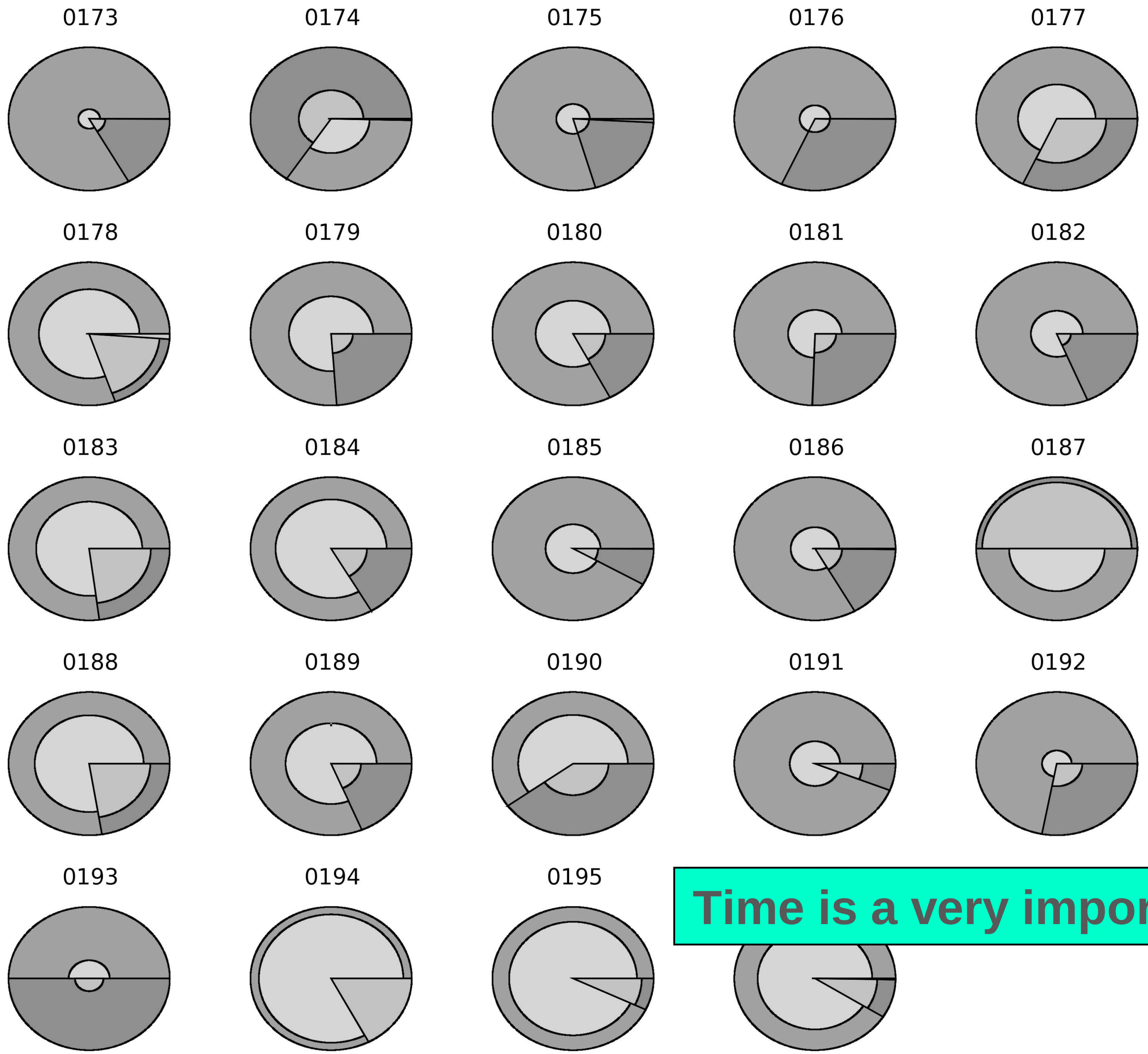
Does help to find additional shots

Decomposition: RotorBrowser



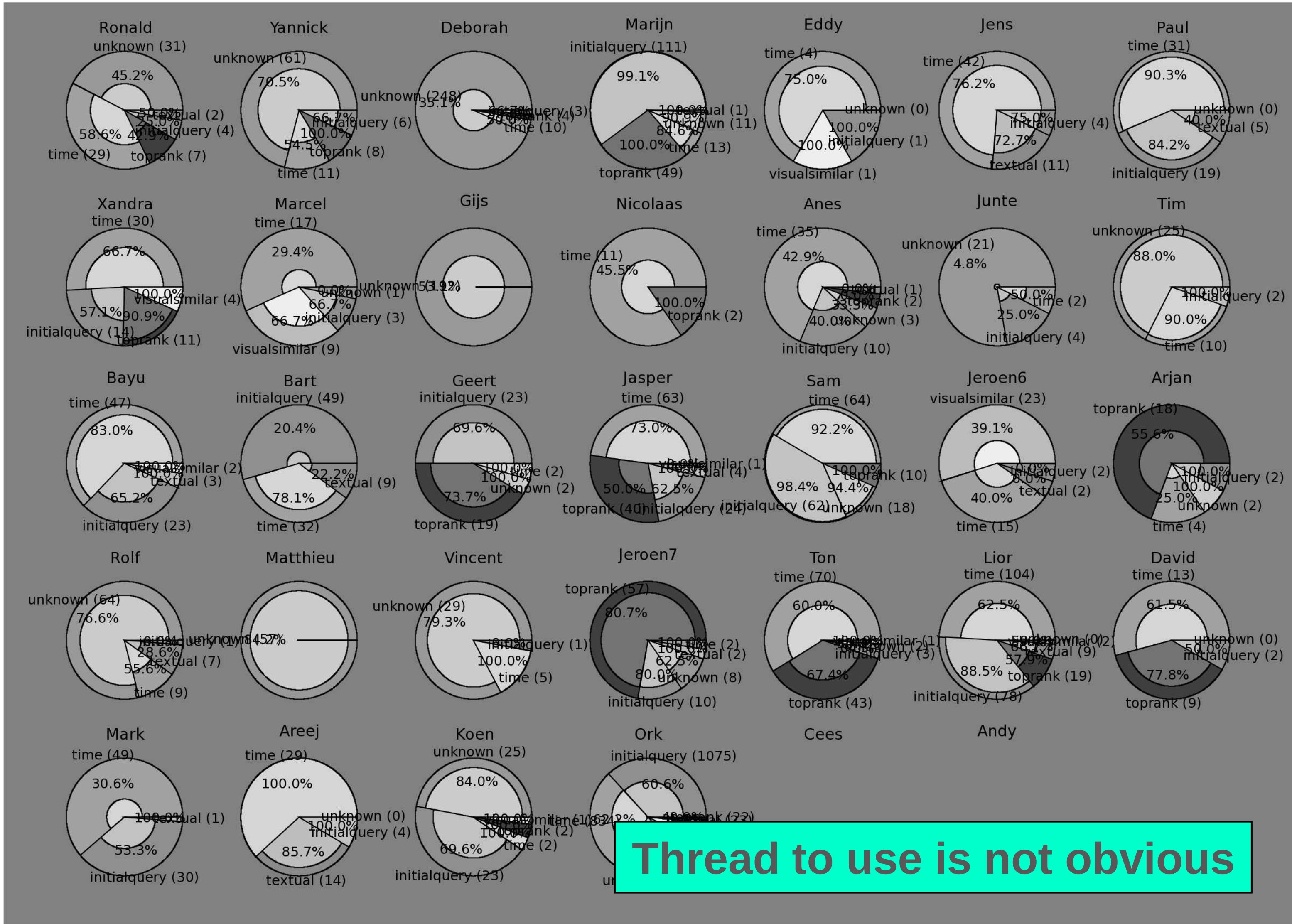
The set of threads to use is very dependent on the topic

Decomposition: CrossBrowser



Time is a very important factor

30 novice users using RotorBrowser



Thread to use is not obvious

Lessons learned

- Region based querying
  - For the current TRECVID topics of limited use
- CrossBrowser versus RotorBrowser
  - For most topics initial query and time are contributing most to the final result, so CrossBrowser often sufficient
  - But in specific cases the use of additional threads can boost performance
- The optimal threads
  - Do not exist, depends on the topic
- Choosing the threads
  - For a novice user not evident from the visualization, performance of threads still too poor

Thanks for your attention