Microsoft Research 微软亚洲研究院

WWW 2009

Tag Ranking

Dong Liu, Xian-Sheng Hua, Linjun Yang, Meng Wang, Hong-Jiang Zhang

Microsoft Research Asia

Harbin Institute of Technology

Social Media and the Associated Tags -
Towards Large-Scale Content-Based Multimedia Search

flickr

www, www2009, madrid, spain
w3c, Don Quixote, Don, Quixote
cervantes, Sancho, …

www2009, w3c, futuro, future, workshop, congreso
palacio, municipal, Madrid, consortium, consorcio
20, aniversario, España, Spain, Vinton, …

YouTube

Social tags are good, but they are

- Noisy
- Ambiguous
- Incomplete
- No relevance information

Two directions to improve tag quality

- During tagging – Tag Recommendation
- After tagging – Tag Refinement/Ranking

The most relevant tag is NOT at the top position in

Social tags for online images are better than automatic annotation in terms of both scalability and accuracy.

This phenomenon is widespread on social media websites such as Flickr.

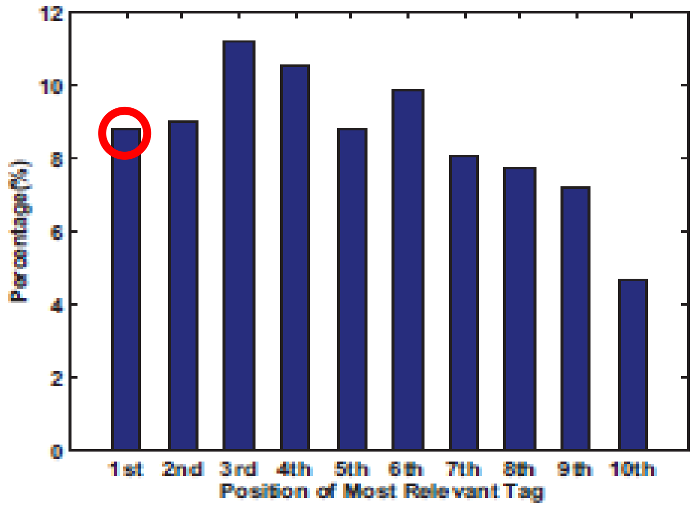
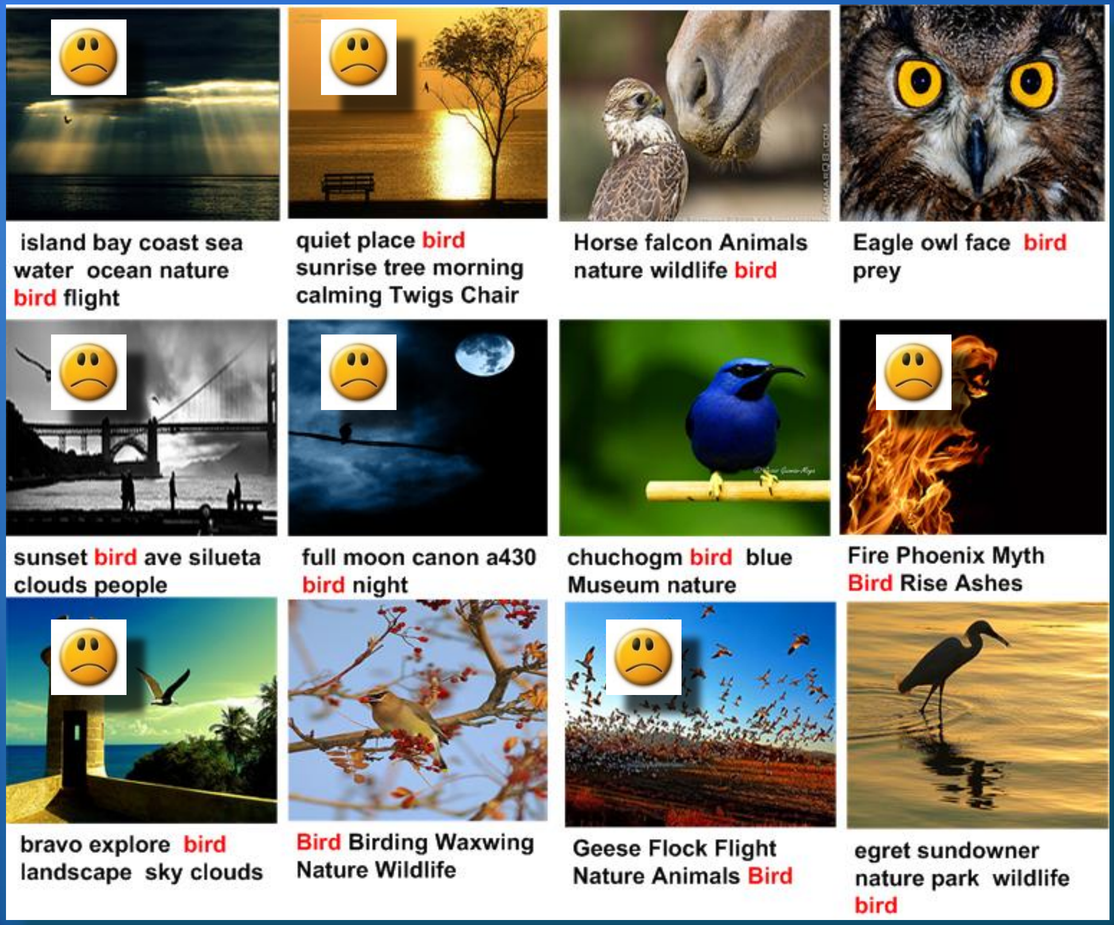


Figure 2: Percentage of images that have their most relevant tag at the n-th position in the associated tag list, where $n = 1, 2, 3 \ldots, 10$.

Only less than 10% images have their most relevant tag at the top position in their tag list.

This has significantly limited the performance of tag-based image search and other applications.

For example, when we search for "bird" on Flickr.



What we are going to do:

Rank the tags according to their relevance to the image.

But how can we make it?

Automatically.

Image Ranking v.s. Tag Ranking

- Image Ranking
  - Order images according to the relevance (of the images) to the query term
- Tag Ranking
  - Order tags according to the relevance (of the tags) to the image

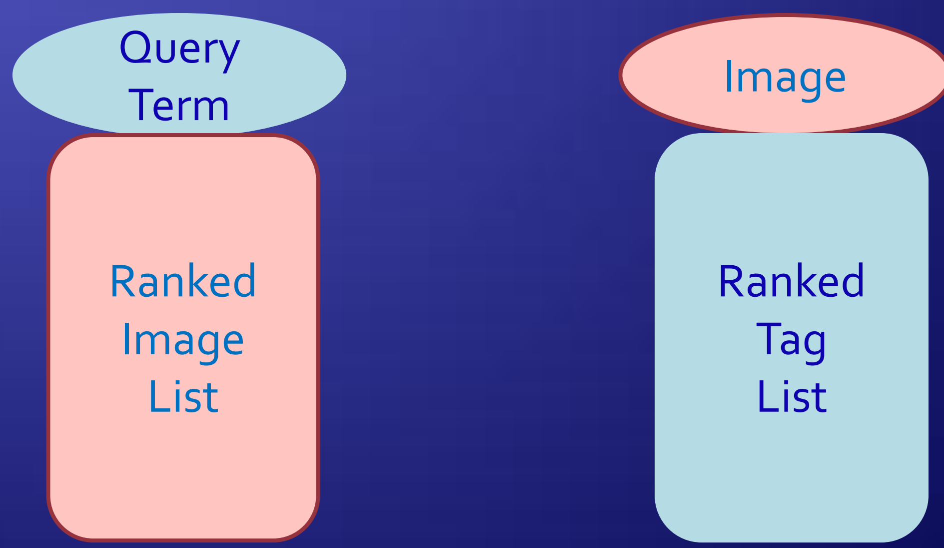


Image Ranking   Tag Ranking

Image *Re*ranking v.s. Tag Ranking

- Image Reranking
  - Initial image ranking list → Improved ranking list
- Tag Ranking
  - Initial tag list (no order)→ Ranked tag list

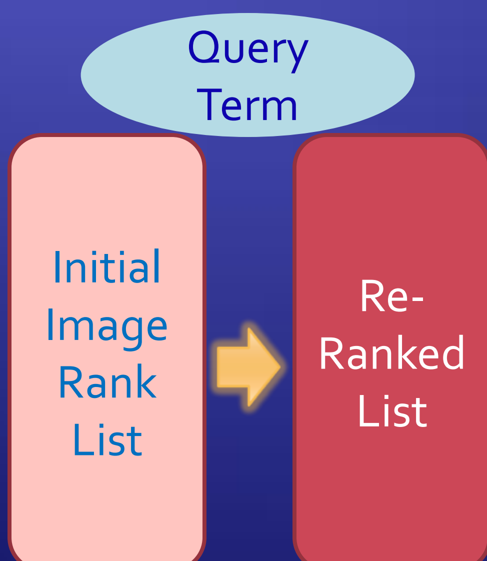
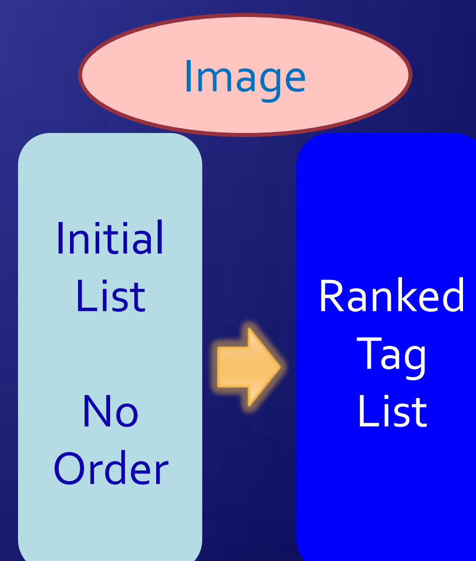


Image ReRanking



Tag Ranking

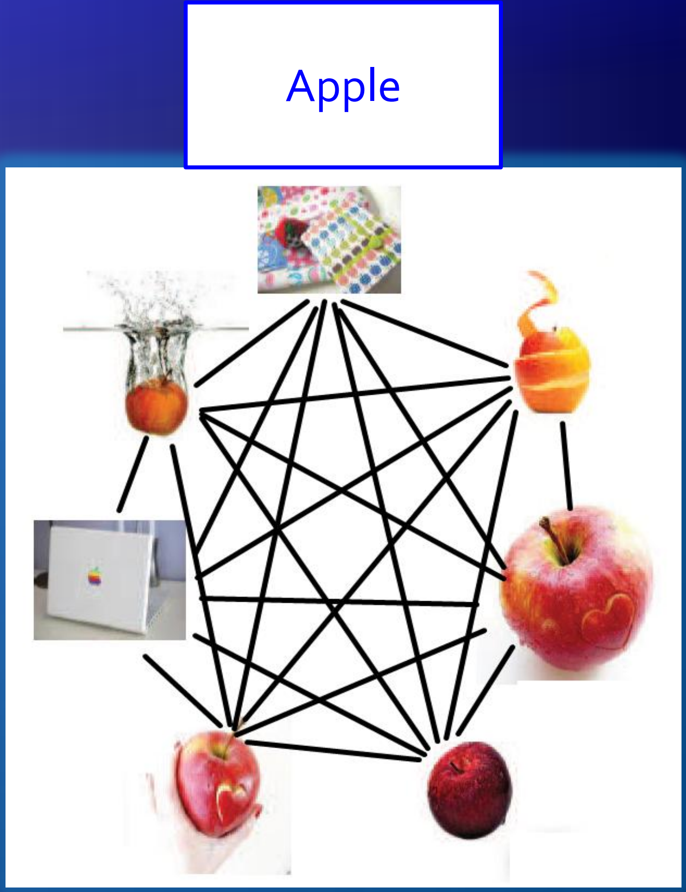
Can we borrow some idea from image reranking?

Basic Assumptions

- Image Reranking
  - Large image clusters should be promoted
  - Visually similar images should be ranked closely

Typical Image ReRanking – Random Walk

- Graph construction
  - Images as nodes
  - Rank or ranking score of an image as the value of the node
  - Visual similarities of images are the edges
  - Transition probability between two nodes
- Graph Iteration
  - To refine the relevance scores step by step
  - With the help of the scores of the visually similar images

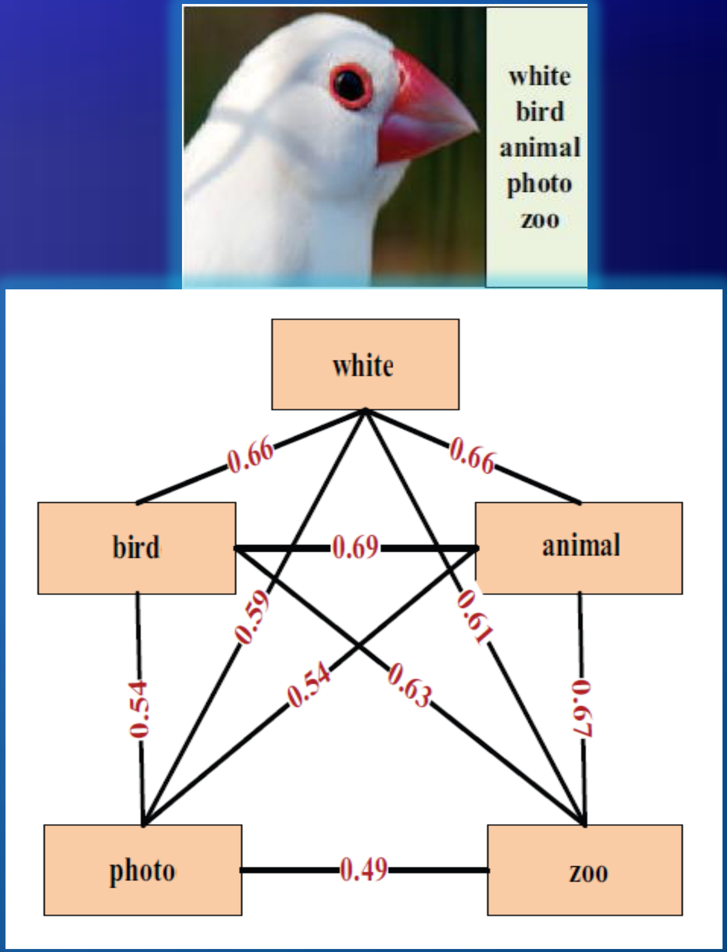


Basic Assumptions

- Image Reranking
  - Large image clusters should be promoted
  - Visually similar images should be ranked closely
- Tag Ranking
  - Large tag clusters should be promoted
  - Semantically close tags should be ranked closely

Tag Ranking (for each image)

- Graph construction
  - Tags as nodes
  - Rank of a tag as the value of the node
  - Semantic similarities of tags are the edges
  - Transition probability between two nodes
- Graph Iteration
  - To refine the relevance scores step by step
  - With the help of the scores of the semantically close tags



The problem is:

How can we calculate the similarity or distance of two tags?

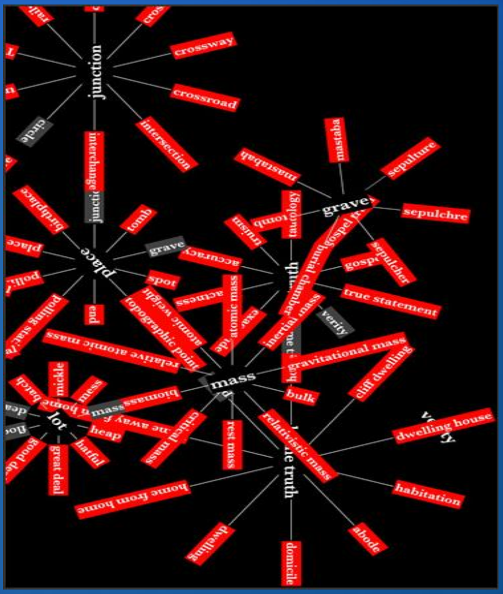
What We Can Use

- WordNet distance
- Google distance

WordNet Distance

PRINCETON UNIVERSITY

- WordNet
  - 150,000 words
- WordNet Distance
  - Quite a few methods to get it in WordNet
  - Basic idea is to measure the length of the path between two words
- Pros and Cons
  - Pros: Built by human experts, so close to human perception
  - Cons: Coverage is limited and difficult to extend



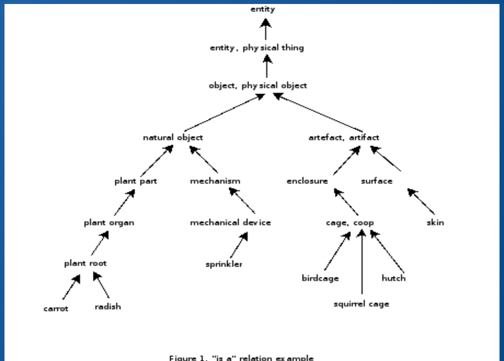


Figure 1. "is a" relation example

Google Distance

University of Amsterdam

- Normalized Google Distance (NGD)
  - Reflects the concurrency of two words in Web documents
  - Defined as

$$NGD(x,y) = \frac{max(\log f(x), \log f(y)) - \log f(x,y)}{\log N - min(\log f(x), \log f(y))}$$

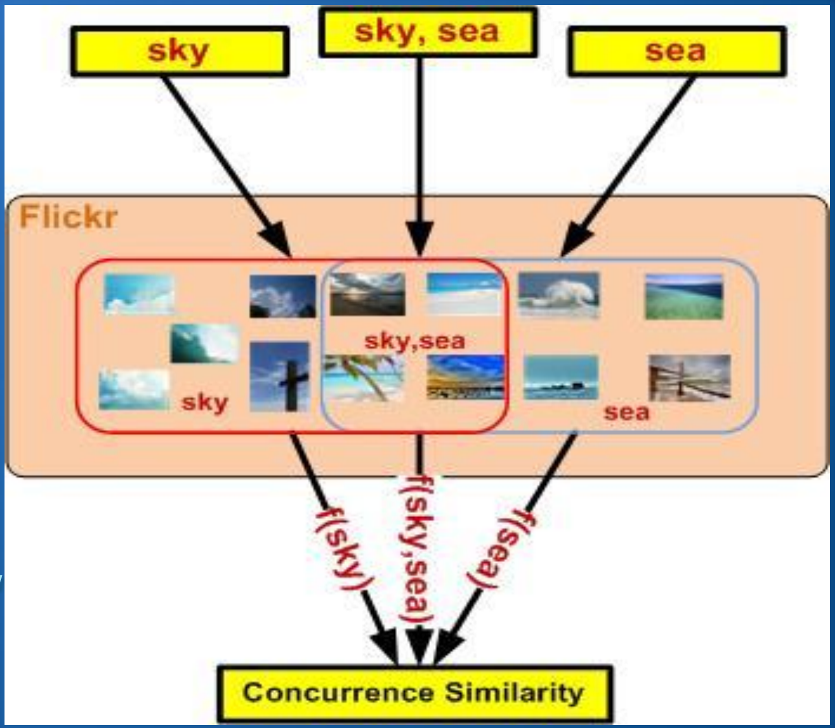
- Pros and Cons
  - **Pros**: Easy to get and huge coverage
  - **Cons**: Only reflects concurrency in textual documents. Not really concept distance (semantic relationship)

What We Can Use

- WordNet distance
- Google distance
- Tag Concurrence Distance (Google Image Distance)
- Tag2Image Distance

Tag Concurrence Distance

- Image Tag Concurrence Distance
  - Reflects the frequency of two tags occur in the same images
  - Based on the same idea of NGD so we can also call it "Google Image Distance"



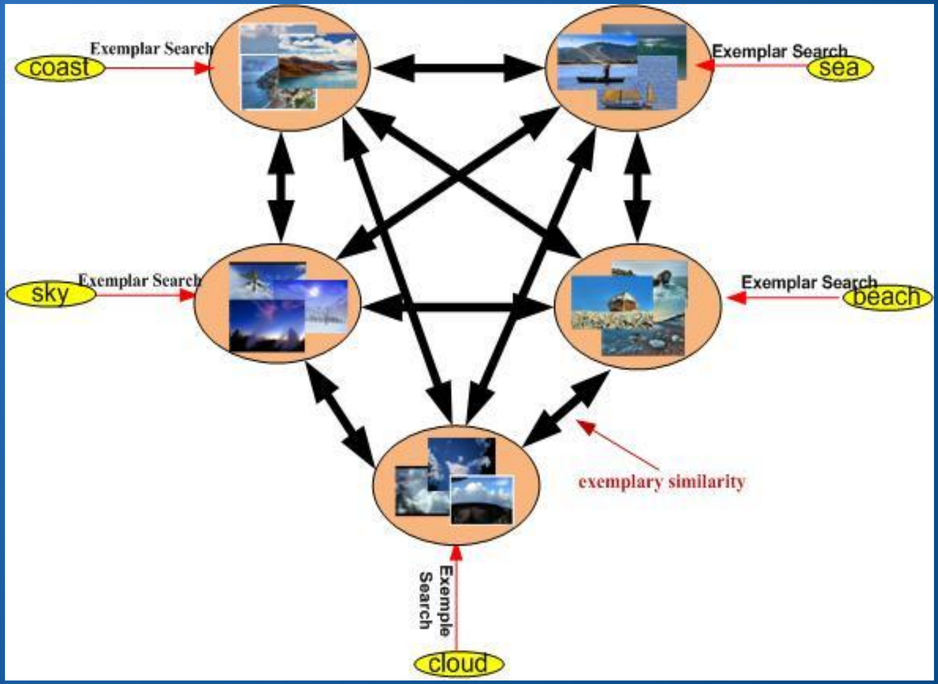
- Pros and Cons
  - Pros: Images are taken into account
  - Cons: Tags are not complete and noisy so visual concurrency is not well reflected. In addition, the distance is image independent

dog , grass, tree, leaf

tree , grass, dog, leaf

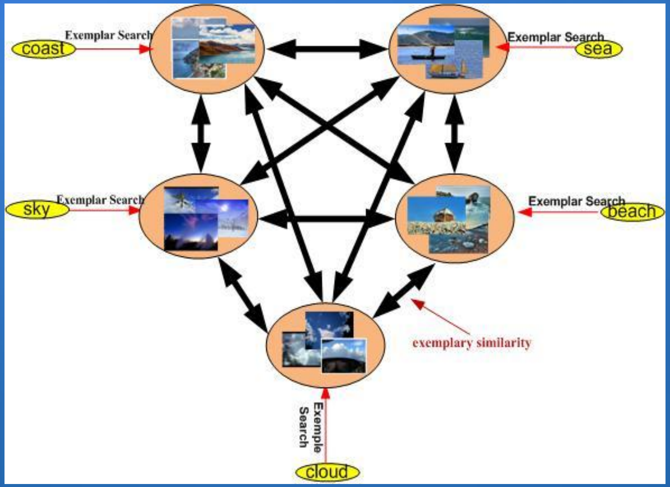
Tag2Image Distance

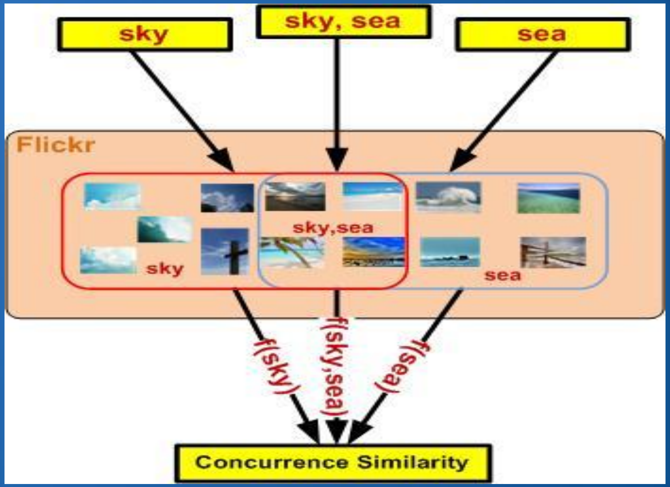
- Tag2Image
  - Find images with a particular tag
  - Keep those close to the target image (finding N-neighborhood)
  - Named as "Tag2Image Set"
- Tag2Image Distance
  - Distance between the corresponding tag2image sets of the two tags
- Pros and Cons
  - **Pros**: Images are taken into account and the distance is image dependent
  - **Cons**: Finding neighbors may be expensive



Random Walk Based Tag Ranking

- Tag Graph Construction
  - Tag2Image similarity & Concurrence similarity





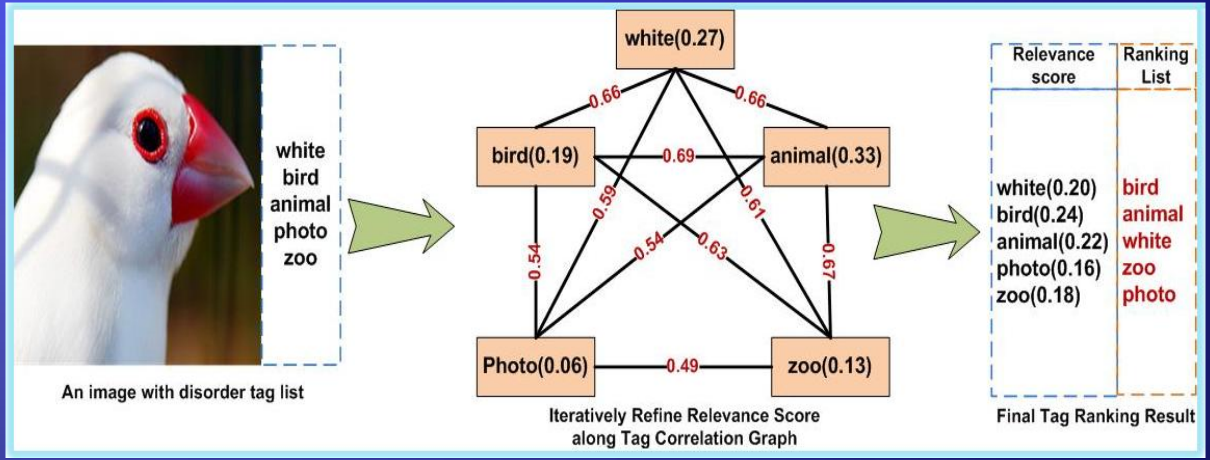
- Combination

$$s_{ij} = s(t_i, t_j) = \lambda \cdot \varphi_e(t_i, t_j) + (1 - \lambda) \cdot \varphi_c(t_i, t_j)$$

Visual similarity   Concurrence similarity

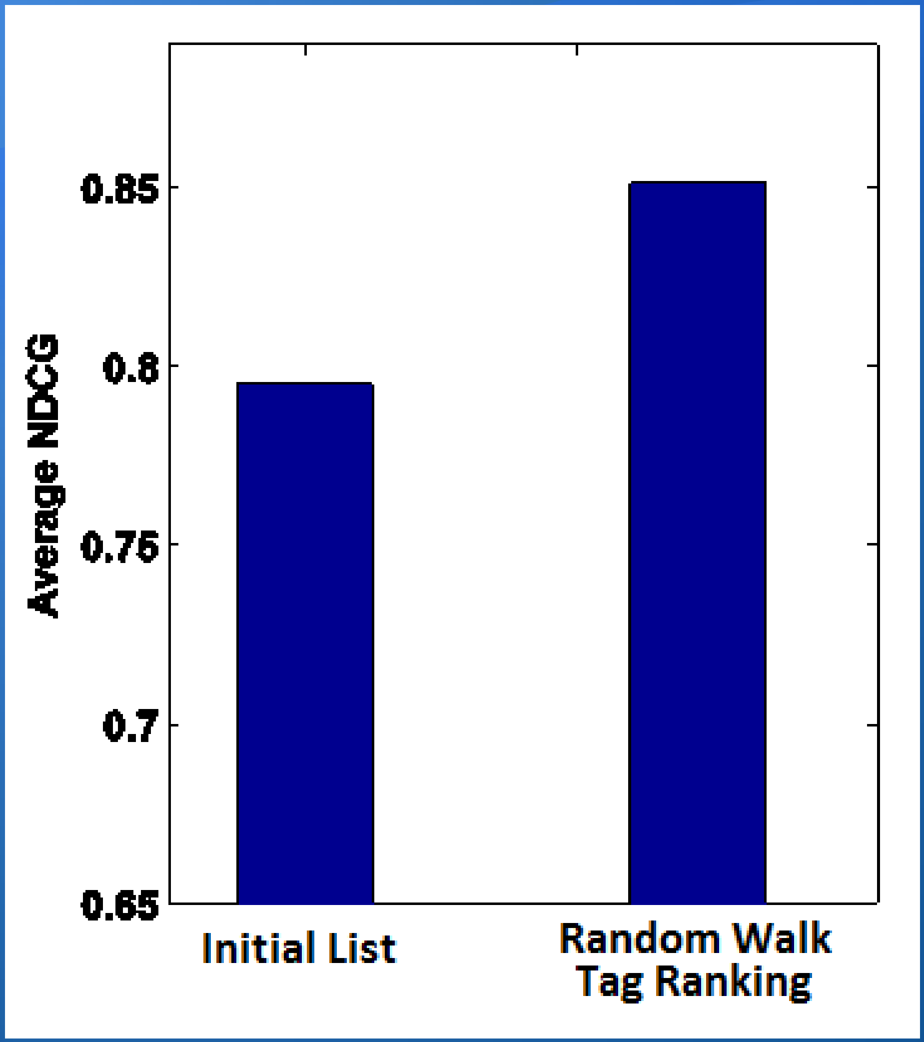
Random Walk Based Tag Ranking

- Random walk over tag graph



$$\mathbf{r}_k = \alpha \mathbf{P} \mathbf{r}_{k-1} + (1-\alpha)\mathbf{v}$$

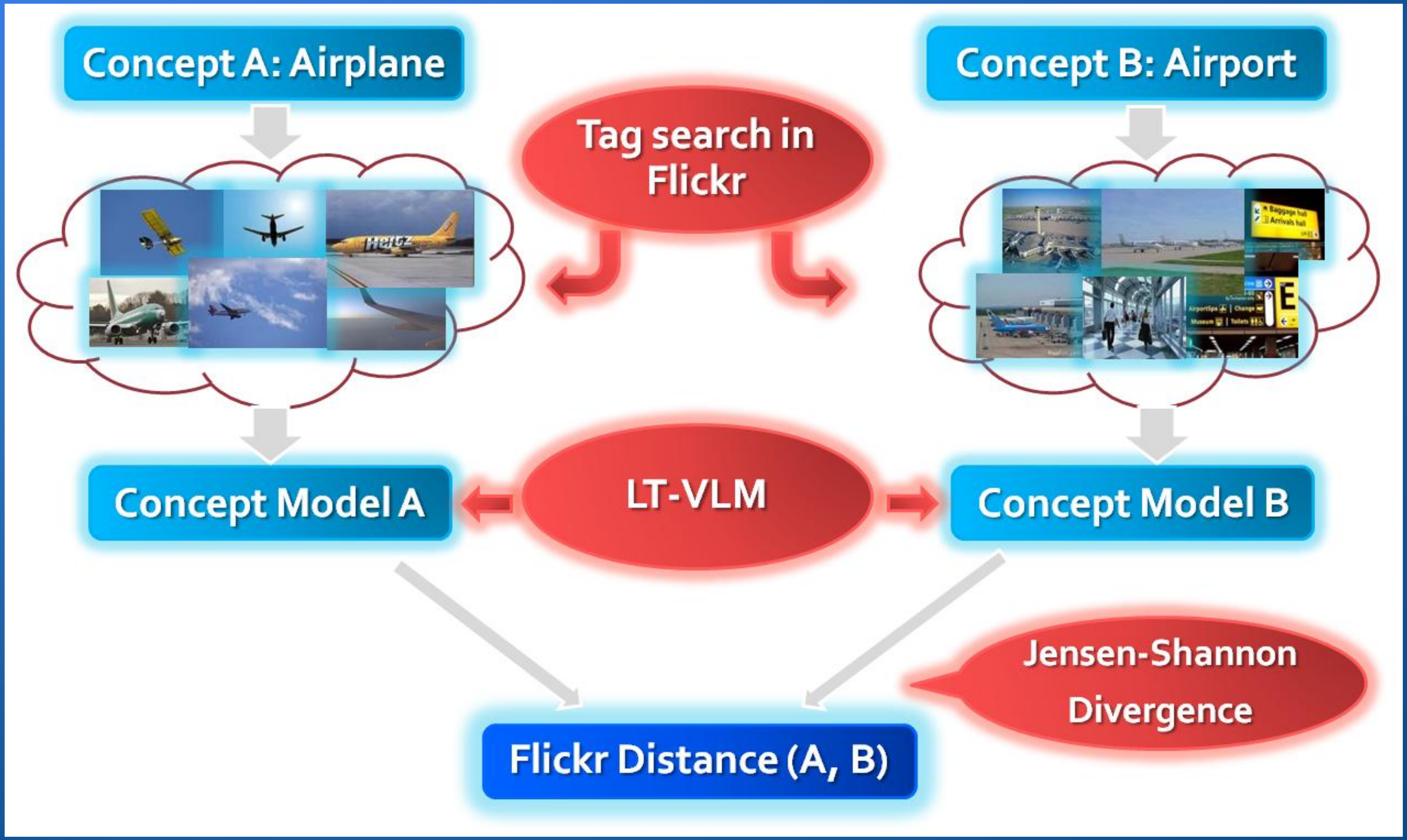
- Transition matrix P denotes the row- normalized matrix of similarity matrix S.
- r is the vector of relevance score for each tag of the image.
- v is the vector of relevance score obtained by initial probabilistic tag relevance estimation.
- Alpha is the weighting parameter.



Average NDCG
0.85
0.8
0.75
0.7
0.65
Initial List
Random Walk
Tag Ranking

A Better Measure: Flickr Distance

- Lei Wu, Xian-Sheng Hua, et al. Flick Distance. *ACM Multimedia 2008* (**ACMMM 2008**). Vancouver, Canada, October 2008. (Best Paper Award Candidate)



Microsoft Research 微软亚洲研究院

Is It Enough?

RECALL

Basic Assumptions

- Image Reranking
  - Large image clusters should be promoted
  - Visual similar images should be ranked closely
  - Initial ranks need to be kept as much as possible

Typically got from text-based ranking

- Tag Ranking
  - Large tag clusters should be promoted
  - Semantically close tags should be ranked closely
  - We don't have initial rank

How can we get it?

Initial Relevance Estimation

- A possible estimation

$$s(t,x) = p(t|x)$$

- A better estimation (normalized by frequency)

$$s(t,x) = p(t|x)/p(t)$$

- After some calculation based on Bayesian Rule

$$s(t,x) = \frac{p(x|t)p(t)}{p(x)p(t)} = \frac{p(x|t)}{p(x)}$$

- It is about a particular image *x*, so *p*(*x*) is a constant, therefore

$$s(t,x) \doteq p(x|t)$$

- What is it now?
  - Density of image *x* in the image space with tag *t*

Initial Rank Estimation

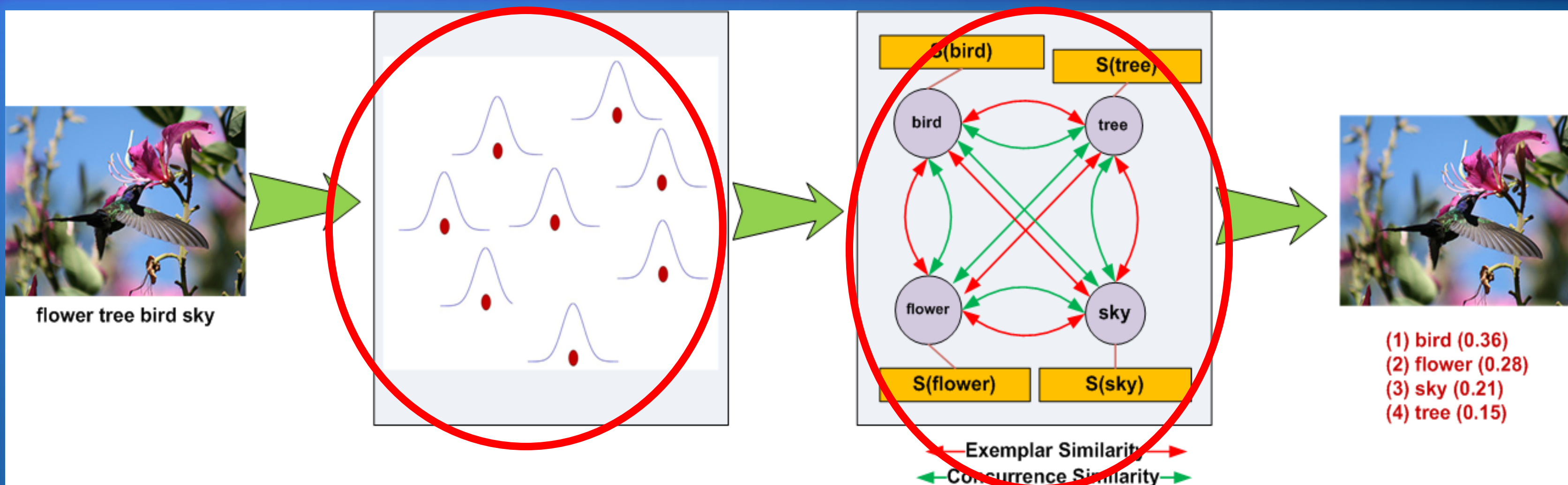
- Can be estimated by Kernel Density Estimation

$$s(t_i, x) = p(x|t_i) = \frac{1}{|X_i|} \sum_{x_k \in X_i} K_\sigma(x - x_k)$$

- An intuitive explanation
  - For image x, X_i can be regarded as x's friends with tag t_i
  - The sum of the similarities estimated based on Gaussian kernel can be regarded as the soft voting from the friends
  - So the initial relevance is actually estimated based on "collective intelligence" from its friend images

In Summary: Tag Ranking

- Two-step strategy

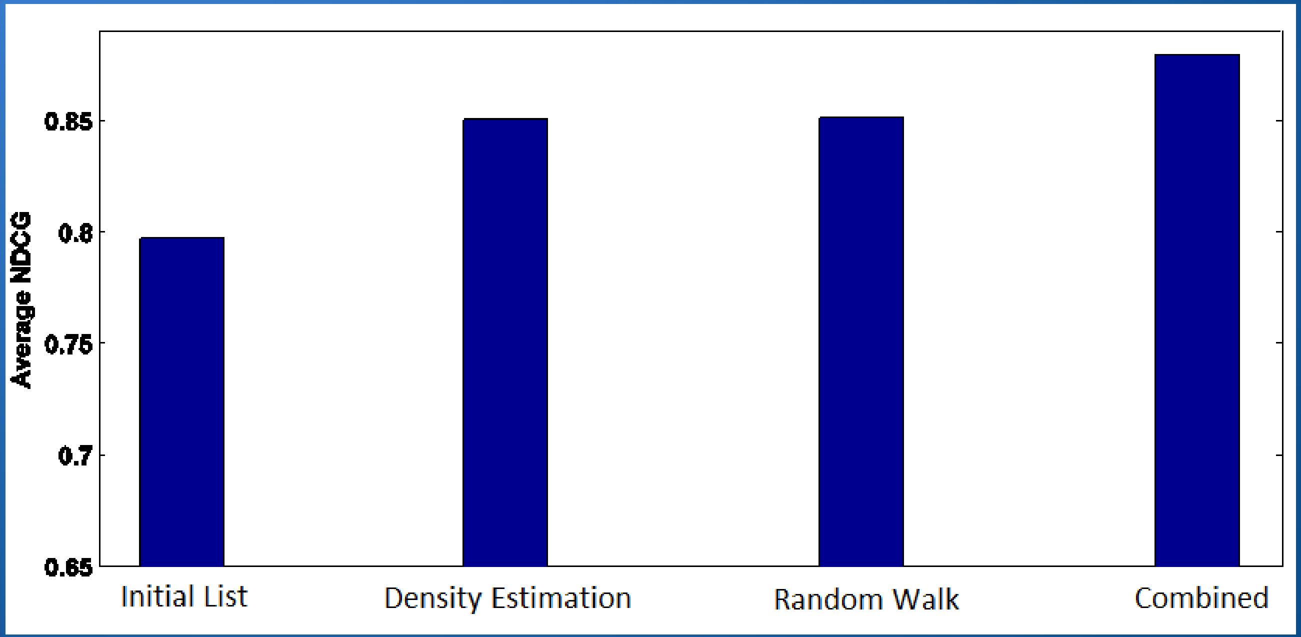


Probabilistic Tag Relevance Estimation

Random Walk Refinement

Performance Evaluation

- In term of average NDCG
  - 50,000 Flickr images (to mine distance and estimate density)
  - 13,330 unique tags
  - 10,000 test images (each was labeled by 5 persons with five levels of relevance)

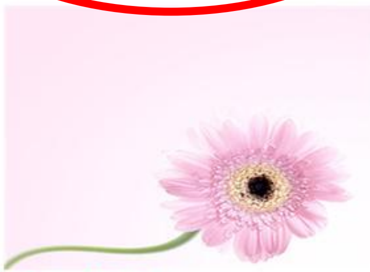


Microsoft Research 微软亚洲研究院

Original Tag List:
blue winter sky white mountain snow photography gold nikon paradise view top greece drama
Ranked Tag List:
mountain sky white snow winter blue nikon photography view paradise gold greece top drama

Original Tag List:
ocean city summer brazil praia beach water architecture fantastic warm paradise desert great playa best resort arena
Ranked Tag List:
beach water ocean summer architecture fantastic paradise great resort playa city brazil best desert praia arena warm

Original Tag List:
blue pakistan portrait green bird nature yellow gold powershot karachi
Ranked Tag List:
bird nature blue green yellow portrait gold powershot pakistan karachi



Original Tag List:
pink light white flower green nature yellow spring flora gerbera
Ranked Tag List:
flower white pink nature light green yellow spring flora gerbera

Original Tag List:
sun sunlight animal cat kitten kitty gata gatto
Ranked Tag List:
cat kitty kitten animal sunlight sun gata gatto

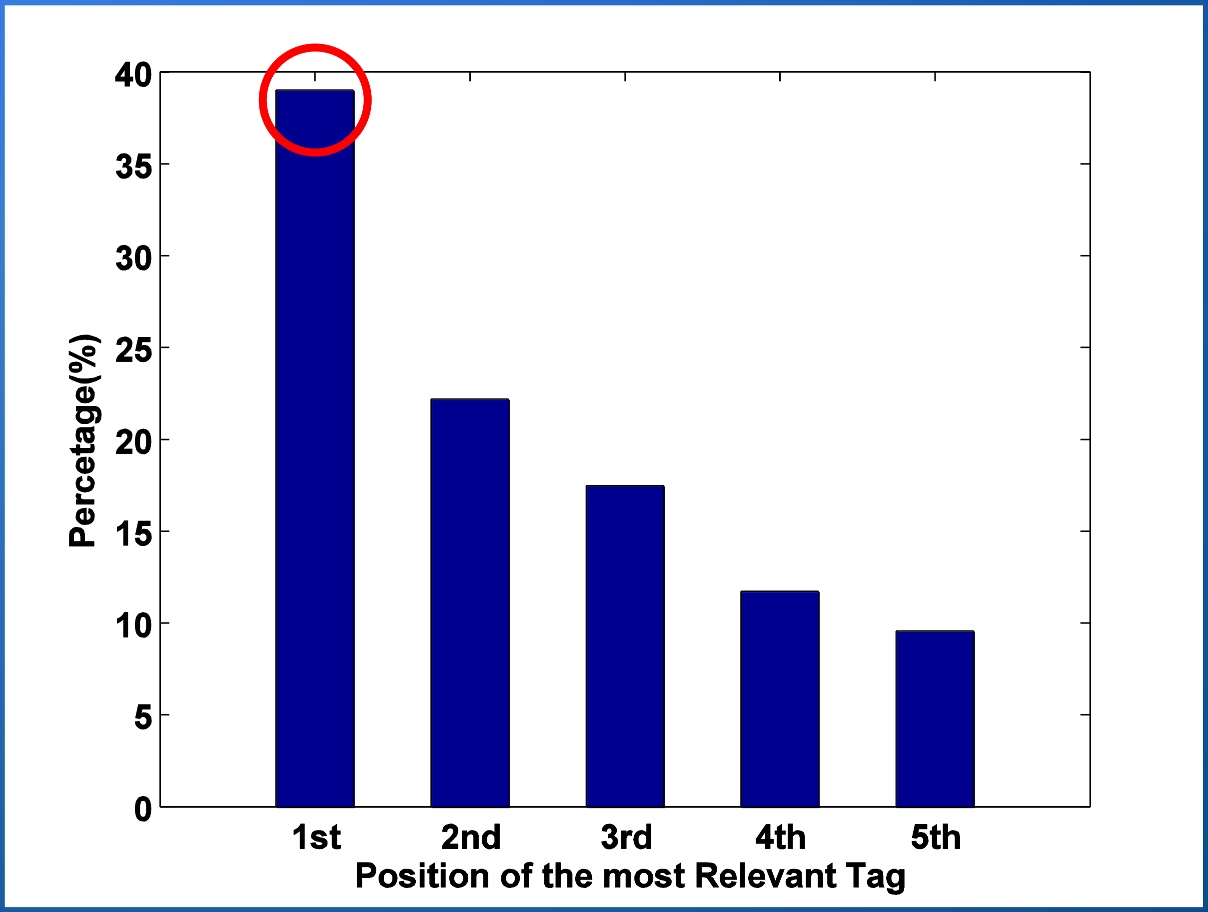
Original Tag List:
family wedding friends sunset red sea love beach silhouette nikon flickr day colours maldives
Ranked Tag List:
sunset sea red beach nikon silhouette maldives love colours flickr friends family day wedding

Original Tag List:
park morning mist holland tree bird water fog duck baum
Ranked Tag List:
tree water bird fog park mist morning duck holland baum

Original Tag List:
ocean travel blue sea water philippines adventure
Ranked Tag List:
sea water ocean blue travel philippines adventure

Original Tag List:
ferrari concept car auto automobile
Ranked Tag List:
automobile car auto ferrari concept

After tag ranking, almost 40% images have their most relevant tag appear at the top position in their tag list.



Application 1: Tag-based search

- Use tag position as relevance measure

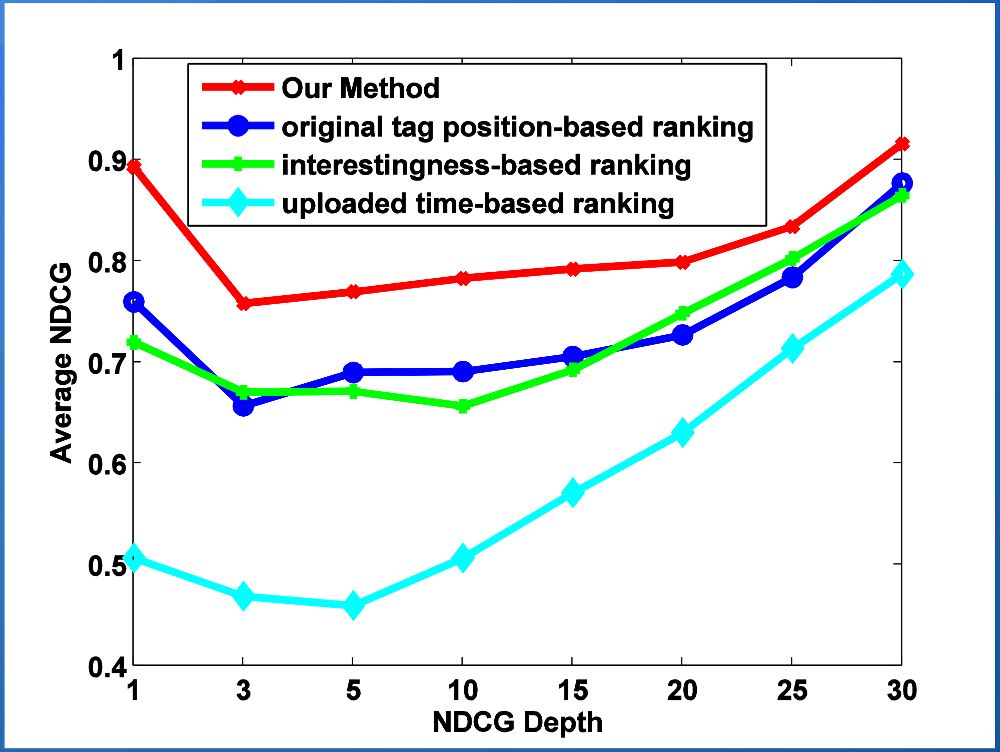
$$r(x_i) = -\tau_i + 1/n_i$$

- Ranking result for query "water"

Application 1: Tag-based search

- Use tag position as relevance measure $r(x_i) = -\tau_i + 1/n_i$
- Ranking result for query "bird"

Performance of Tag-Based Search

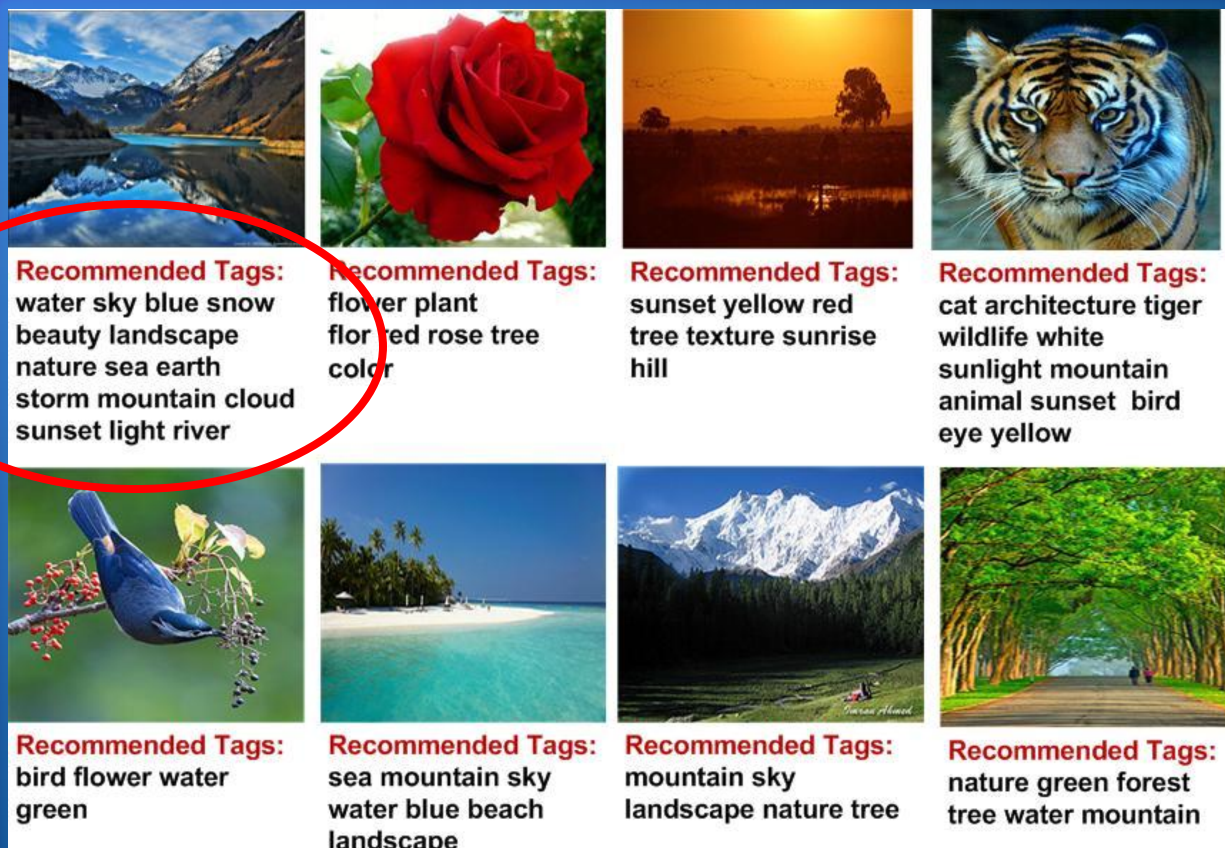


Our tag position-based ranking strategy outperforms all other image ranking strategies on Flickr

Microsoft Research 微软亚洲研究院

Application 2: Auto Tagging

- Use top tags of similar images as tags for a new uploaded image



Performance of Auto Tagging

	Prec@1	Prec@5	Prec@all
Original(Baseline)	0.5858	0.4980	0.4980
Recommendation	0.7255	0.5799	0.5772
Improvement(%)	23.9	16.5	15.9

Using top tags after tag ranking to perform auto tagging

even outperforms human being

Application 3: Group recommendation

- Use the top tags of an image as query keywords to search for its potentially suitable groups.

bird nature wildlife black flight action

Tags	Recommended Groups
bird:	Birds and Wildlife UK | Birds Photos | British Birds
nature:	Nature's Beauty | The World of Nature | Arizona Nature
wildlife:	we love wildlife | California Wildlife | The Wildlife Photography

Performance of Group Recommendation

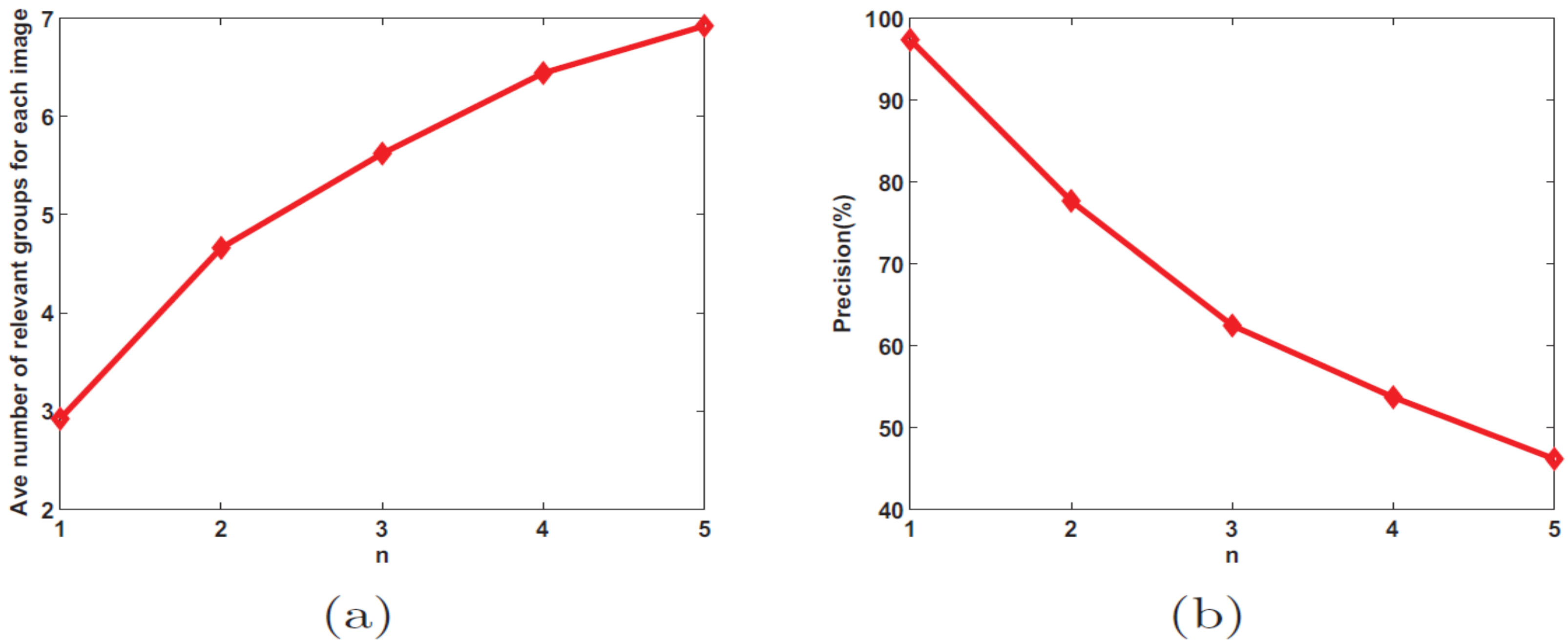


Figure 18: Performance of group recommendation with different n. (a) illustrates the average numbers of relevant recommended groups and (b) illustrates the recommendation precisions.

Tag ranking based group recommendation can help users better share their media content

Conclusion

- Initial tags are orderless in term of the relevance which limits the performance of tag-based search and other applications based on tags
- We propose a tag ranking strategy to solve this problem:
  - Density estimation to obtain initial rank scores
  - Refined by random walk based on image-dependant tag graph
- Tag ranking benefits a series of tag-based applications on social media websites

Future Work

- Extending it to surrounding text
- Extending it to videos
- Extending it to textual documents
- Using more sophisticated distance measures

Thank You

Backup Slides

Benefit of ranking tags

- If tags associated with social images can be ordered according to their relevance to the image, many tag based applications can be enhanced., e.g.
  - Tag based image search
    - Ranking images according to tag position
  - Auto tagging
    - Annotate top ranked tags to new uploaded images
  - Group recommendation
    - Use the top ranked tags of an image to search for suitable groups to recommend to the user

1st step: Probabilistic Tag Relevance Estimation

- We define the relevance score of a tag *t* with respect to an images *x* as

$$s(t,x) = p(t|x)/p(t)$$

- We further derive the formula based on Bayes' rule.

$$s(t,x) = \frac{p(x|t)p(t)}{p(x)p(t)} = \frac{p(x|t)}{p(x)}$$

- Thus the relevance of *t* with respect to image *x* becomes

$$s(t,x) \doteq p(x|t)$$

- Kernel Density Estimation is then adopted to obtain this score.

1st step: Probabilistic Tag Relevance Estimation (cont.)

- The relevance score of tag t_i with respect to image x can be estimated by KDE as follows.

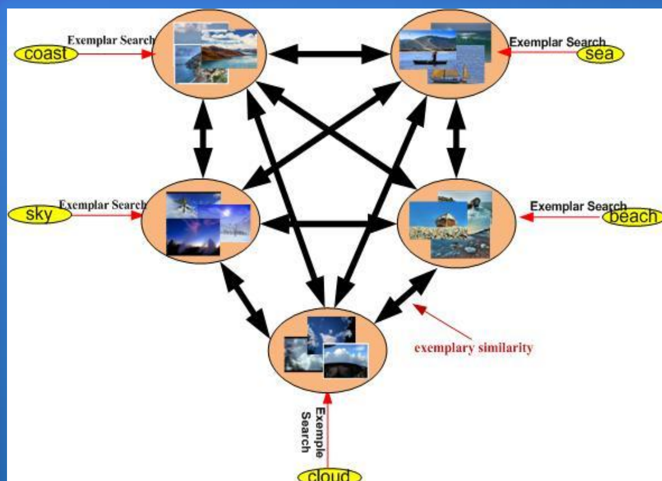
$$s(t_i, x) = p(x|t_i) = \frac{1}{|X_i|} \sum_{x_k \in X_i} K_\sigma(x - x_k)$$

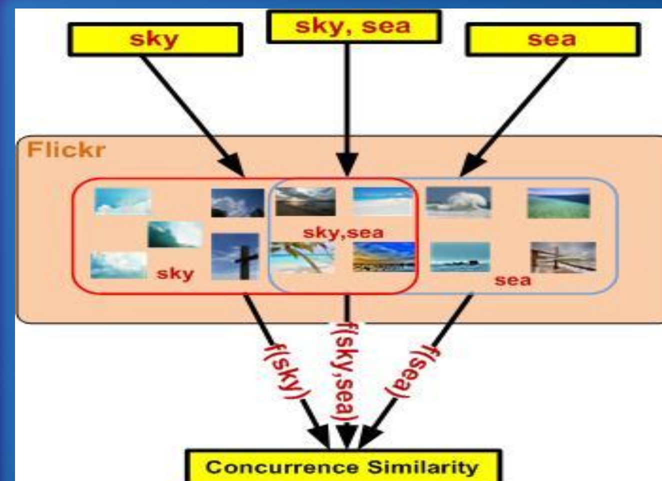
where $|X_i|$ the cardinality of the set of images tagged with tag t_i, and K_σ the Gaussian kernel function with radius parameter σ e.,

$$K_\sigma(x - x_k) = exp(-\frac{||x - x_k||^2}{\sigma^2})$$

2nd step: Random walk based refinement

- Tag Graph Construction
  - Exemplar similarity & Concurrence similarity





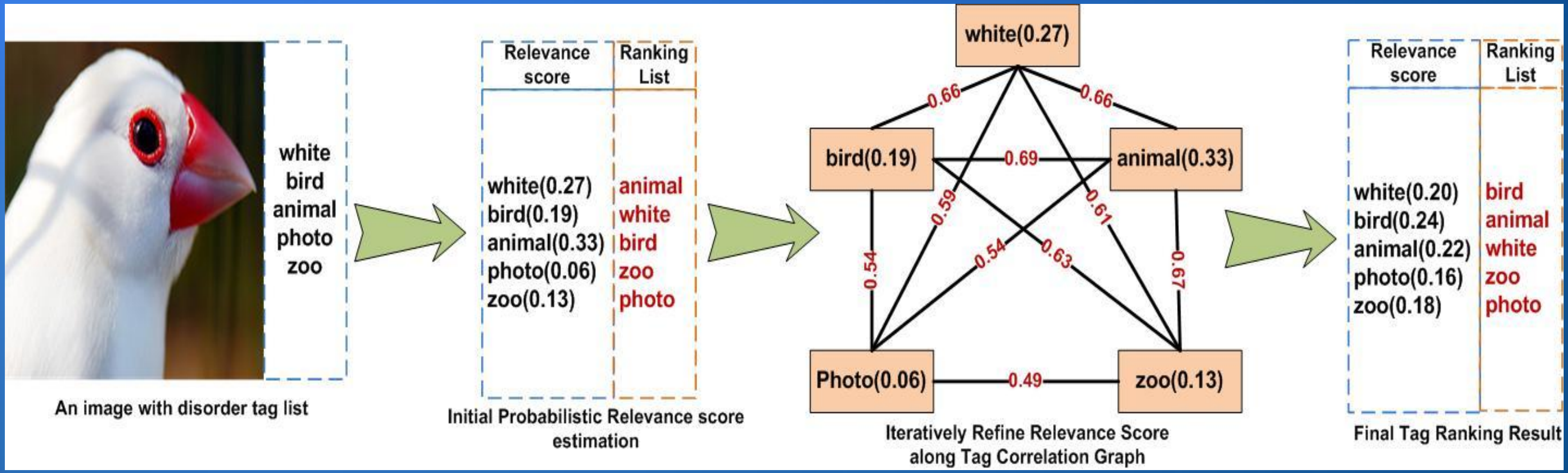
- Combination

$$s_{ij} = s(t_i, t_j) = \lambda \cdot \varphi_e(t_i, t_j) + (1 - \lambda) \cdot \varphi_c(t_i, t_j)$$

Visual similarity Concurrence similarity

2^{nd} step: Random walk based refinement (cont)

- Random walk over tag graph



$$\mathbf{r}_k = \alpha \mathbf{P} \mathbf{r}_{k-1} + (1-\alpha)\mathbf{v}$$

- Transition matrix P denotes the row- normalized matrix of similarity matrix S.
- r is the vector of relevance score for each tag of the image.
- v is the vector of relevance score obtained by initial probabilistic tag relevance estimation.
- Alpha is the weighting parameter.