Exploring the synergy between knowledge graph and computer vision for personalisation systems

Vincent Lully, Philippe Laublet, Milan Stankovic, Filip Radulovic

Sorbonne Université & Sépage, Paris, France

Imagine that
you like this image:

What interests can we infer?

- Headgear?
- Colourful clothes?
- Tz'utujil people?
- Lake Atitlan?

2 problems

- Image user profiling
- Image selection in recommendation banners (A personalisation application)

Demo


QuickTime Player File Edit View Window Help
Tue 16:08
Image user profiling demo
localhost:9000/image-user-profiling
110 %
Rechercher
Les plus visités
Débuter avec Firefox
Please click on an image and check the created user profile.

Background

- Computer vision (CV) applications
  - Face detection
  - Content-based image retrieval
  - Automatic photo annotation
  - Autonomous cars
- Knowledge graph (KG) applications
  - Semantic search
  - Exploratory search
  - Document similarity calculation
  - Question answering
- CV and KG applications
  - Object detection with external knowledge graphs
  - Scene description with triples
  - Knowledge graph completion with visual features
  - Visuo-semantic search

Background

- CV and KG applications in personalisation systems insufficiently studied
- Efforts concentrated on analysing textual data
- Lots of multimedia data available on the Web and being produced continuously
  - Publication of videos and photos on the Social Web
  - Photos play an important role in decision making on e-commerce websites
- Modern websites should be armed with facilities which can understand users’ interests through their interactions with multimedia data and adapt the services accordingly in order to provide a better user experience.

Workflow

user profiling

personalisation tasks

Input images

Knowledge graph entities

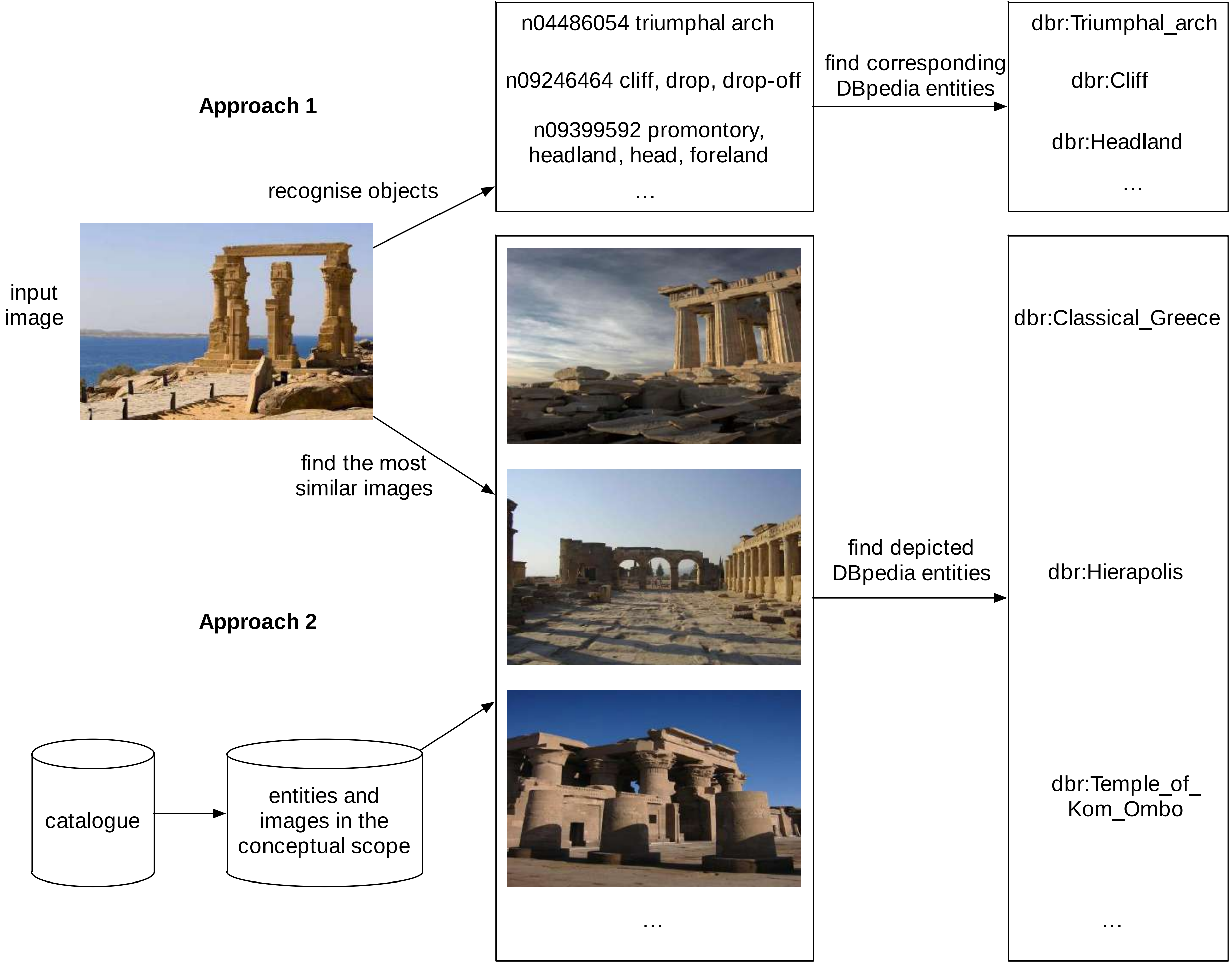
- Item recommendations
- Image selection in recommendation banners
- Reverse image search

Image2Entity : semantic image user profiling

- Input : an image
- Output : Top-n knowledge graph entities
- Knowledge graph : DBpedia, Wikidata or other large-scale KGs
- 2 approaches :
  - Object detection and entity liking
  - Catalogue-driven visual similarity

Approach 1 : object detection and entity linking

- Map an image to entities corresponding to the objects appearing in it.
- Step 1 object detection
  - CV tool Inception-V3 convolutional neural network model trained for ImageNet Large Visual Recognition Challenge using the data from 2012
  - Map to 1000 WordNet synsets like “gazelle” and “patio, terrace”
- Step 2 entity linking
  - Map synsets to DBpedia entities


Approach 1
recognise objects
n04486054 triumphal arch
n09246464 cliff, drop, drop-off
n09399592 promontory, headland, head, foreland
…
find corresponding DBpedia entities
dbr:Triumphal_arch
dbr:Cliff
dbr:Headland
…
input image
find the most similar images
Approach 2
catalogue
entities and images in the conceptual scope
…
find depicted DBpedia entities
dbr:Classical_Greece
dbr:Hierapolis
dbr:Temple_of_Kom_Ombo
…

Approach 1 : object detection and entity linking

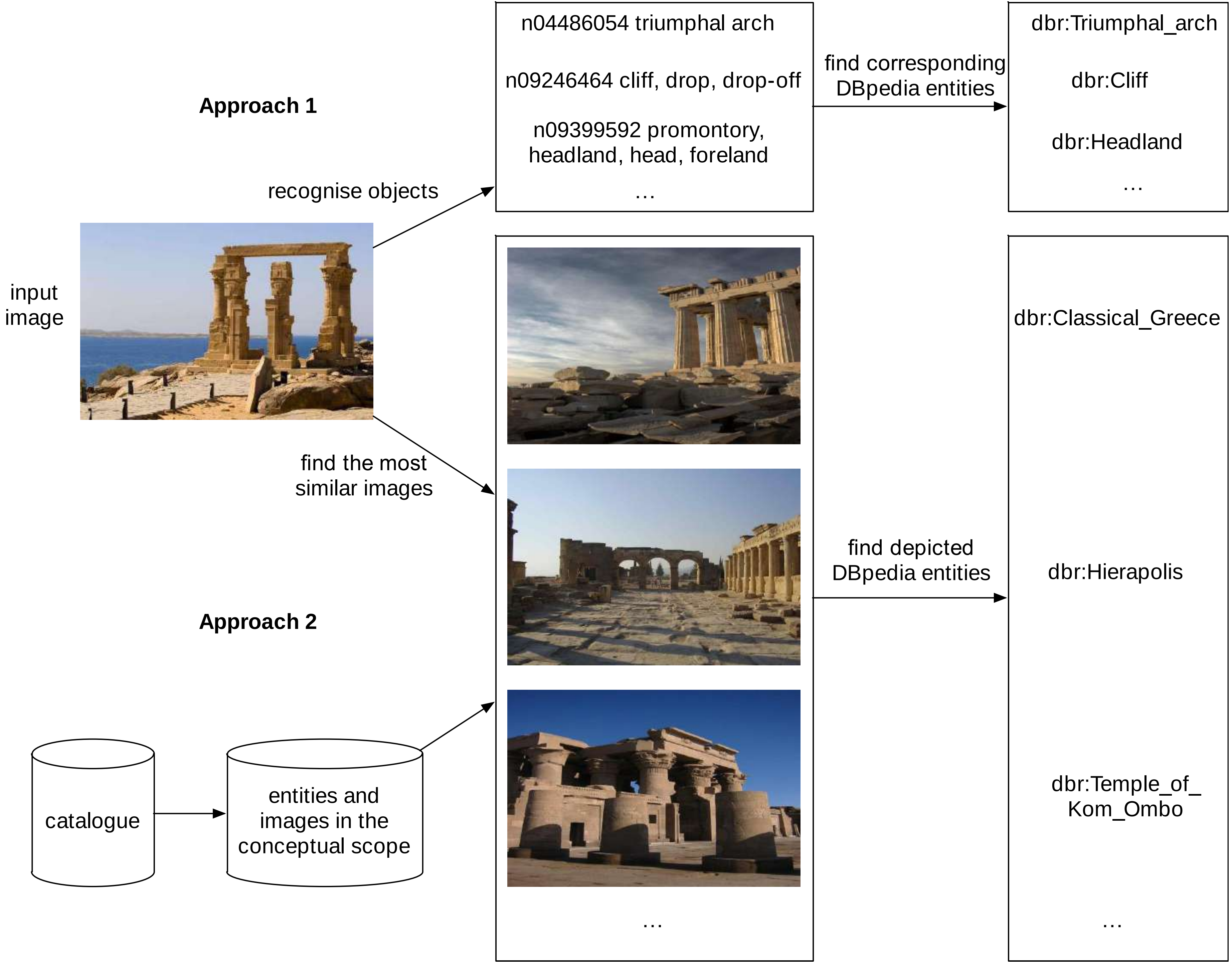
- Basic and obvious approach but not found in the literature
- 2 shortcomings
  - limited number of mappable synsets, large qualitative training data needed
  - 1000 mappable sysnets not necessarily in line with the image and not within the conceptual scope of the catalogue

Approach 2 : catalogue-driven visual similarity

- Catalogue conceptual scope
  - All knowledge graph entities which directly appear in the catalogue.
  - Direct item linking (for items having corresponding knowledge graph entries like films and artists) and item description linking (for items with rich textual descriptions)
  - In case of need, e.g. the number of appearing entities is too small, we may enrich with the entities which are closely related to the appearing ones.
  - The main idea is to map to entities which can contribute to the semantic similarity calculation, in other words, which are useful in further personalisation tasks. Thus, ideally, the conceptual scope should be defined by considering the semantic similarity calculation to be adopted further. For example, we may enrich with the entities by a set of selected object properties, or the ones by 1-hop category enrichment or the ones used as dimensions in embeddings.

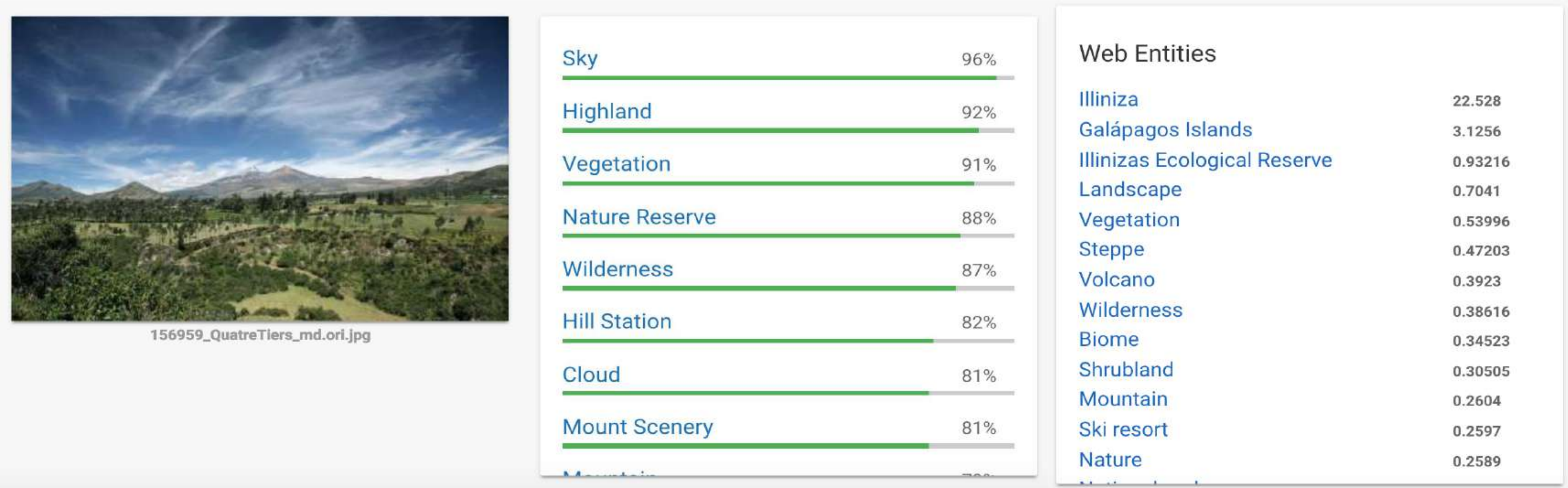
Approach 2 : catalogue-driven visual similarity

- Map an image to entities which are depicted by visually similar images and exist in the conceptual scope of the catalogue within which further personalisation tasks are conducted
- Step 1 create the conceptual scope of the catalogue
- Step 2 retrieve images depicting the entities in the scope (foaf:depiction)
- Step 3 compute pairwise visual similarity
  - Penultimate layer outputted by Inception-V3 which is a 2048-dimensional vector
  - Euclidean distance between vectors
- Step 4 output top-n entities


n04486054 triumphal arch
n09246464 cliff, drop, drop-off
n09399592 promontory, headland, head, foreland
…
dbr:Triumphal_arch
dbr:Cliff
dbr:Headland
…
find corresponding DBpedia entities
Approach 1
recognise objects
input image
find the most similar images
Approach 2
catalogue
entities and images in the conceptual scope
find depicted DBpedia entities
dbr:Classical_Greece
dbr:Hierapolis
dbr:Temple_of_Kom_Ombo
…
…

Evaluation

- Baseline : Google Cloud Vision API
- Label detection (GL) and web entity detection (GE)
- Google’s KG --> DBpedia (“/m/017rgb” to “dbr:Ferris_wheel”)



Evaluation

- Commercial catalogue of a popular French travel agency
- 1,357 world-wide travel packages in 136 countries and regions
- 11,614 distinct images
- 50 diverse and representative images selected by hierarchical clustering
- Evaluation dataset created by 3 annotators
- Each image has a ranked list of 5 entities as ground truth.
- 4 compared approaches : I2E1, I2E2, GL, GE
- Conceptual scope needed for I2E2 is done with an entity linking tool Dandelion.
- Each approach outputs top-5 entities which are compared with the ground truth.

Evaluation

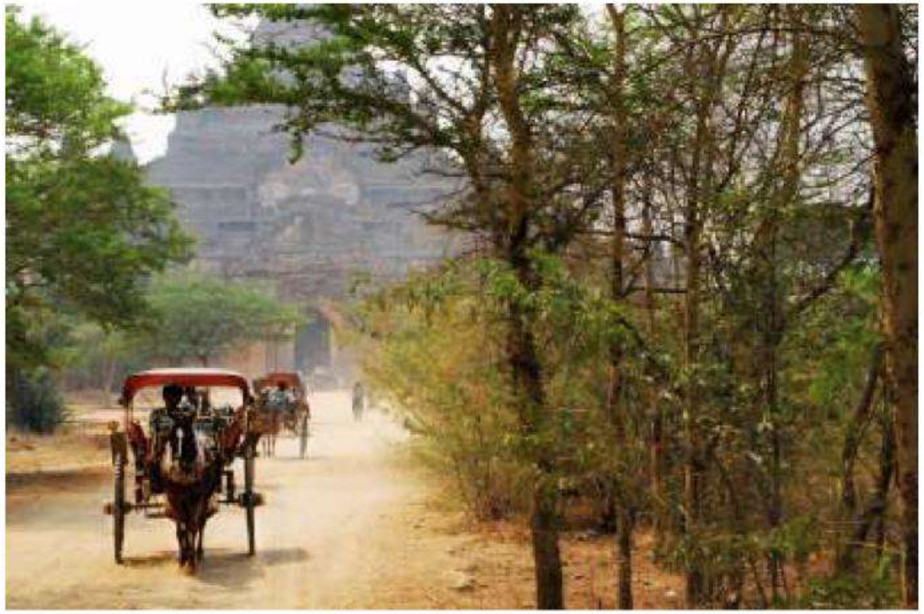
- 4 metrics
  - How relevant are the individual mapped entities? (precision)
  - How many relevant entities are successively mapped? (recall)
  - How early can we find a relevant entity? (mean reciprocal rank)
  - How is the global relevancy of the mapped entities? (normalised discount cumulative gain)

Evaluation

	I2E1	I2E2	GL (baseline)	GE (baseline)
precision	0.384	0.576	0.524	0.376
σ	0.279	0.276	0.328	0.282
p-value GL	< 0.01	> 0.1		
p-value GE	> 0.1	< 0.01		
recall	0.226	0.348	0.288	0.223
σ	0.177	0.152	0.138	0.167
p-value GL	< 0.1	< 0.05		
p-value GE	> 0.1	< 0.01		
mean reciprocal rank	0.692	0.886	0.77	0.551
σ	0.407	0.257	0.361	0.406
p-value GL	> 0.1	< 0.05		
p-value GE	< 0.1	< 0.01		
nDCG	0.311	0.478	0.401	0.3
σ	0.245	0.217	0.232	0.24
p-value GL	< 0.05	< 0.1		
p-value GE	> 0.1	< 0.05		

Evaluation

- I2E2 > GL > I2E1 > GE, differences statistically significant mostly
- 2 reasons may explain the less good performance of the baseline :
  - Correct labels but wrong entities, GL outputs “path”, “dbr:Path_(graph_theory)”
  - Entities too generic “dbr:Building”, “dbr:Travel”
- I2E1 achieves a better performance than expected. Images in line with the synsets: fauna, flora, natural landscape



Evaluation

- The better performance of I2E2 may be explained by its capacity of capturing the general atmosphere of an image rather than the objects.
  - A train in a rural area, “dbr:Mokra_Gora” not train objects
  - Women wearing traditional clothes, “dbr:Wayuu_people” and “dbr:Toraja” instead of “dbr:Sombrero”
- Limitations of I2E2:
  - Dependent to the quality of entity linking tools
  - Depedent to images depicting the entities, “dbr:Oriental_(Morocco)” is only one of the twelve regions of Morocco, depicted by an image of a mountainous landscape of Jebel Tamejout which is not representative of the whole region and it is not very reasonable to determine that a user would be interested in the region. Not working well on inclusive geographic entities (coutries), abstract entities (love, history).

Image selection in recommendation banners

- Multiple studies and applications have shown the influence of images in the perception of items.
  - The consumer’s perception of a hot beverage would be influenced by the color of the plastic vending cup from which it is served.
  - The dating application Tinder shows the best photo as the first photo with the functionality “Smart Photos” and claims to increase matches by 12%.
  - The images hotels choose to display have significant impact on click-through rate.
- Hypothesis : by displaying images more in line with the user’s interests we can improve the user’s perception of the recommended items.

Image selection in recommendation banners

- The user’s interests (user profile) and an item’s images (image profile) are represented by DBpedia entities.
- Thus, they are put in the same conceptual space and we can calculate their similarity.
- Knowledge-based approach
  - User profile : interaction with textual or image data
  - Image profile : I2E2
  - Similarity : Jaccard measure

Evaluation

- 2-stage user study with 32 participants (18 women, 14 men, 22 to 32 years old)
- Same travel dataset
- Baseline:
    - *Random.* We randomly select an image among available ones.
    - *Agent.* We select the image ranked first by the human travel agent who makes the catalogue. This assumes that the travel agent privileges images which are the most attractive in general.

Evaluation – survey design

- Step 1 Participants put themselves in the scenario of searching for a package tour for their next vacation. They simulate a browsing experience on a web interface where they can visualize the tours where each tour is described by some basic information. They make a selection.

Please select the travel products which appeal to you and submit your selection by clicking on the button at the bottom of this page.

☐ KAZAKHSTAN , 11 jours Le Kazakhstan, une terre mystérieuse Voyage découverte du Kazakhstan et de ses richesses culturelles et naturelles,	☐ TURQUIE , 8 jours La Cappadoce en liberté Randonnée entre cheminées de fée et vallées au mille couleurs...	☐ EMIRATS ARABES UNIS ,QATAR , 11 jours Magie de lArabie! Découverte des sites majeurs du Qatar et des Emirats Arabes Unis.
☐ ISRAEL , 8 jours Au coeur du Néguev randonnée nature au coeur du biblique désert du Néguev et découverte de Massada et Jérusalem	☐ SRI LANKA , 13 jours De Colombo à Jaffna Safari, balades au coeur des montagnes, triangle culturel et Péninsule de Jaffna	☐ INDONESIE , 16 jours Belle échappée en Sulawesi (VTT) Découverte en VTT des trésors naturels de Sulawesi : lacs, montagnes et rizières et village ethnique !
☐ EMIRATS ARABES UNIS , 8 jours Perles des Emirats Découverte des Emirats Arabes Unis:Dubaï, Abu Dhabi, oasis dAl Ain et kayak insolite	☐ BIRMANIE , 14 jours Un passeport pour la Terre d'Or! Classiques de la Birmanie et mini treks dans la région de Loikaw et Mont Victoria	☐ MADAGASCAR , 3 jours Trekking à Bekopaka Rando dans les grands tsingy du Parc de Bemaraha
☐ GUINEE BISSAU , 9 jours Bijagos, l'archipel authentique Odyssée dîle en île dans une nature préservée, sa faune riche, son peuple animiste	☐ FINLANDE , 8 jours Vagabondage en Laponie	☐ NAMIBIE , 22 jours Total trip en Namibie Version complète du Kalahari au Fish River Canyon et les chutes d'Epupa entre autres
☐ SAO TOME , 9 jours Trek équatorial sur les traces du cacao ! Trek en forêt équatoriale, plantations de café, cacao, visites et plages sauvages !	☐ MALAISIE , 15 jours Villages, jungle et eaux turquoises LE best of de la Malaisie avec l'authenticité en prime	☐ SEYCHELLES , 10 jours Douceur des îles en Lodge ! Découverte des Seychelles et de ses 3 parcs marins en hébergements de charme !

Evaluation – survey design

- Step 2 Participants see a recommendation banner without image. They rate on a 5-level Likert scale 3 aspects: persuasion, effectiveness and attention, we call this the first rating stage.
- Pesuasion : I am interested in this recommendation.
- Effectiveness : I have sufficient information to decide whether I click on the recommendation or not.
- Attention : The recommendation banner captures my attention.

CANADA , 11 jours

La magie du Québec

Entre faune et fjord, enfilez votre plus belle
chemise à carreaux et partez à la découverte
de la Belle Province canadienne !

Evaluation – survey design

- Step 3 Participants see 3 recommendation banners with images. They rate again persuasion, effectiveness and attention. They also rate efficiency and affinity.
- Efficiency : The image helps me decide more rapidly whether to discover more about it or not.
- Affinity : The image shows things that I am in affinity with.

CANADA , 11 jours
La magie du Québec
Entre faune et fjord, enfilez votre plus belle chemise à carreaux et partez à la découverte de la Belle Province canadienne !

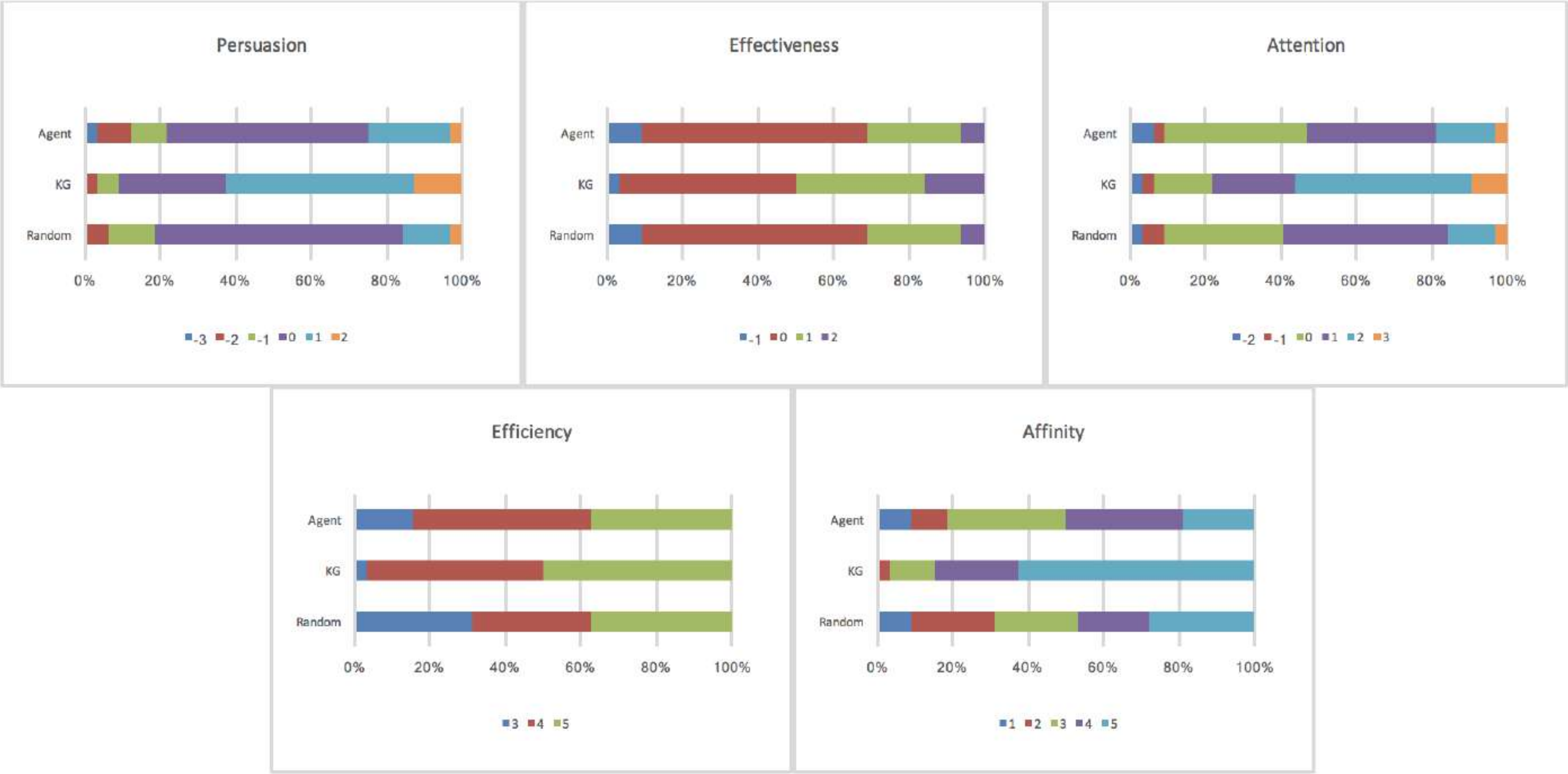
CANADA , 11 jours
La magie du Québec
Entre faune et fjord, enfilez votre plus belle chemise à carreaux et partez à la découverte de la Belle Province canadienne !

CANADA , 11 jours
La magie du Québec
Entre faune et fjord, enfilez votre plus belle chemise à carreaux et partez à la découverte de la Belle Province canadienne !

Evaluation

Aspect	Statement	Stage	Metric
Persuasion	I am interested in this recommendation.	1, 2	Rating change
Effectiveness	I have sufficient information to decide whether I click on the recommendation or not.	1, 2	Rating change
Attention	The recommendation banner captures my attention.	1, 2	Rating change
Efficiency	The image helps me decide more rapidly whether to discover more about it or not.	2	Rating
Affinity	The image shows things that I am in affinity with.	2	Rating

Evaluation


Persuasion
Agent
KG
Random
0% 20% 40% 60% 80% 100%
-3 -2 -1 0 1 2
Effectiveness
Agent
KG
Random
0% 20% 40% 60% 80% 100%
-1 0 1 2
Attention
Agent
KG
Random
0% 20% 40% 60% 80% 100%
-2 -1 0 1 2 3
Efficiency
Agent
KG
Random
0% 20% 40% 60% 80% 100%
3 4 5
Affinity
Agent
KG
Random
0% 20% 40% 60% 80% 100%
1 2 3 4 5

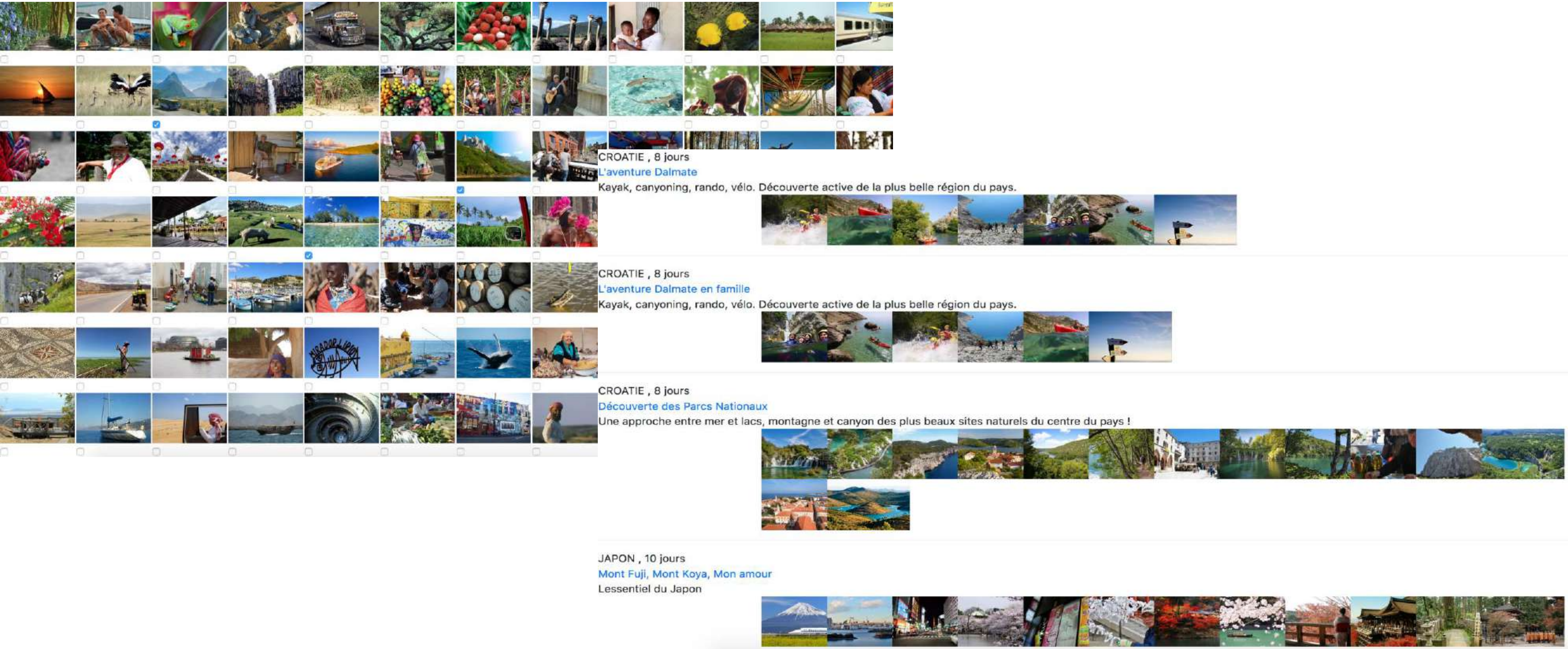
Evaluation

- KG outperforms other approaches on affinity and the difference is statistically significant. This shows that KG is actually capable of selecting an image corresponding to users' interests. Several participants comment that the images selected by KG reflect exactly the trip they imagine during the browsing simulation phase.
- We observe a net increase for all approaches on attention and efficiency. This shows that displaying an image in a recommendation banner can better capture users' attention. Images can help users decide more rapidly whether to discover more about the recommendation or not, as voiced by the participants with the majority positive ratings.
- The results on effectiveness are not conclusive. For all three approaches, most participants did not find that the images provide much more information to help them decide. The other information displayed in a textual form was sufficient to help them decide whether to click or not.
- On persuasion, certain participants see their interest in the recommendation decrease after seeing the banners with certain images. Some participants comment that some images are so uninpirational that they even negate the teasing effect of textual descriptions. On this aspect, KG actually enhances the perception of the recommendations while the other two approaches have rather neutral or negative impact. This shows the importance of carefully selecting the images which correspond to users' interests.

Conclusion

- Exploring the synergy between knowledge graph and computer vision for personalisation systems
- 2 semantic image user profiling approaches which create accurate and useful profiles
- 1 knowledge-based approach for image selection in recommendation banners which can enhance the perception of recommended items
- Future work
  - Research on the impact of the conceptual scope
  - Application : travel reverse image search

Travel reverse image search



Demo


QuickTime Player File Edit View Window Help
Tue 16:08
Image user profiling demo
localhost:9000/image-user-profiling
110 %
Rechercher
Les plus visités
Débuter avec Firefox
Please click on an image and check the created user profile.