

Vyhledávání informací

ImR a CBVR systémy

1. 12. 2023

Rozehrivací vyhledávačka



Dali mu prostřední jméno Gamaliel a mezi lety 1915-1921 to dotáhl až na pozici amerického senátora (a později možná ještě dál).

Rozehrivací vyhledávačka



Dali mu prostřední jméno Gamaliel a mezi lety 1915-1921 to dotáhl až na pozici amerického senátora (a později možná ještě dál).

Najděte obrázek tohoto muže
v co největším rozlišení/kvalitě.

Opáčko...

- kolekce obrazů rostou velkým tempem
- nutnost v nich vyhledávat roste s nimi
- celá oblast hnána snahou o automatizaci
- automaticky rozpoznat co/kdo je na obrazu
- mnoho rovin využití...
- *image retrieval (ImR)*

Opakování: Jak indexovat obraz?

- nutné přidat metadata
- interní/externí
- *EXIF*
- *alt*

EXIF

<http://exif.regex.info/exif.cgi>

exchangeable image file format

specifikace metadat pro uložení do obrazu

1. Ruční anotace

- ruční tvorba různých metadat
- archivy, muzea, knihovny, galerie
- komerční subjekty, které potřebují přesnost
- většinou kombinují ruční a automatizované přístupy

Title

Description

Object type

Collection

Catalogue title

 ▼

Material

Technique

Objectnumber

Only with image

Now on display

Made between

Inscription / marks

Catalogue reference

Acquisition credits

Provenance name

IMAGE RESOLUTION

PEOPLE

NUMBER OF PEOPLE

No people

One person

Two people

Group of peo...

AGE

PEOPLE COMPOSITION

ETHNICITY

Black

Caucasian

East Asian

Hispanic/Lat...

Middle Easte...

Mixed Race ...

Multi-Ethnic ...

Native Amer...

Native American/First Nations

Pacific Island...

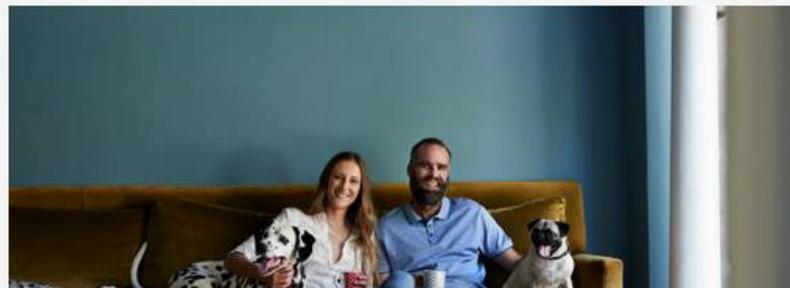
South Asian

Southeast As...

IMAGE STYLE

COLOR

LOCATIONS

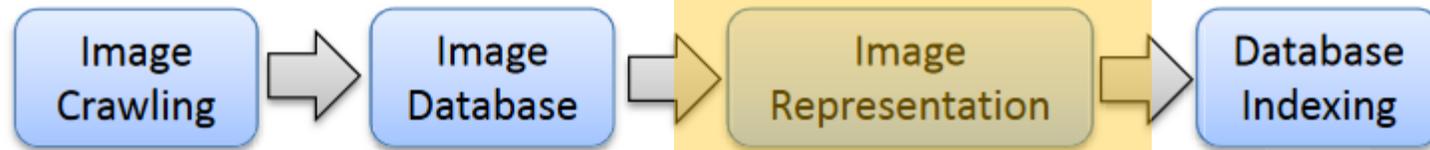


2. Automatická anotace

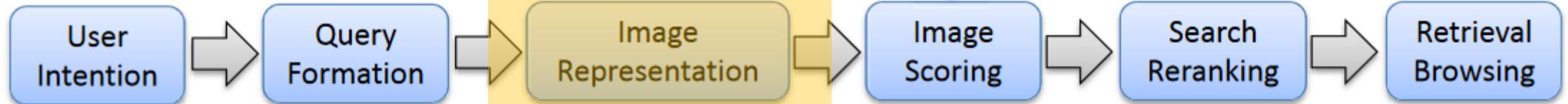
- různé úrovně složitosti
- *content-based ImR* – vyhledávání založené na obsahu
- jiný přístup než TBIImR (*text-based*)
- IBM QBIC – [původní článek](#) (1993)
- automatické rozpoznávání obrazu
- počítačové vidění

In the QBIC (Query By Image Content) project we are studying methods to query large on-line image databases using the images' content as the basis of the queries. Examples of the content we use include color, texture, and shape of image objects and regions. Potential applications include medical ("Give me other images that contain a tumor with a texture like this one"), photo-journalism ("Give me images that have blue at the top and red at the bottom"), and many others in art, fashion, cataloging, retailing, and industry. Key issues include derivation and computation of attributes of images and objects that provide useful query functionality, retrieval methods based on similarity as opposed to exact match, query by image example or user drawn image, the user interfaces, query refinement and navigation, high dimensional database indexing, and automatic and semi-automatic database population. We currently have a prototype system written in X/Motif and C running on an RS/6000 that allows a variety of queries, and a test database of over 1000 images and 1000 objects populated from commercially available photo clip art images. In this paper we present the main algorithms for color texture, shape and sketch query that we use, show example query results, and discuss future directions.

Offline Stage



Online Stage



Content-based ImR

- *content-based image indexing* – indexování
- indexování mimo textové charakteristiky
- prvním krokem je identifikace těchto charakteristik
- poté hledání způsobů, jak je zachytit/popsat

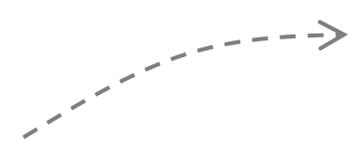
Co může charakterizovat
obrázek kromě *txt* metadat?



TBImR

Camera:	Canon EOS 400D Digital
Exposure:	0.005 sec (1/200)
Aperture:	f/11
Focal Length:	100 mm
ISO Speed:	100
Exposure Bias:	0/3 EV
Flash:	Flash fired
Orientation:	Horizontal (normal)
X-Resolution:	72 dpi
Y-Resolution:	72 dpi
Software:	Adobe Photoshop CS2 Windows
Date and Time:	2008:01:06 20:59:54
YCbCr Positioning:	Co-Sited
Exposure Program:	Manual
Date and Time (Original):	2008:01:06 14:47:06
Date and Time (Digitized):	2008:01:06 14:47:06
Shutter Speed:	500948/65536
Metering Mode:	Pattern
Color Space:	sRGB
Focal Plane X-Resolution:	4433.295 dpi
Focal Plane Y-Resolution:	4453.608 dpi
Exposure Mode:	Manual
Compression:	JPEG

CBImR



např. barva, textura, tvar...
získávání příznaků

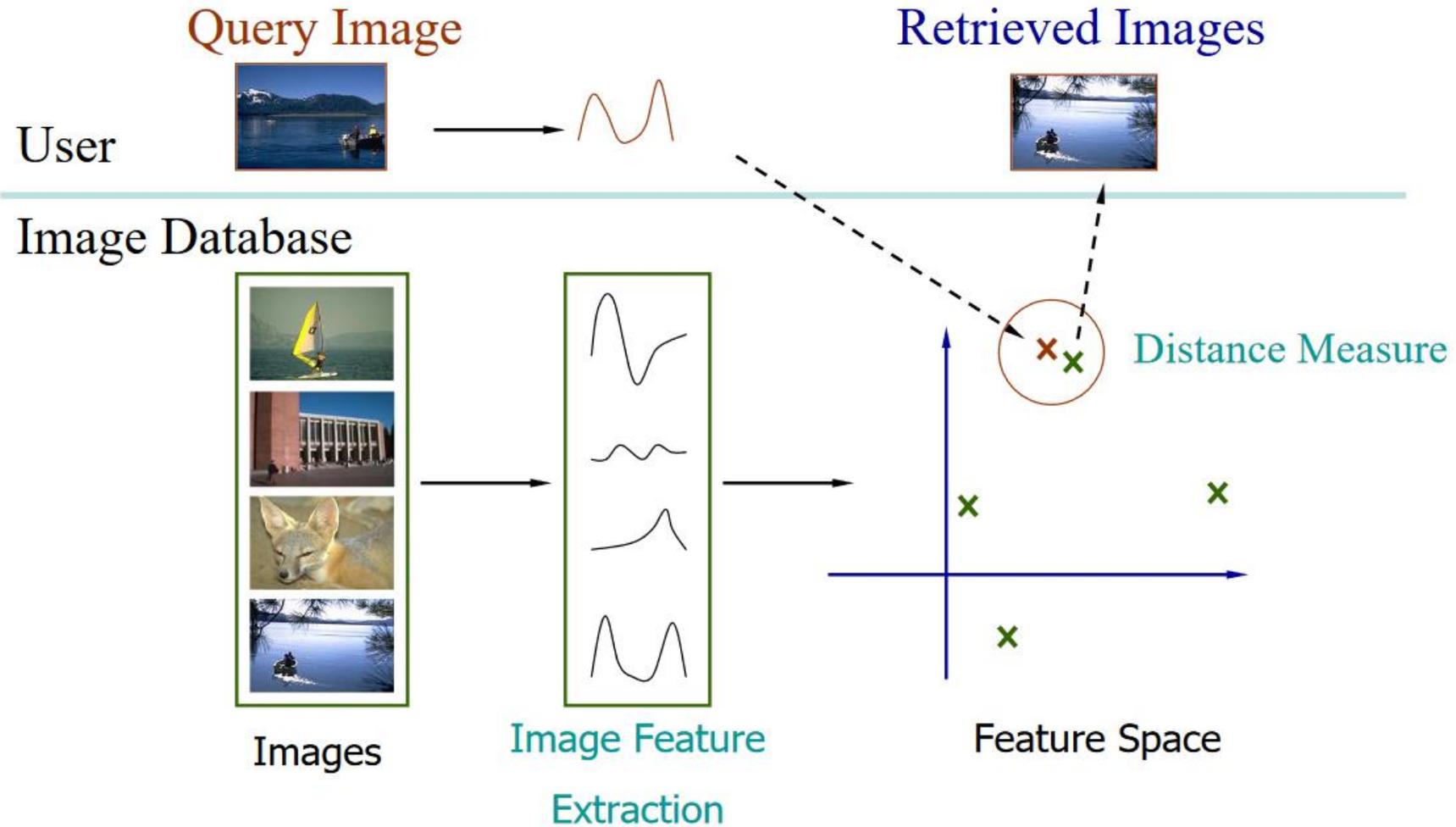


databáze obrázků



query





barevný histogram

Table 1: Grouping of the features into different types. (a) color representation, (b) texture representation, (c) local features, (d) shape representation. The table also gives the time to extract the features from 10 images and to query 10 images in a 10 image database to give an impression of the computational costs of the different features (experiments were performed on a 1.8GHz machine).

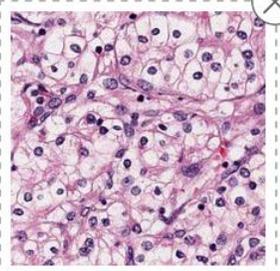
Feature name	Section	comp. measure	type	extr.[s]	retr.[s]
Appearance-based Image Features					
32×32 image	3.1	Euclidean	abcd	0.25	0.19
X×32 image	3.1	IDM	abcd	0.25	9.72
Color Histograms	3.2	JSD	a	0.77	0.16
Tamura Features	3.3	JSD	b	14.24	0.13
Global Texture Descriptor	3.4	Euclidean	b	3.51	0.16
Gabor histogram	3.5	JSD	b	8.01	0.12
Gabor vector	3.5	Euclidean	b	8.68	0.17
Invariant Feature Histograms					
w. monomial kernel	3.6	JSD	ab	28.93	0.16
w. relational kernel	3.6	JSD	ab	18.23	0.14
LF Patches					
global search	3.7	-	ac	4.69	7.13
histograms	3.7	JSD	ac	4.69+5.17	0.27
signatures	3.7	EMD	ac	4.69+3.37	0.55
LF SIFT					
global search	3.7	-	cd	11.91	9.23
histograms	3.7	JSD	cd	11.91+6.23	0.27
signatures	3.7	EMD	cd	11.91+4.50	1.03
MPEG 7: scalable color	3.8.1	MPEG7-internal	a	0.48	0.42
MPEG 7: color layout	3.8.2	MPEG7-internal	ad	0.20	0.33
MPEG 7: edge histogram	3.8.3	MPEG7-internal	b	0.16	0.43

https://en.wikipedia.org/wiki/List_of_CBIR_engines

Luigi Luigi-Genome

Search Image

Search with demo images

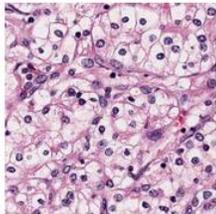
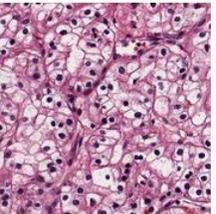
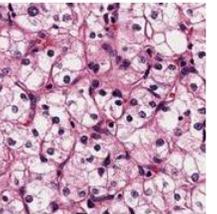
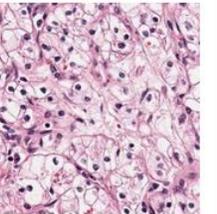
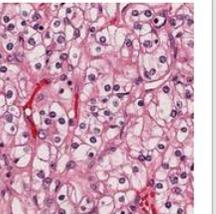
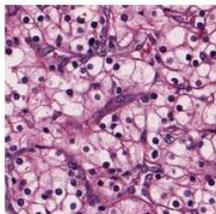
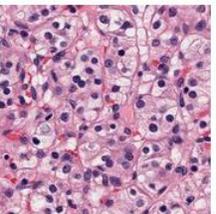
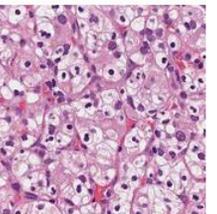
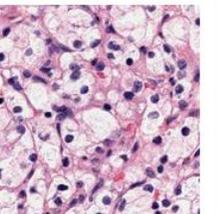
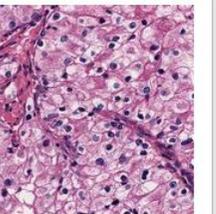


Filename: 05.jpg

Search

Results of image retrieval

Select a body location/system

 <p>Kidney renal clear cell ca... Similarity: 0.999</p> <p>TCGA-B0-4814-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.961</p> <p>TCGA-B0-4821-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.954</p> <p>TCGA-B0-4852-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.952</p> <p>TCGA-CZ-4863-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.951</p> <p>TCGA-B0-5106-01Z-00-DX1</p>
 <p>Kidney renal clear cell ca... Similarity: 0.950</p> <p>TCGA-BP-4327-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.950</p> <p>TCGA-EU-5907-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.949</p> <p>TCGA-A3-3362-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.948</p> <p>TCGA-EU-5906-01Z-00-DX1</p>	 <p>Kidney renal clear cell ca... Similarity: 0.947</p> <p>TCGA-B4-5838-01Z-00-DX1</p>



Color-based ImR

- analýza na nejnižší úrovni
- *barevné momenty* (ze statistických centrálních m.)
- *barevné histogramy, color signatures,...*
- nenesou žádnou informaci o prostorovém rozložení
- stane se něco, když obrázek otočím?

- proto se využívají tzv. *spatial color features*
- např. *barevné korelogramy*, *vektory koherence barev,...*

barevné momenty



Index Image



Test Image 1



Test Image 2

$$\begin{bmatrix} 0.1016 & 0.1149 & 0.1779 \\ 0.8583 & 0.1139 & 0.0563 \\ 0.6416 & 0.2994 & 0.0974 \end{bmatrix}$$

Index Image

$$\begin{bmatrix} 0.1718 & 0.0986 & 0.1400 \\ 0.7619 & 0.1508 & 0.0455 \\ 0.7062 & 0.2242 & 0.0772 \end{bmatrix}$$

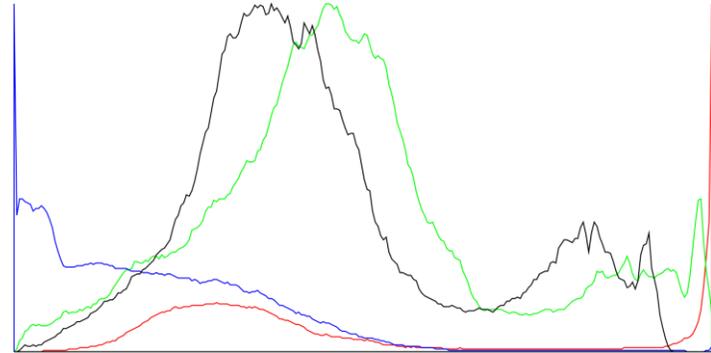
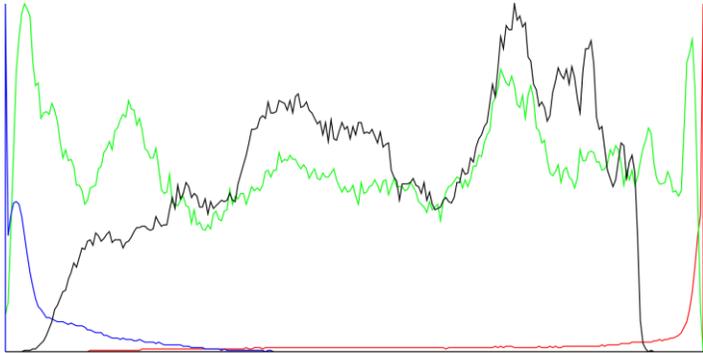
Test Image 1

$$\begin{bmatrix} 0.1878 & 0.1671 & 0.2331 \\ 0.2462 & 0.2281 & 0.2492 \\ 0.6052 & 0.3532 & 0.1534 \end{bmatrix}$$

Test Image 2

$$d_{mom}(Index, Test1) < d_{mom}(Index, Test2)$$

barevný histogram vs. korelogram



vektor koherence barev



Histogram rank: 50. CCV rank: 26.



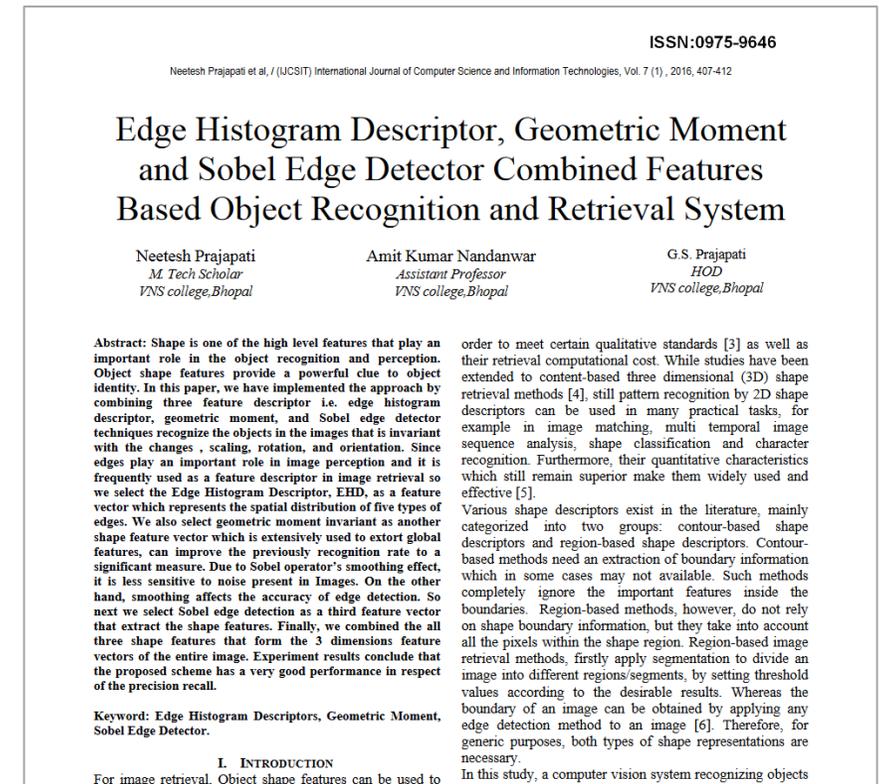
Histogram rank: 35. CCV rank: 9.

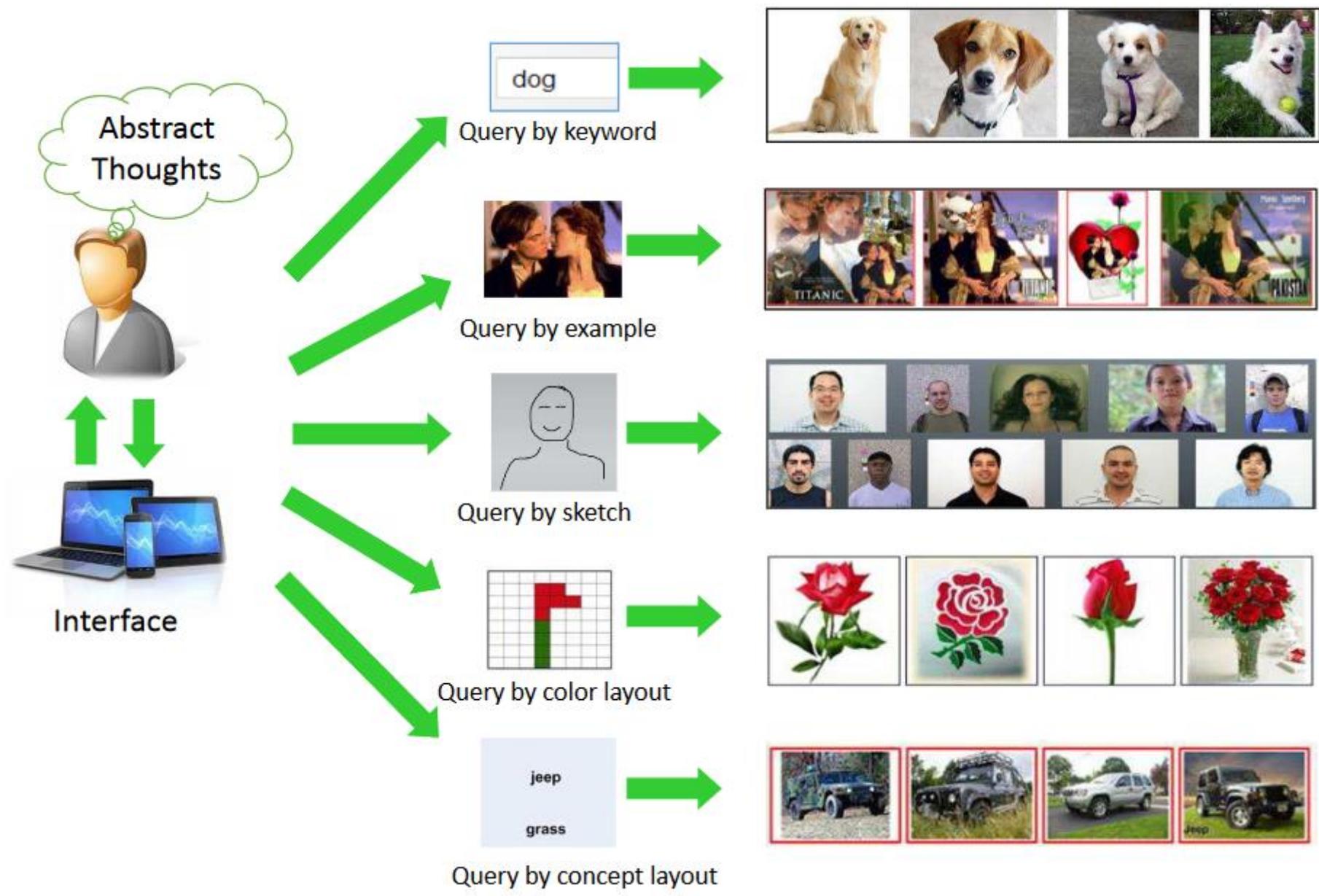
Indexace podle textury

- lokální binární vzor
- co-occurrence matrix

Indexace podle tvaru

- edge histogram
- image moments

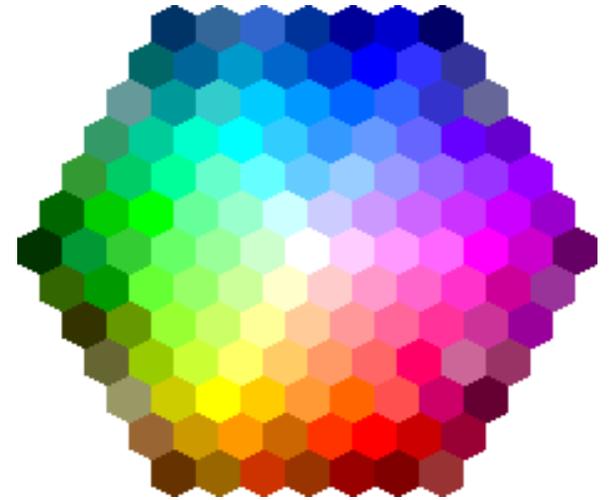




Query-by-color

- vyhledávání pomocí barvy
- implementováno v různých systémech
- využívá se často u digitalizovaných sbírek
- používá např. Europeana

- [TinEye Multicolor Engine](#)
- [MultiColr](#)
- [Google ArtPalette](#)



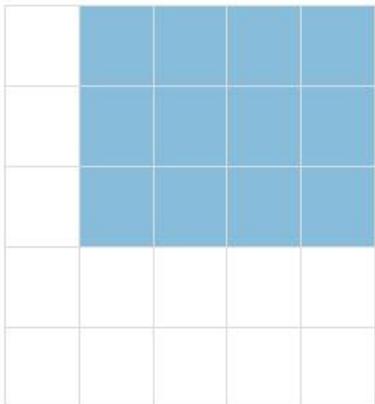
QBE | Query-by-example

- dotaz příkladem, reverzní vyhledávání
- *query = obrázek*
- Google Image Search, TinEye
- [Google Lens](#)

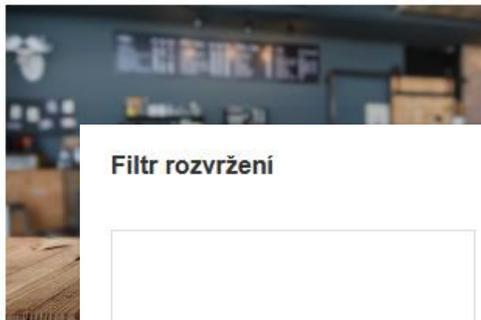
Layout filter

Reset

Copy Space



Click and drag within the grid to select where within the image copy space is needed.



Filtr rozvržení



Vyhledávání kompozice



Zahrnout místo pro text

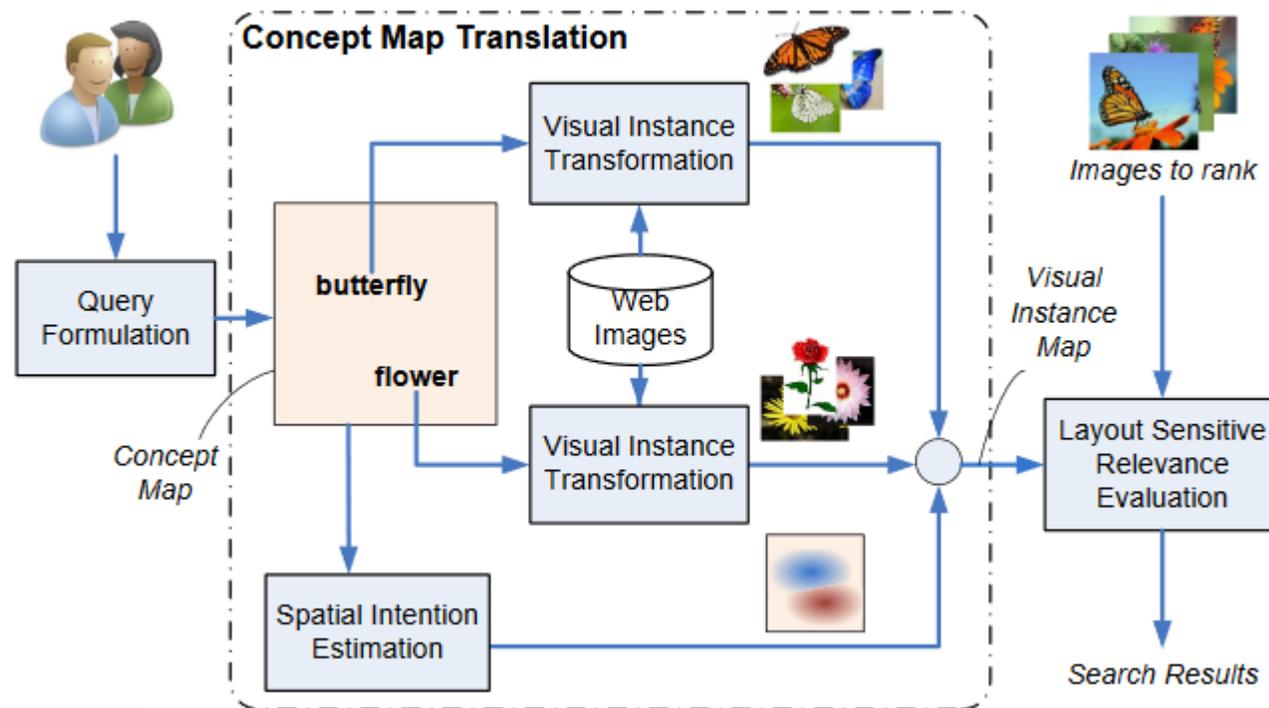
Vyhledávejte a klíčová slova se zobrazí v poli výše, připravena k přesunutí na místo.

Proč je tato stránka částečně v angličtině?



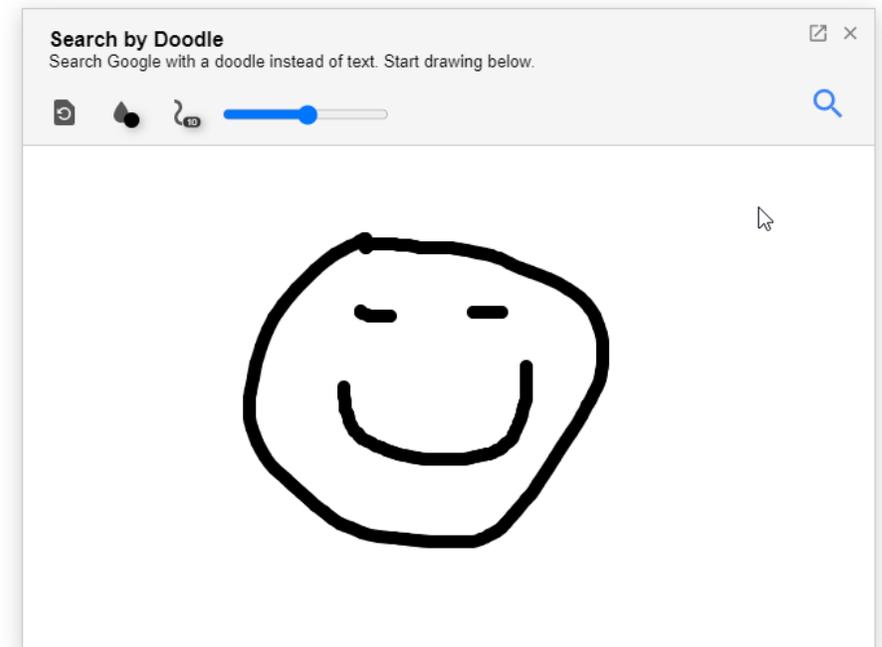
Query-by-concept-map

- Microsoft
- *concept layout*
- záleží i na rozložení

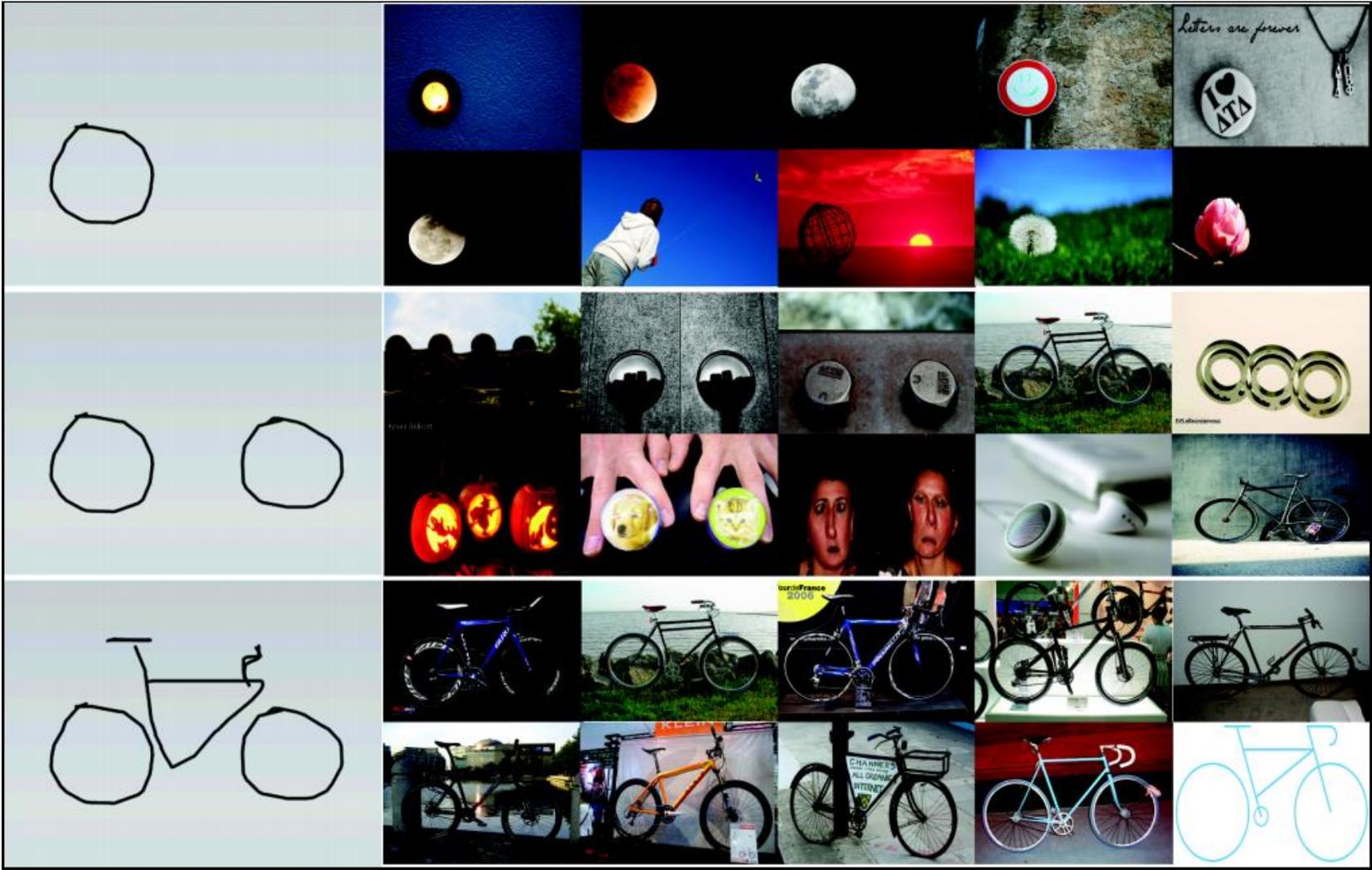


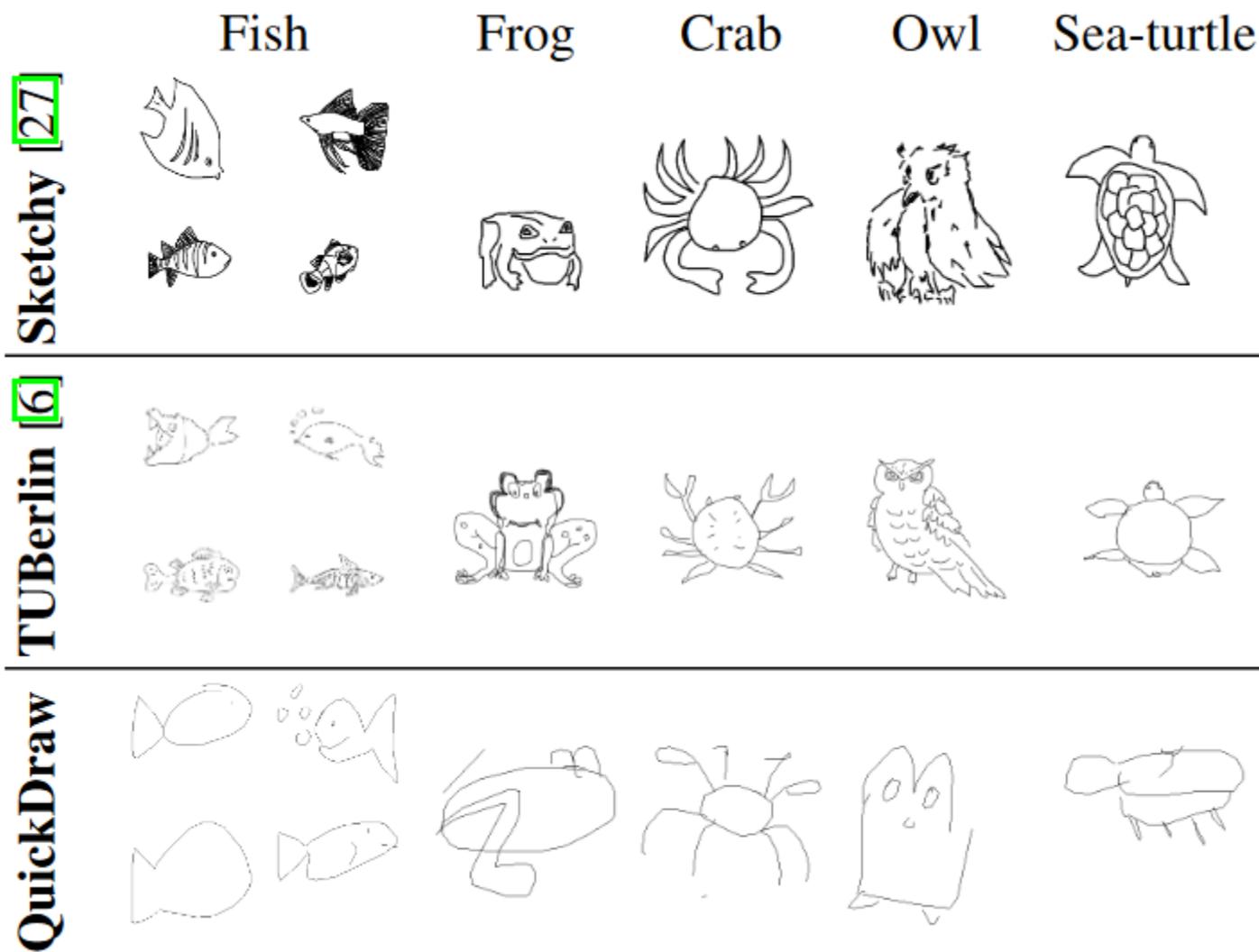
SBIR | Sketch-based ImR

[Google QuickDraw](#)



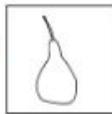
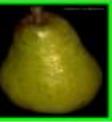
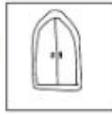
MS Mind Finder

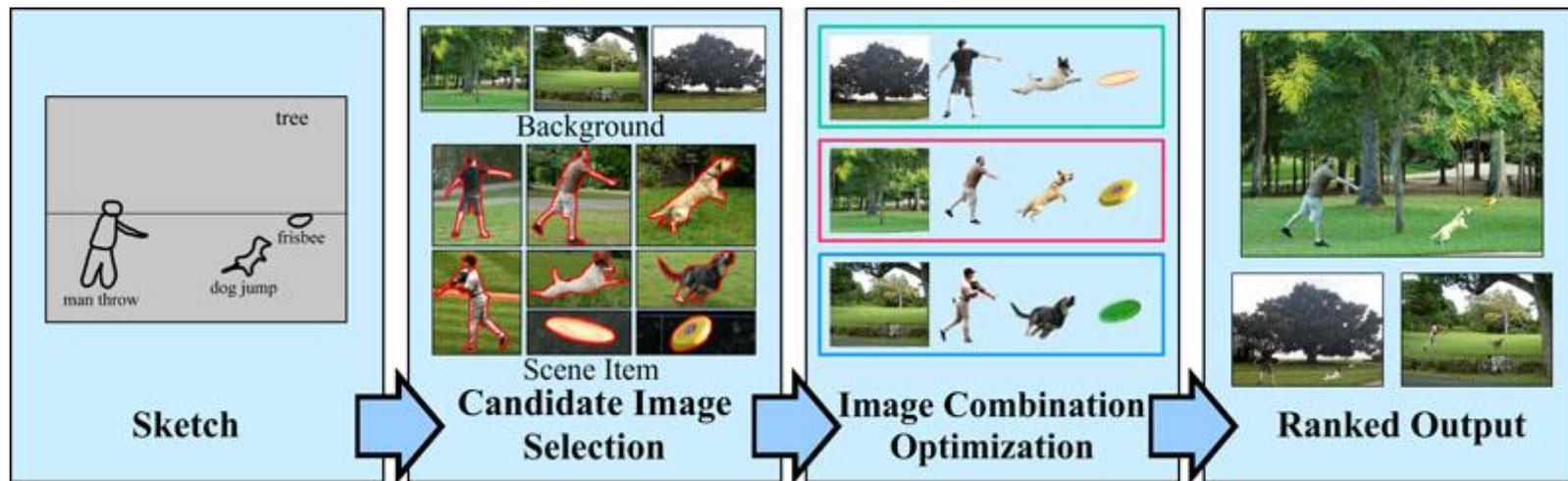




Sketchy [27]

QuickDraw

	Query	Top-8 retrieved candidates									Query	Top-8 retrieved candidates							
CVAE [36]																			
Ours	pear									skyscraper									
CVAE [36]																			
Ours	door									helicopter									



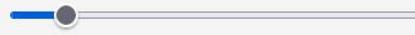
Fill/brush color:



Brush shape:



Brush size: 24



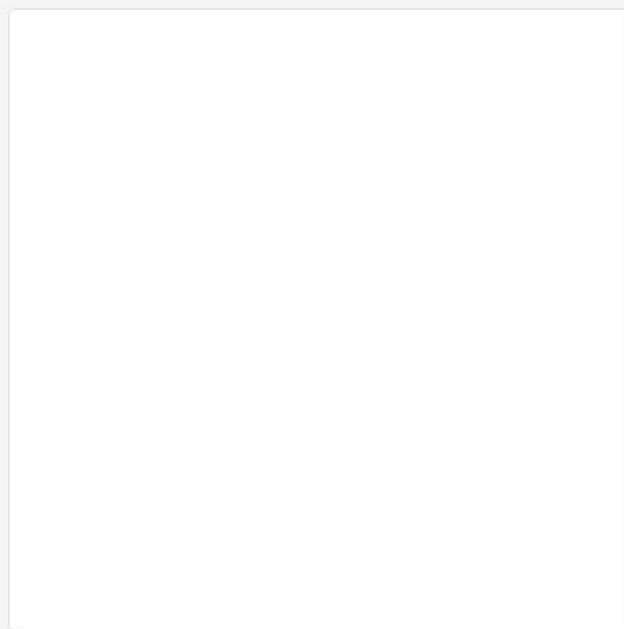
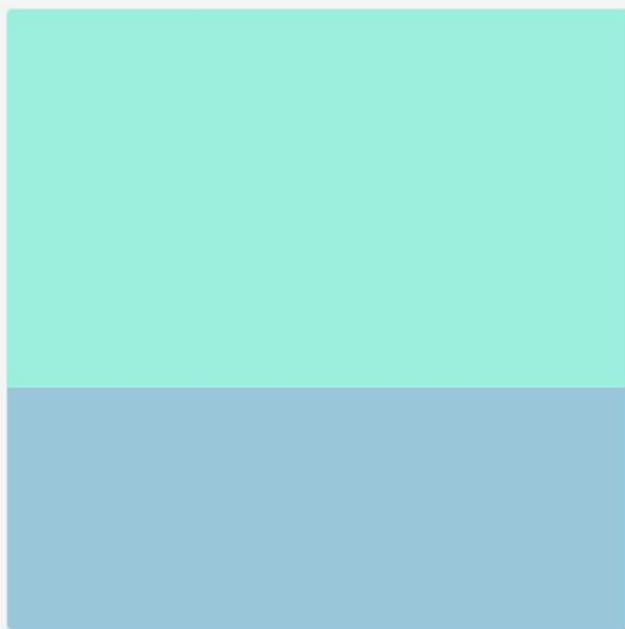
GauGAN Beta

Building

Ground

Landscape

Plant



Upload Segmentation Map ⓘ

Browse...

No file selected.

Upload

Upload Landscape Image ⓘ

Browse...

No file selected.

Upload

Upload Custom Style Filter ⓘ

Browse...

No file selected.

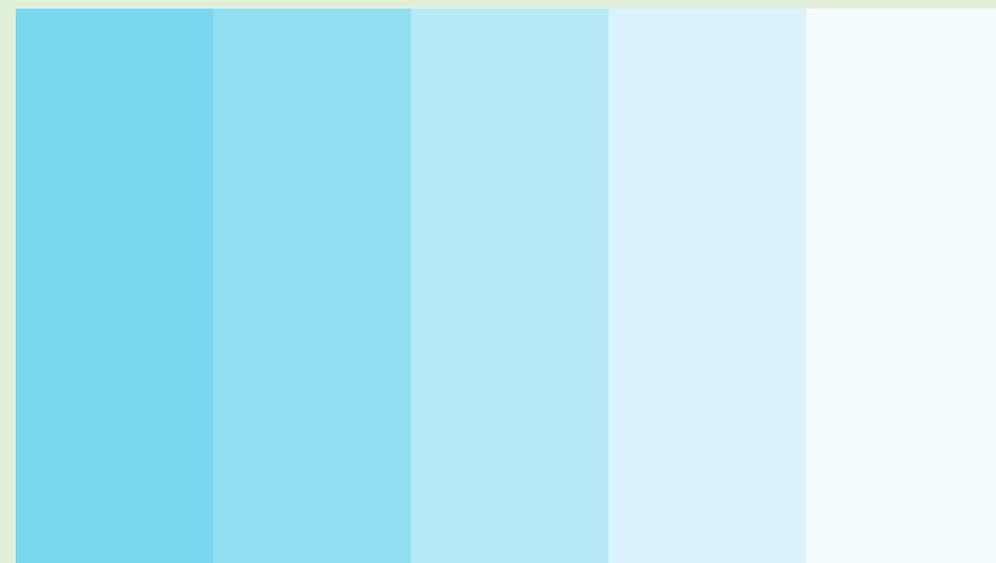
Upload

---> generativní umění

Vyhledávačka



Najděte obrázky z digitalizovaných sbírek švédských paměťových institucí jejichž dominantní barvou je modrá.



Vyhledávačka

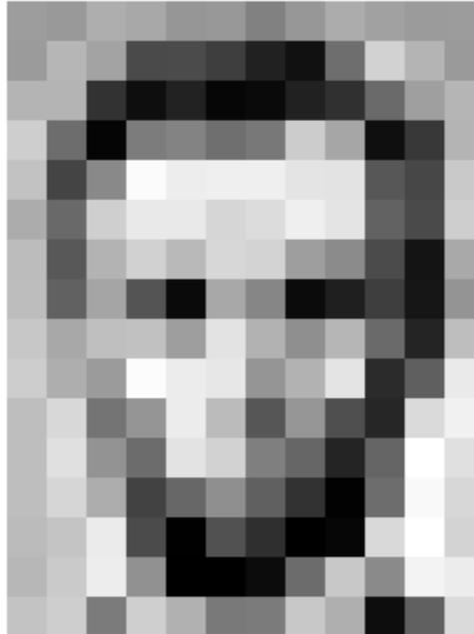


Najděte obrázky z digitalizovaných sbírek švédských paměťových institucí jejichž dominantní barvou je modrá.



Automatická indexace obsahu

- automatické rozpoznání obsahu
- jak naučit počítač rozpoznat, co je na obrázku?
- počítačové vidění



157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	106	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218



man in black shirt is playing guitar.



construction worker in orange safety vest is working on road.



two young girls are playing with lego toy.



boy is doing backflip on wakeboard.



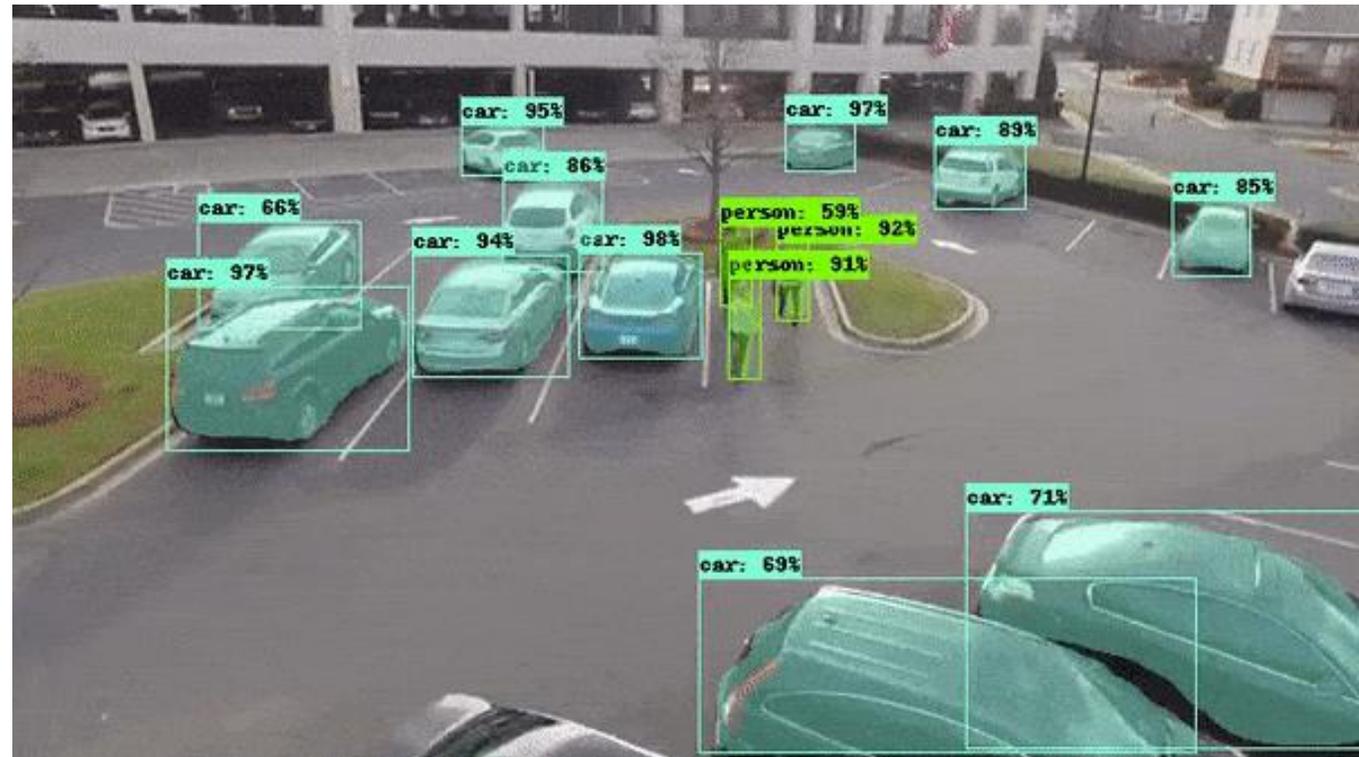
"a young boy is holding a
baseball bat."

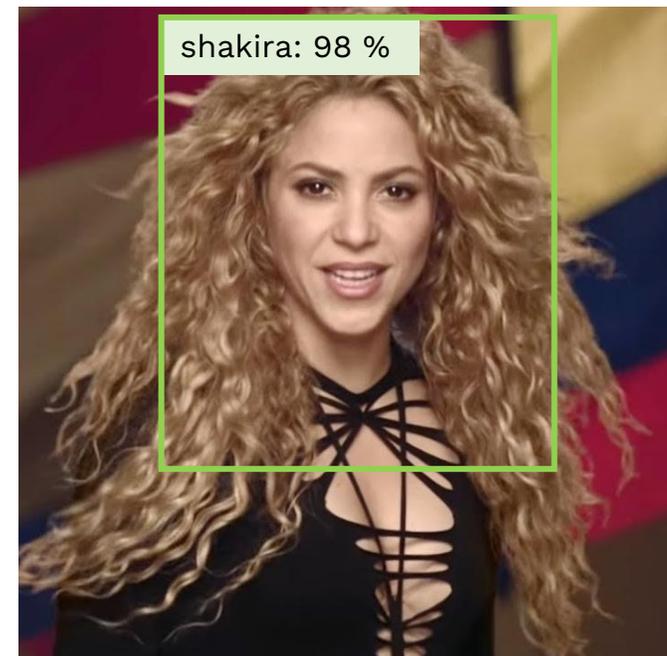
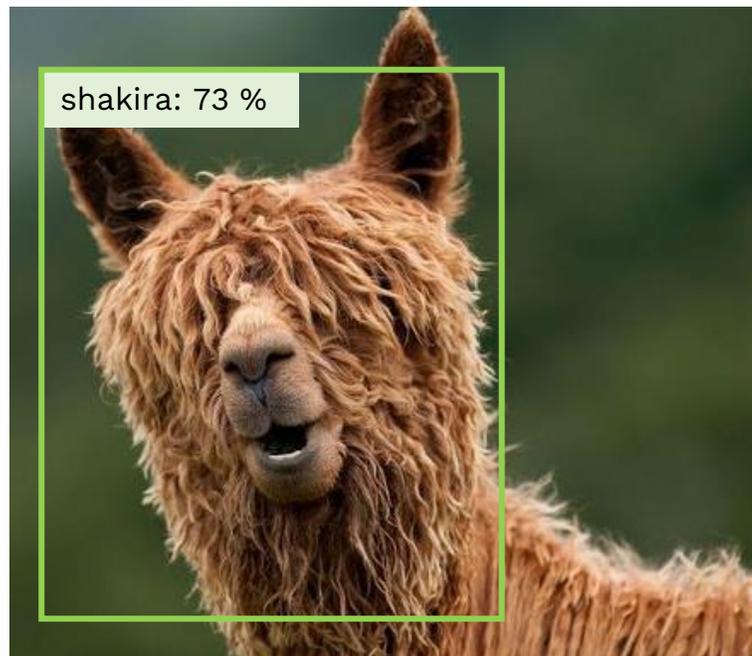
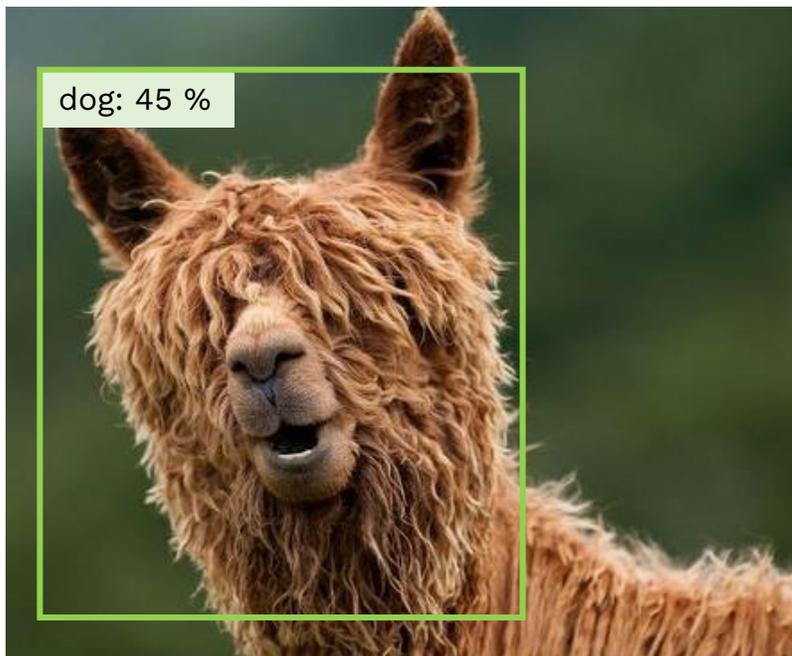


"a cat is sitting on a couch with a
remote control."

Rozpoznávání obrazu

- konvoluční neuronové sítě
- *vstup* – obraz
- *výstup* – třída, s určitou pravděpodobností
- všude, od taggování na FB po autonomní auta
- *supervised learning*
- *unsupervised learning*





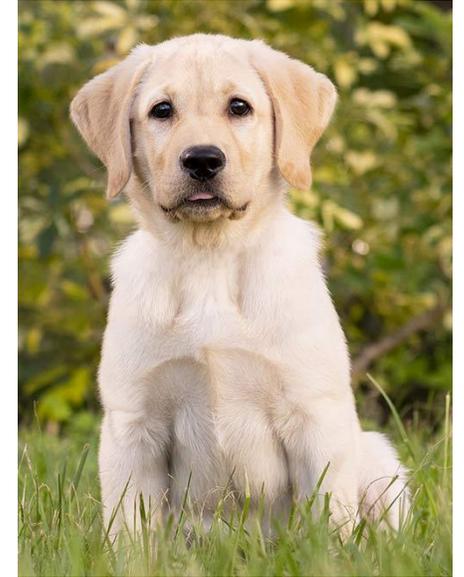


na, tady máš obrázky...

group_A



group_B





A PELICAN
BOOK

Artificial Intelligence

A Guide for Thinking Humans

Melanie Mitchell

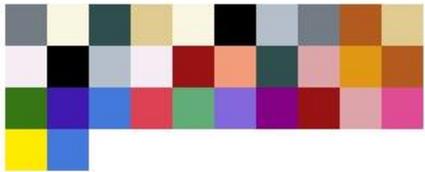
8	5	6	5	0	21	24	11	20	23	7	0	4	5	4	7
0	0	0	0	20	0	28	26	21	24	18	4	0	0	0	0
3	0	0	11	1	125	230	0	119	13	26	9	9	0	0	2
5	0	10	10	95	239	254	229	96	21	28	15	33	0	0	3
0	0	0	8	197	251	243	253	201	58	12	12	11	0	0	0
0	0	7	18	180	183	234	248	176	79	9	8	20	0	0	0
0	0	5	62	216	168	228	235	173	175	126	6	6	0	0	0
0	0	6	145	267	241	232	231	231	237	227	41	5	30	0	0
0	0	9	143	253	242	221	225	247	246	222	186	12	17	0	0
0	0	9	40	226	225	222	222	235	225	194	160	12	14	0	0
0	0	16	6	196	230	223	217	224	228	119	6	15	25	0	0
0	0	22	6	94	244	232	232	231	228	137	6	29	11	0	0
0	0	16	10	9	210	246	238	204	241	138	5	30	6	0	0
0	0	10	11	11	162	229	227	221	250	150	7	8	13	0	0
0	0	7	8	11	170	228	238	238	243	183	159	125	6	0	0
0	0	8	5	8	195	215	225	229	228	231	241	100	4	5	0

– Refine

› Images only (450,318)

- + Maker
- + Object type
- + Period
- + Place
- + Material
- + Technique

Color



Dolls' house of Petronella Oortman

anonymous, c. 1686 - c. 1710

On display in room 2.20



Chest with nine bottles

anonymous, c. 1680 - c. 1700



Seated Cupid



Blue Macaw

Meissener Porzellan Manufaktur, 1731

On display in room 1.4



Weepers from the tomb of Isabella...

Borman workshop (attributed to), Renier van Thienen (I) (attributed to), c. 1475 - c. 1476



Cupboard

Herman Doomer, c. 1635 - c. 1645

On display in room 2.8

› Images only (450,318)

+ Maker

+ Object type

+ Period

+ Place

+ Material

+ Technique

Color



Co pújde zautomatizovat?

+ Material

- › paper (502,648)
- › photographic paper (104,304)
- › cardboard (65,867)
- › ink (40,436)
- › pencil (37,250)
- › chalk (27,007)
- › watercolor (papers) (16,665)
- › linen (material) (15,250)
- › wood (plant material) (14,899)
- › porcelain (8,800)
- › silver (metal) (8,700)

[Load more](#)

+ Technique

- › etching (175,090)
- › engraving (124,116)
- › albumen print (59,542)
- › pen (38,251)
- › letterpress printing (29,574)
- › brush (29,311)
- › gelatin silver print (29,250)

+ Maker

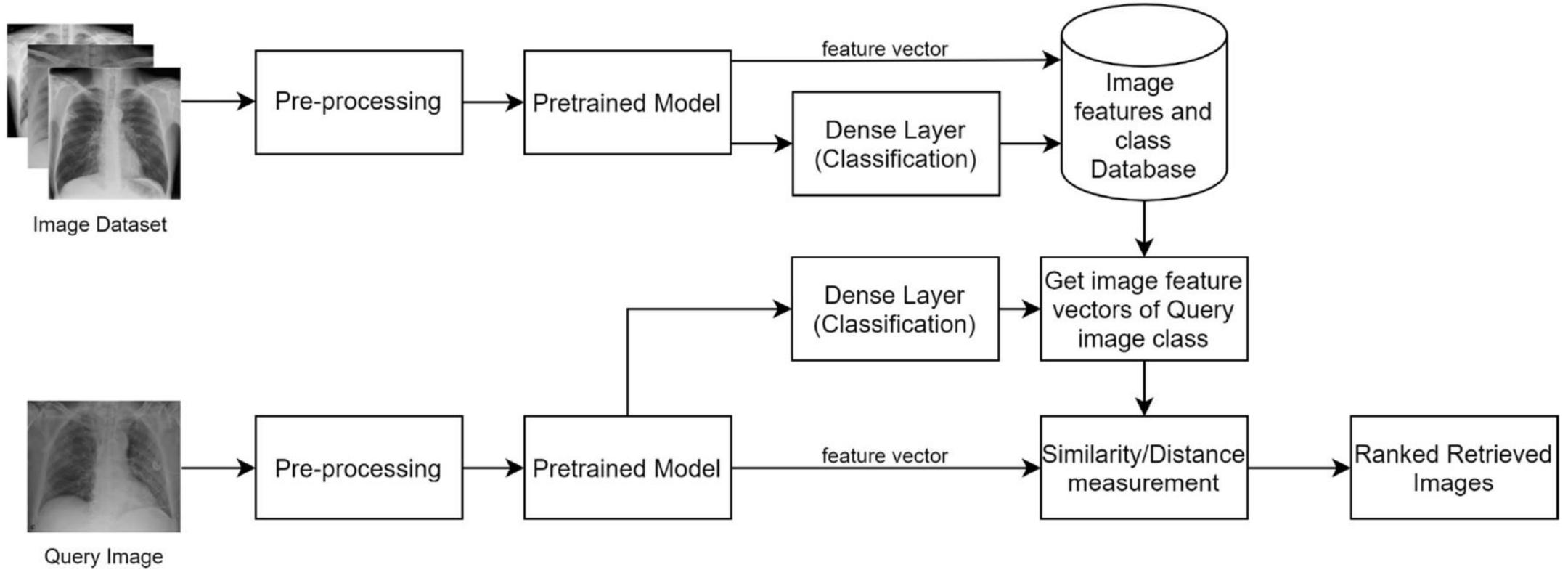
- › anonymous (161,471)
- › George Hendrik Breitner (5,576)
- › Jan Luyken (4,895)
- › Reinier Vinkeles (I) (4,347)
- › Marius Bauer (3,339)
- › Isaac Israels (2,864)
- › unknown (2,859)
- › Johannes Tavenraat (2,649)
- › Daniel Nikolaus Chodowiecki (2,481)
- › Aat Veldhoen (2,430)
- › Bernard Picart (2,351)

[Load more](#)

+ Object type

- › print (329,964)
- › photograph (113,983)
- › drawing (48,099)
- › photomechanical print (29,055)
- › carte-de-visite (16,612)
- › letter (11,140)
- › stereograph (10,709)
- › painting (4,783)
- › cabinet photograph (4,349)
- › book (4,027)
- › history medal (3,783)

[Load more](#)



Vyhledávačka na doma



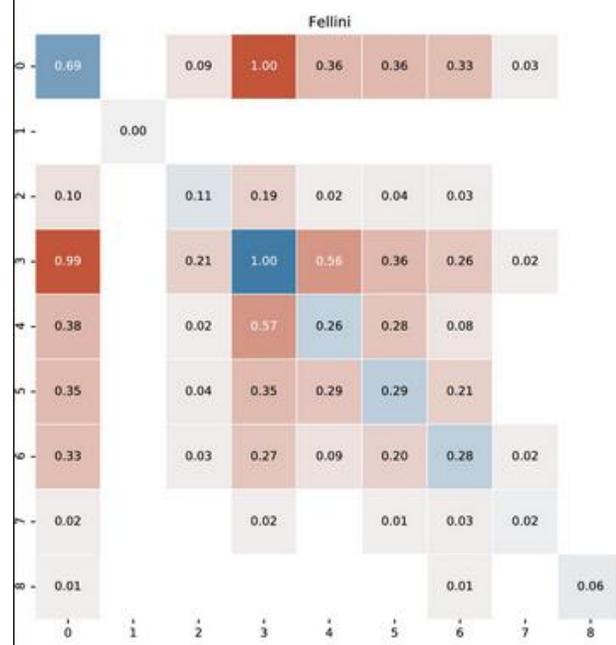
Porovnejte výsledky a funkce jednotlivých vyhledávačů na základě q .

- A. [Google Images](#)
- B. [Yahoo Image Search](#)
- C. [Bing Image Search](#)
- D. [Flickr Search](#)

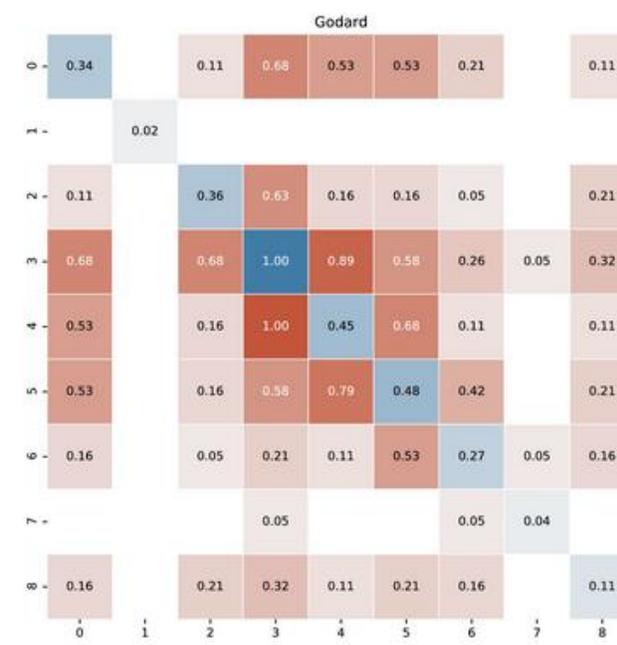
1. Najděte obrázek psa na růžovém pozadí s malou hloubkou ostrosti (tj. pozadí bude rozostřené).
2. Najděte ruční náčrtek/kresbu psa nebo kočky ve veřejné doméně (*public domain*).
3. Najděte kvalitní portrétní fotky vašeho oblíbeného herce/herečky.

Specifika vyhledávání videa

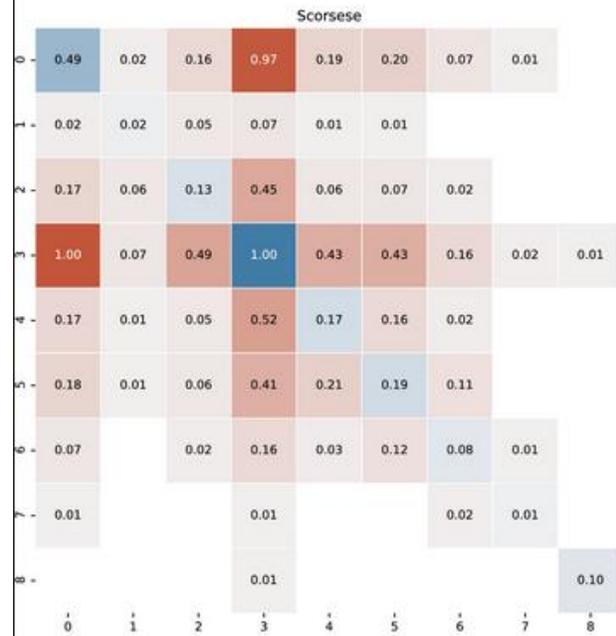
- textová metadata
- speciální sbírky – *televizní archivy, dohledové systémy*
- obrovský nárůst dat
- *content-based video retrieval*
- i zde vede směr automatizace



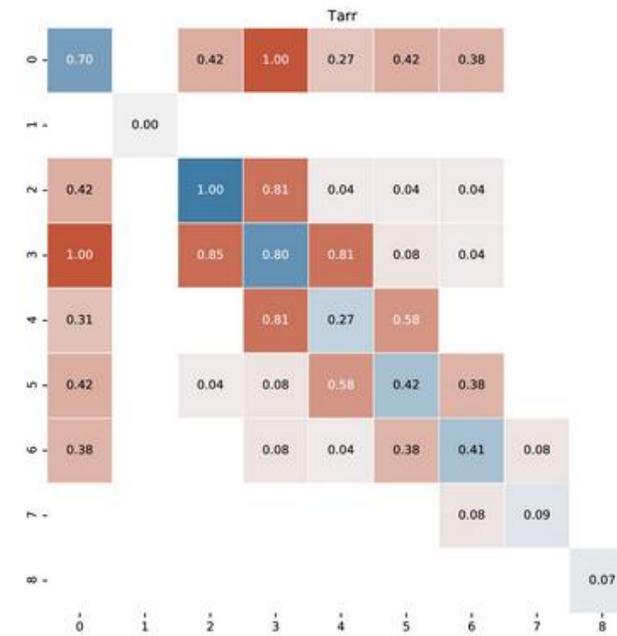
(c) Fellini



(d) Godard



(e) Scorsese



(f) Tarr



netřeba si pamatovat!

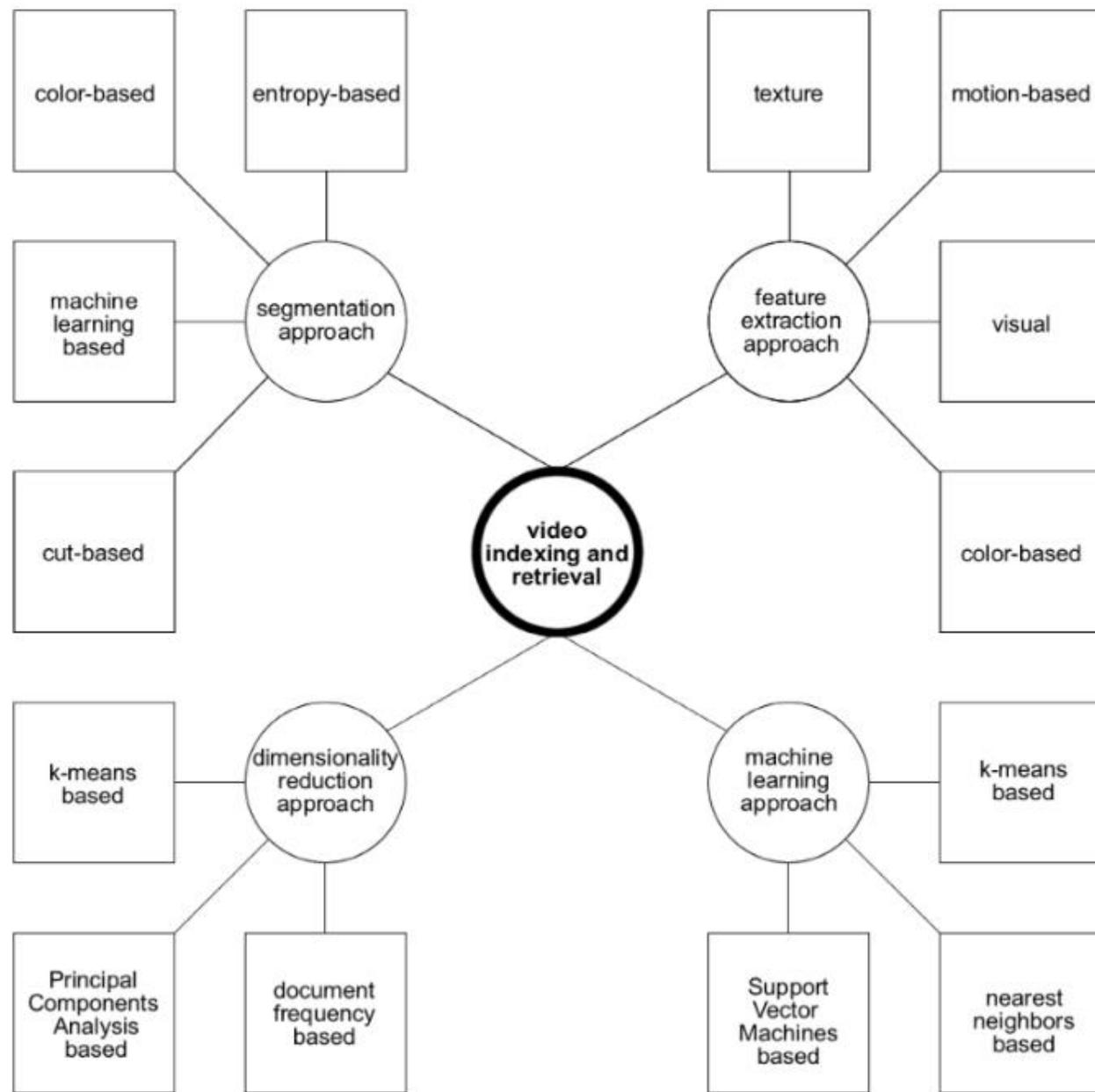


Fig. 3. Usual approaches found in content-based video indexing and retrieval.

Engineering Applications of Artificial Intelligence 90 (2020) 103557

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A systematic review on content-based video retrieval[☆]

Newton Spolaôr^a, Hwei Diana Lee^{a,*}, Weber Shoitly Resende Takaki^{a,b}, Leandro Augusto Ensina^a, Claudio Saddy Rodrigues Coy^b, Feng Chung Wu^{a,b}

^a Laboratory of Bioinformatics, Western Paraná State University, Fox do Iguaçu, Paraná, Presidente Tancredo Neves Avenue, 6731, ZIP code: 85867-900, Brazil

^b Service of Cytogenetics, Faculty of Medical Sciences, University of Campinas, Campinas, São Paulo, Brazil

ARTICLE INFO

Keywords:
Color features
Unsupervised learning
Shot boundary detection

ABSTRACT

Content-based video retrieval and indexing have been associated with intelligent methods in many applications such as education, medicine and agriculture. However, an extensive and replicable review of the recent literature is missing. Moreover, relevant topics that can support video retrieval, such as dimensionality reduction, have not been surveyed. This work designs and conducts a systematic review to find papers able to answer the following research question: "what segmentation, feature extraction, dimensionality reduction and machine learning approaches have been applied for content-based video indexing and retrieval?". By applying a research protocol proposed by us, 153 papers published from 2011 to 2018 were selected. As a result, it was found that strategies for cut-based segmentation, color-based indexing, k-means based dimensionality reduction and data clustering have been the most frequent choices in recent papers. All the information extracted from these papers can be found in a publicly available spreadsheet. This work also indicates additional findings and future research directions.

1. Introduction

Multimedia documents composed of different media types have increasingly been published and consumed (Guo et al., 2015; Benoit-Pineau et al., 2012). This fact is due to the larger access to computational resources and the Internet, among other reasons (Bhaumik et al., 2016). Video in particular consists in a usual way to capture and share information, as it is able to represent moving objects in space and time accordingly. These benefits come at the price of reasonable storage and processing costs (Priya and Shanmugam, 2013).

In general, video content is richer than single image content (Hu et al., 2011). A video file typically has much raw data, but little prior structure. Moreover, information available in video occasionally include textual metadata and captions, images (frames) and audio. Due to the crescent interest in video, automatic indexing and retrieval are usually considered in multimedia research. In particular, the former specifies indexes (features) to describe a video, whereas the latter allows one to search for relevant videos. These tasks can be combined, for example, in find video in an indexed data-base that 2015), education (Yang and Meinel, 2014), geo-referenced video (Yin et al., 2015), human action recognition (Shao et al., 2014), journalism (Younessian and Rajan, 2012), marketing (Sharma et al., 2013), medicine (Charrrière et al., 2014), sports (Al Kabary and Schukit, 2014) and television broadcast (Mühling et al., 2016).

To support video indexing and retrieval, some additional topics can be useful (Puthenpathussery et al., 2017; Priya and Shanmugam, 2013; Hu et al., 2011). Video segmentation is a classical preliminary step typically implemented to separate a video into several units that potentially improve indexing (Pereira et al., 2015; Yuan et al., 2007; Lelesci and Schönfeld, 2001). Dimensionality Reduction (DR) techniques, usual in data mining research (Ian and Kanber, 2011), represent an alternative to keep only relevant and non-redundant video indexes or combine the original indexes to create new features (Shao et al., 2014; Huang and Chen, 2011). Machine Learning (ML), in turn, can be associated with the remaining topics, for example, to discover segment boundaries or learn meaningful video indexes (Yang and Meinel, 2014; Choi et al., 2013; André et al., 2012). ML has also been useful to support video retrieval by searching for the nearest videos of a user over performing

CBVR

- content-based video retrieval
- osobní využití / institucionální využití (TV producent)
- dohledové technologie
- *surveillance video, forensic video analysis*
- lidé, auta, aktivity
- *City Surveillance*



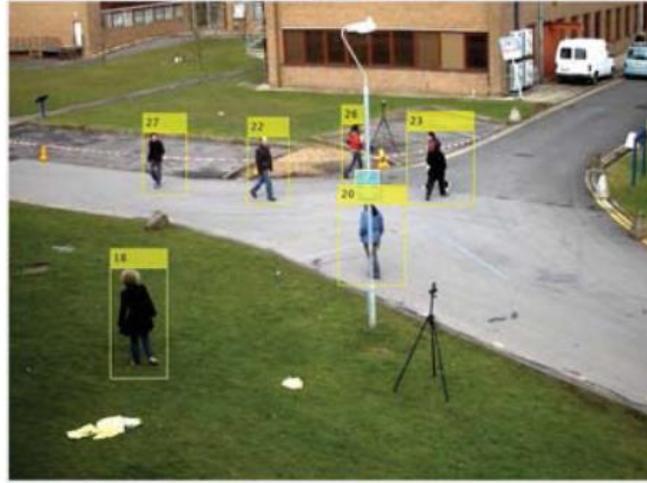


Fig. 4. Final results by the system.

TABLE II
RETRIEVAL RESULT CORRESPONDING TO DOMINANT COLOR (DC), EDGE HISTOGRAM (EH), COVARIANCE MATRIX (CM) AND SIFT DESCRIPTORS. THE QUERY BLOB IS IN THE LEFT WHILE THE FIVE RELEVANT BLOBS AND THEIR RANKS ARE IN THE RIGHT.

Query	Retrieved objects				
DC	1	36	31	25	49
EH	1	21	45	35	25
CM	1	7	24	2	10
SIFT	1	39	31	28	29

Unsupervised Surveillance Video Retrieval based on Human Action and Appearance

David Gerónimo and Hedvig Kjellström
Computer Vision and Active Perception Lab
KTH Royal Institute of Technology
Stockholm, Sweden
(dgero, hedvig)@kth.se

Abstract—Forensic video analysis is the offline analysis of video aimed at understanding what happened in a scene in the past. Two of its key tasks are the recognition of specific actions, e.g., walking or fighting, and the search for specific persons, also referred to as re-identification. Although these tasks have traditionally been performed manually in forensic investigations, the current growing number of cameras and recorded video leads to the need for automated analysis. In this paper we propose an unsupervised retrieval system for surveillance videos based on human action and appearance. Given a query window, the system retrieves people performing the same action as the one in the query, the same person performing any action, or the same person performing the same action. We use an adaptive search algorithm that focuses the analysis on relevant frames based on the inter-frame difference of foreground masks. Then, for each analyzed frame, a pedestrian detector is used to extract windows containing each pedestrian in the scene. For each detection, we use optical flow features to represent its action and color features to represent its appearance. These extracted features are used to compute the probability that the detection matches the query according to the specified criterion. The algorithm is fully unsupervised, i.e., no training or constraints on the appearance, actions or number of actions that will appear in the test video are made. The proposed algorithm is tested on a surveillance video with different people performing different actions, providing satisfactory retrieval performance.

1. INTRODUCTION

In the last decades, communications and storage technologies have undergone a huge progress, experiencing a rapid increase in the amount and quality of data generated, acquired and stored. This is certainly the case in the field of visual surveillance. Nowadays, the amount of cameras aimed at investigating crimes or causes of accidents has raised significantly, establishing themselves as a relevant tool in the task of solving crimes such as the recent Boston bombings. The field that analyzes such videos in the search for specific persons, recognize actions and interactions is called forensic video analysis [1]. Traditionally, such analysis has been carried out manually. However, during the last years, extensive research has been developed in content-based image and video mining in order to automatize the process. Related topics such as

future, potentially leaving events or objects undiscovered.

Video surveillance systems perform two key tasks, recognizing actions and searching for specific persons (e.g., re-identification), among others such as people tracking or anomaly detection. In the context of forensics we refer to these two tasks as action retrieval and appearance retrieval. Action retrieval is focused on searching for clips in which a person performs a given action, e.g., running or fighting, as defined by a textual or visual query. Appearance retrieval is focused on the same kind of search but constrained to appearance, e.g., similar clothing or biometric measures.

Although forensic video analysis in surveillance can also deal with the same scenarios as traditional online scene monitoring, e.g., shopping malls or underground stations, the nature

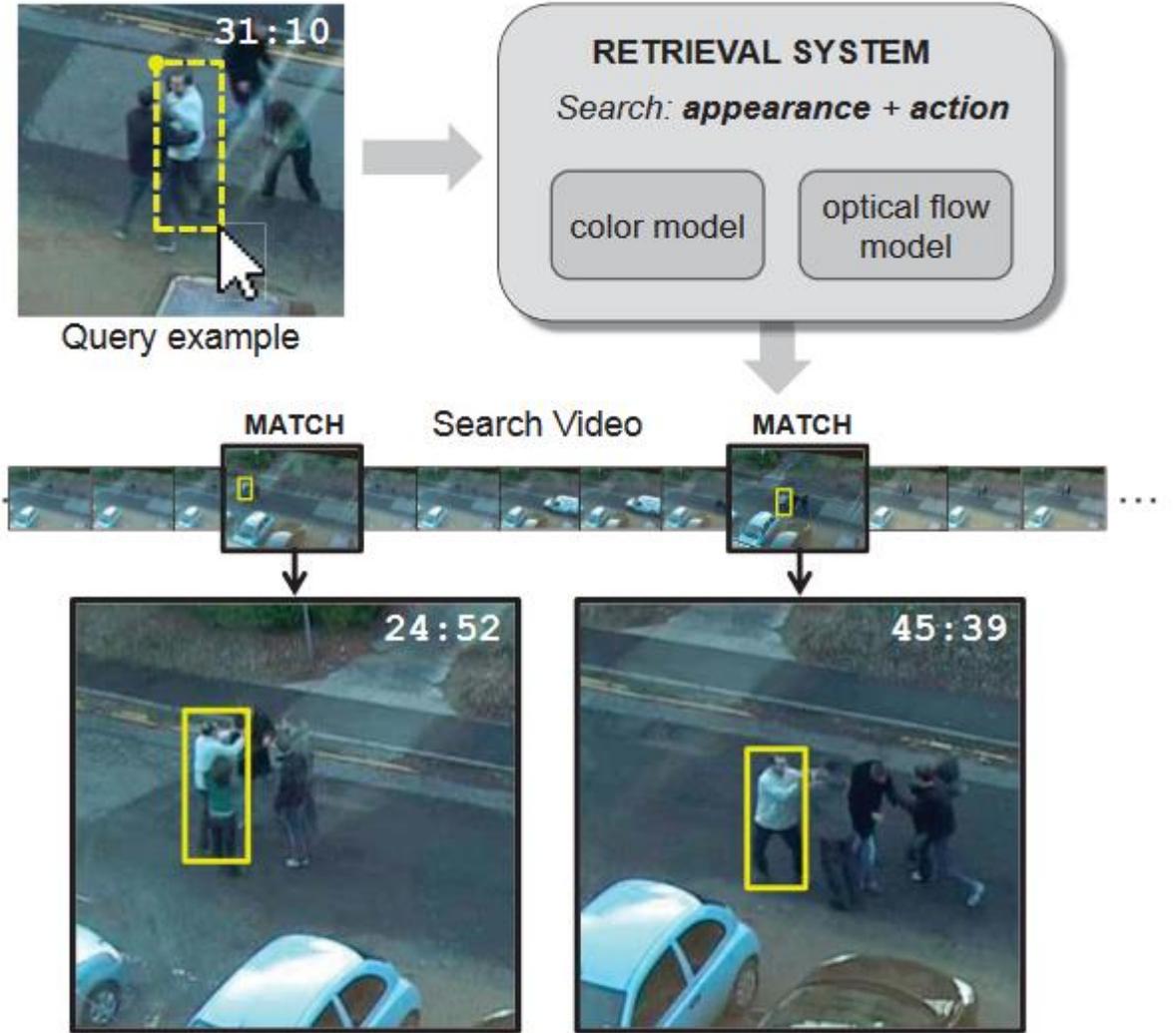


Fig. 1. Proposed surveillance image retrieval system. Given a query bounding box, the system outputs matches to the query according to its appearance and/or action in a search video.

Fig. 1. Proposed surveillance image retrieval system. Given a query bounding box, the system outputs matches to the query according to its appearance and/or action in a search video.

Retrieve by action



Retrieve by appearance

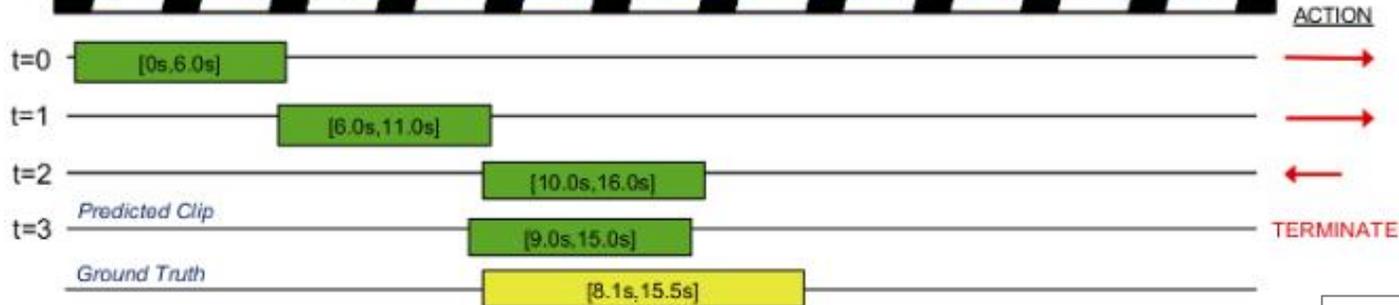


Retrieve by action+appearance



Fig. 4. Examples of queries and retrieved results. The cross marks the incorrect matches.

Query: the person begins eating the sandwich



Query: the person puts on their shoes

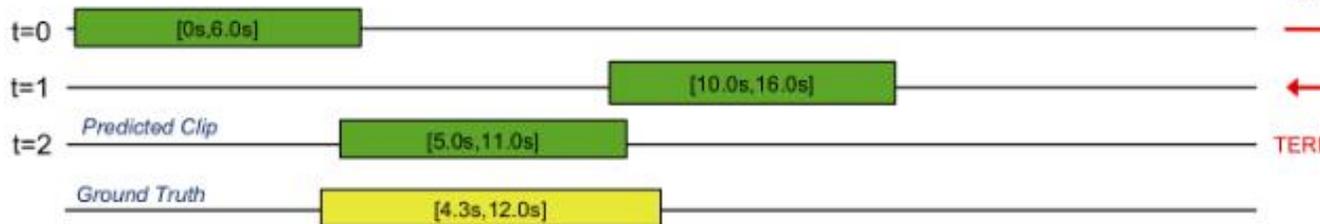


Figure 3: Qualitative performance of TripNet-GA: We show two examples where the agent skips through the video looking at different candidate windows before terminating search. Both these videos are from the Charades-STA dataset.

Tripping through time: Efficient Localization of Activities in Videos

Meera Hahn¹
meerahahn@gatech.edu
Asim Kadav²
asim@nec-labs.com
James M. Rehg¹
rehg@gatech.edu
Hans Peter Graf²
hogg@nec-labs.com

¹ College of Computing
Georgia Institute of Technology
Atlanta, GA
² Machine Learning Department
NEC Labs America
Princeton, NJ

Abstract

Localizing moments in untrimmed videos via language queries is a new and interesting task that requires the ability to accurately ground language into video. Previous works have approached this task by processing the entire video, often more than once, to localize relevant activities. In the real world applications of this approach, such as video surveillance, efficiency is a key system requirement. In this paper, we present TripNet, an end-to-end system that uses a gated attention architecture to model fine-grained textual and visual representations in order to align text and video content. Furthermore, TripNet uses reinforcement learning to efficiently localize relevant activity clips in long videos, by learning how to intelligently skip around the video. It extracts visual features for few frames to perform activity classification. In our evaluation over Charades-STA [14], ActivityNet Captions [26] and the TACoS dataset [36], we find that TripNet achieves high accuracy and saves processing time by only looking at 32-41% of the entire video.

Past TRECVID guidelines

- TRECVID 2019
- TRECVID 2018
- TRECVID 2017
- TRECVID 2016
- TRECVID 2015
- TRECVID 2014
- TRECVID 2013
- TRECVID 2012
- TRECVID 2011
- TRECVID 2010
- TRECVID 2009
- TRECVID 2008
- TRECVID 2007
- TRECVID 2006
- TRECVID 2005
- TRECVID 2004
- TRECVID 2003
- TREC-2002 Video Track
- TREC-2001 video Track

TREC Video Retrieval Evaluation: TRECVID

The [TREC](#) conference series is sponsored by the National Institute of Standards and Technology ([NIST](#)) with additional support from other U.S. government agencies. The goal of the conference series is to encourage research in information retrieval by providing a large test collection, uniform scoring procedures, and a forum for organizations interested in comparing their results. In 2001 and 2002 the TREC series sponsored a video "track" devoted to research in automatic segmentation, indexing, and content-based retrieval of digital video. Beginning in 2003, this track became an independent evaluation (**TRECVID**) with a workshop taking place just before TREC.

TRECVID 2020

TRECVID 2020 new registration is now closed. If you would like to join **TRECVID 2021** please send a request by [email](#)

TRECVID 2020 guidelines

[TRECVID 2020 guidelines](#)

TRECVID Statement on Product Testing and Advertising:

The annual "TREC Video Retrieval Evaluation" (TRECVID) is an event in which organizations with an interest in information retrieval research take part in a coordinating evaluation

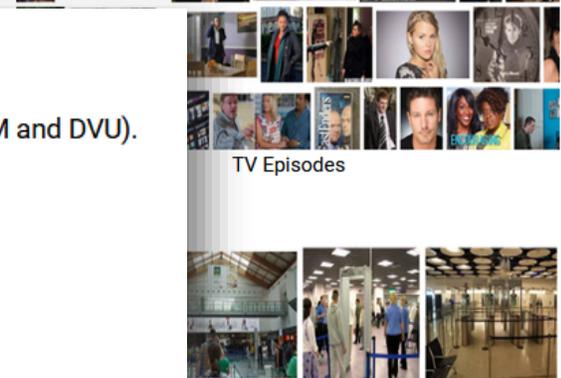
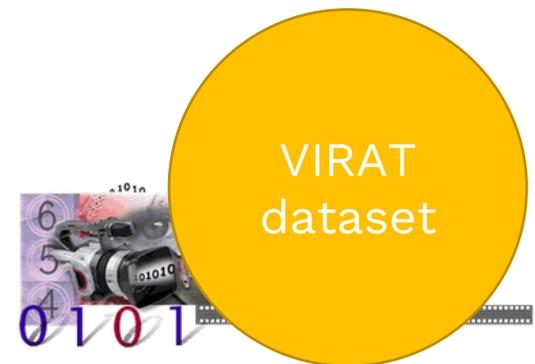
By design execution teams are from research aspect.

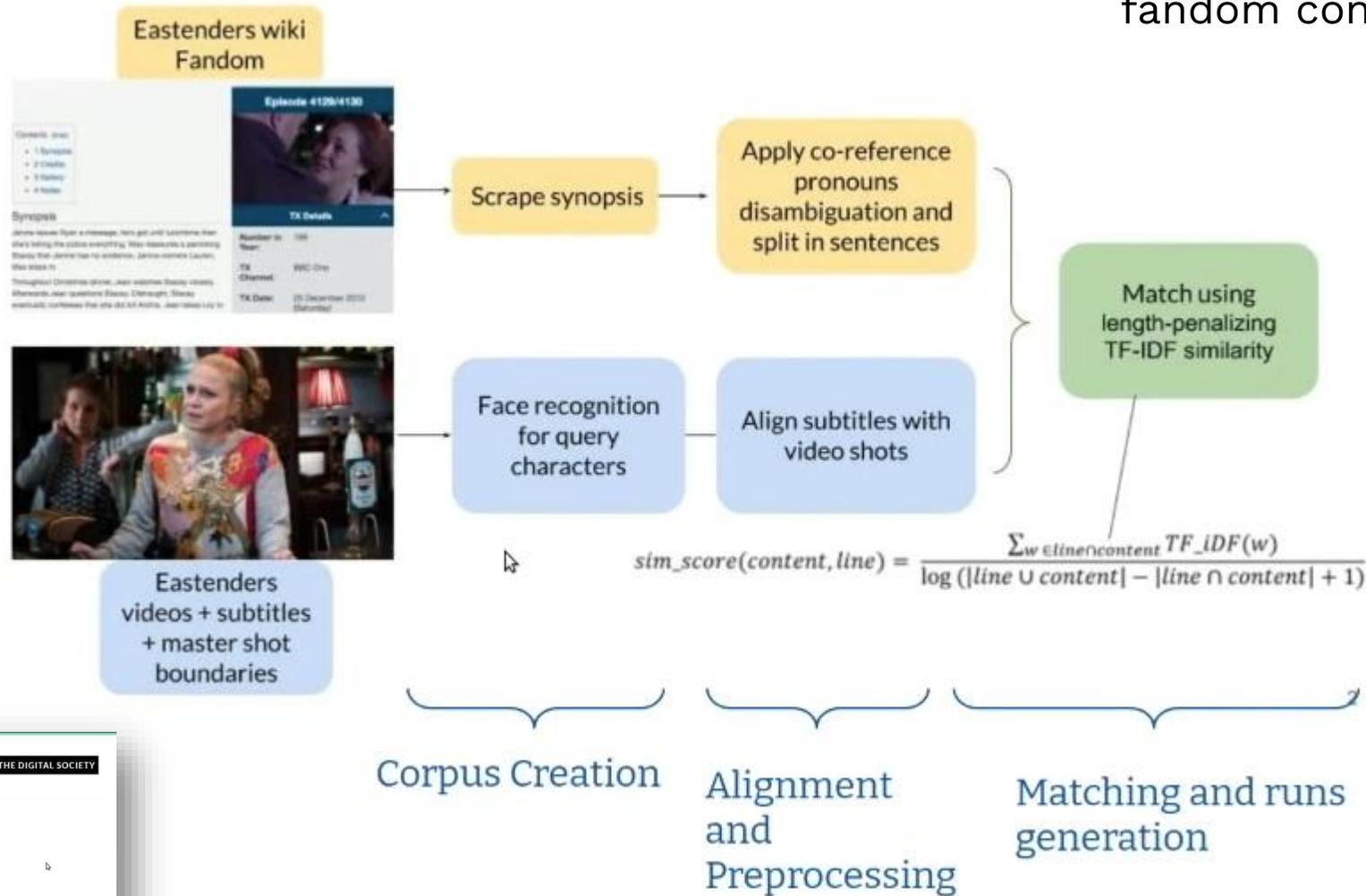
To present the dissemination or inaccurate TRECVID

TRECVID 2022 Tasks

In **TRECVID 2022**, 4 tasks (AVS, VTT, ActEV, and DSDI) will continue with some revisions, and 2 new tasks will start (MSUM and DVU).

- Ad-hoc Video Search (AVS) [**Retrieve videos from text query**]
- Activities in Extended Video (ActEv) [**Detect activities from long surveillance videos**]
- Deep Video Understanding (DVU) [**Answer questions about movies**]
- Video to Text (VTT) [**Provide a description for short videos**]
- Movie Summarization (MSUM) [**Summarize the main events of movie characters**]
- Disaster Scene Description and Indexing (DSDI) [**Classify features in Low Altitude Disaster Imagery**]





alison

IOOL AND RESEARCH CENTER AT THE HEART OF THE DIGITAL SOCIETY

EURECOM
SUDRIA UNIVERSITA

MeMAD
Methods for Managing Audiovisual Data

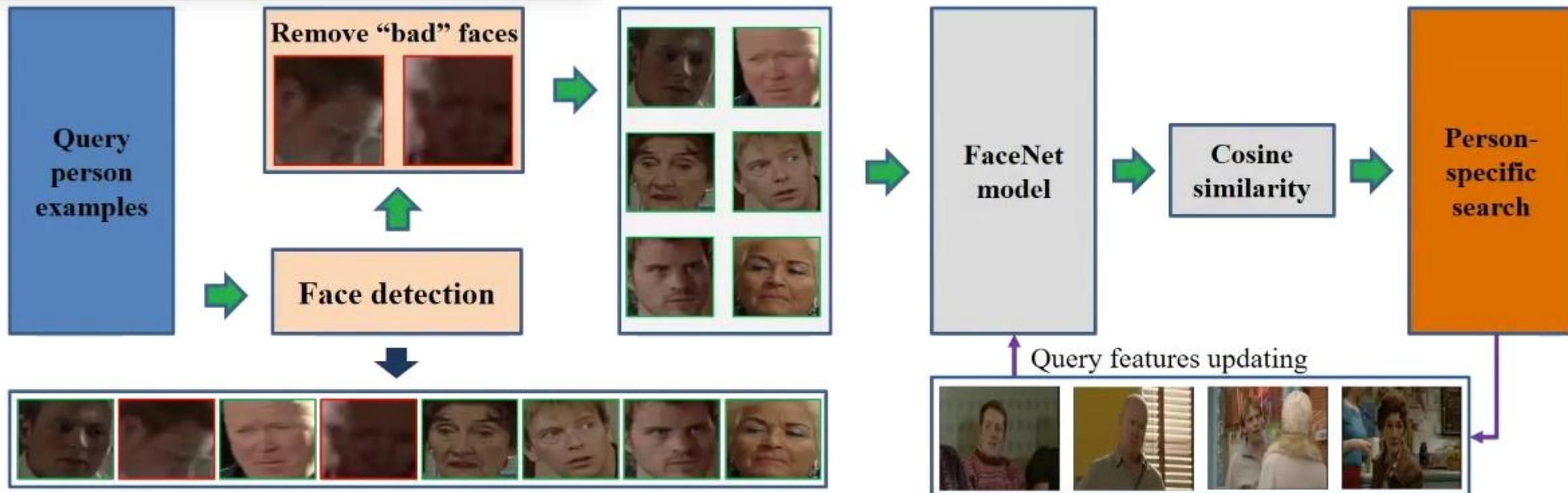
Using Fan-Made Content, Subtitles and Face Recognition for Character-Centric Video Summarization

Ismail Harrando, Alison Reboud, Jorma Laaksonen, Pasquale Lisena, Anja Virkkunen, Mikko Kurimo, Raphaël Troncy

PKU_WICT at TRECVID 2020:
Instance Search Task

**A two-stage approach with similarity computing and
result re-ranking**

Yuxin Peng, *Zhaoda Ye*, Junchao Zhang, and Hongbo Sun
Wangxuan Institute of Computer Technology,
Peking University, Beijing 100871, China
{pengyuxin@pku.edu.cn}



Frozen in Time Video Search Demo

e.g. empty street in nepz

display: 

Visual search of ~2.6M videos are based on research described in
[Frozen in time: A joint video and image encoder for end-to-end retrieval.](#)

[Previous](#)

[Next](#)

Shot 631

Frames 81666 to 81716



[Animate](#)
[Video](#)
[Thumbnails](#)
[Search](#)

Shot 632

Frames 81717 to 81836



[Animate](#)
[Video](#)
[Thumbnails](#)
[Search](#)

Shot 633

Frames 81837 to 82324



[Animate](#)
[Video](#)
[Thumbnails](#)
[Search](#)

Shot 634

Frames 82325 to 82905



[Animate](#)
[Video](#)
[Thumbnails](#)
[Search](#)

Shot 635

Frames 82906 to 82980



[Animate](#)
[Video](#)
[Thumbnails](#)
[Search](#)

Shot 636

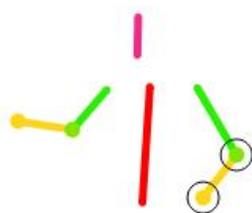
Frames 82981 to 83095



[Animate](#)
[Video](#)
[Thumbnails](#)
[Search](#)

Pose based Video Retrieval

Move the stickman joints

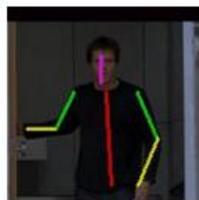


Predefined Poses

Hands On Hips At Ease Arms Folded

Restrict the search to... All Movies

Advanced options



About a Boy
01:17:47 - 01:17:53



Notting Hill
01:08:04 - 01:08:28



Love Actually
00:29:48 - 00:29:59



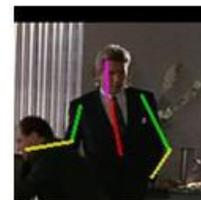
Lost in Translation
00:23:27 - 00:23:30



Apollo 13
01:01:17 - 01:01:26



Rainman
02:02:35 - 02:02:51



Pretty Woman
01:35:33 - 01:35:36



Buffy_s5e3
00:12:42 - 00:12:47



Forrest Gump
01:21:43 - 01:22:18



A fish called Wanda
00:35:56 - 00:36:05



Desperately seeking Susan
00:51:05 - 00:51:09



Groundhog Day
01:02:34 - 01:02:39

Results - mouse-over thumbnails for full frames (check the "Show Full Frames").

Go to page: 0

vitrivr

rome



FILTER

UPLOAD DATE	TYPE	DURATION	FEATURES	SORT BY
Last hour	Video	Short (< 4 minutes)	Live	Relevance
Today	Channel	Long (> 20 minutes)	4K	Upload date
This week	Playlist		HD	View count
This month	Movie		Subtitles/CC	Rating
This year	Show		Creative Commons	
			360°	
			VR180	
			3D	
			HDR	
			Location	
			Purchased	



Dermot Kennedy - Rome (Lyrics)

1.8M views • 1 year ago



Dermot Kennedy - Rome Lyrics » Stream: <https://open.spotify.com/artist/5KNNV...> Stream / Download Dermot Kennedy's debut ...



Rome Vacation Travel Guide | Expedia

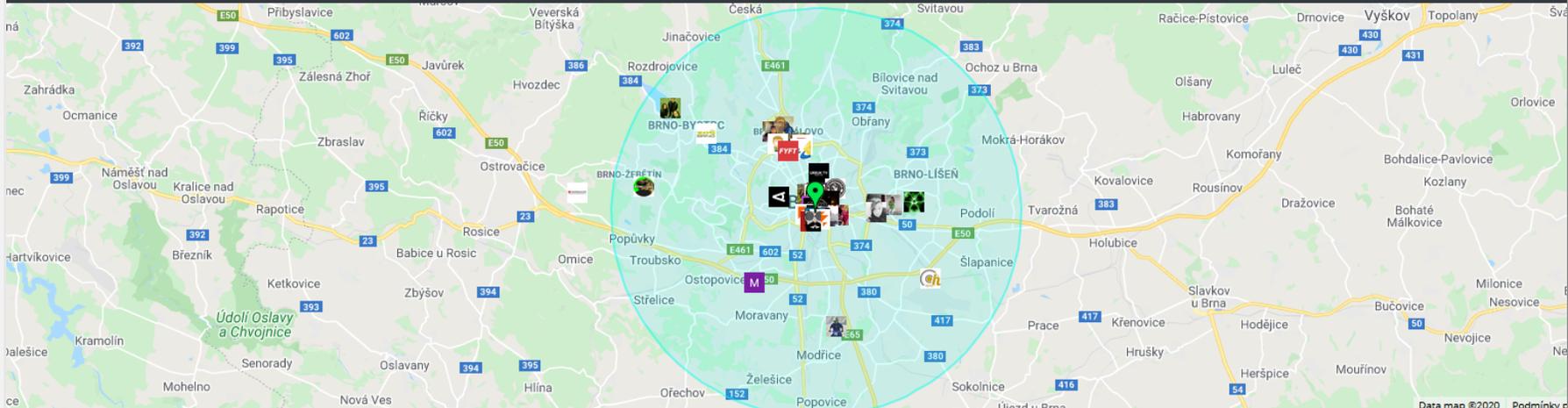
1.7M views • 7 years ago



Rome – A city of staggering history and beauty. Check out its best sights in our travel video and start planning a trip of your own so ...

YouTube Geofind

Search by: Channel Topic Location [View on GitHub](#)



The map shows the Brno region in the Czech Republic, centered on Brno-Líšeň. A light blue circular radius is drawn around the center. Numerous geotags are visible as small icons with text labels, including 'BRNO-BY...', 'BRNO-ZE...', 'BRNO-LÍSEŇ', and 'BRNO-MLOVO'. The map includes road networks, green spaces, and various town names.

By Location

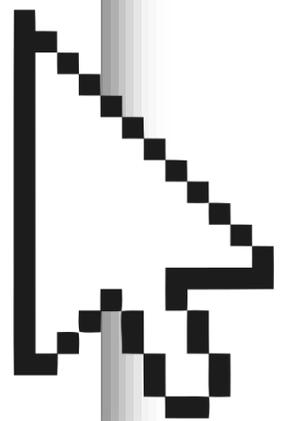
Update December 15th! Location and topic searching now include language filter on the geotags results. ✕

Location 📍
Drag and drop the green point or type desired location and hit enter.

Radius ⌵
YouTube may return results close yet outside the specific search radius.

Keywords
Keywords as if actually searching YouTube.

Sort by ⌵



<https://www.youtube.com/watch?v=tCFWraMxc-o>

Vyhledávačka na doma



Dorazil uživatel se zájmem o lidové kroje a je rozmrzelý, že je nemůže prohlížet přímo v muzeích, do kterých měl ve zvyku jezdit.

Pokusme se pro něj identifikovat vhodné zdroje digitalizovaných krojů – aby je mohl zkoumat a prohlížet alespoň jako obrázky v co nejvyšší kvalitě.

Vyhledávačka na doma



V knihovně jste našli jediný bestseller autora Sherwooda Andersona (*Dark Laughter*) v anglickém vydání z roku 1925. Na předsádce byla ilustrace, ale ve vašem vydání se desky slepily a při odtrhnutí vám z celé ilustrované předsádky zbyl jen malý nepoškozený kousek – zbytek je nenávratně pryč. Pokuste se dohledat, co se na obrázku v předsádce této knihy dělo...

