Image Region Entropy: A Measure of "Visualness" of Web Images Associated with One Concept

Keiji Yanai and Kobus Barnard
Univ. of Electro-Communications, JAPAN / Univ. of Arizona, USA

C-07

1. Objective

Select "words" for image-annotation automatically

We have a lot of words, but not all words are appropriate for image-annotation. We should use "visual" words.

"Visual" : being associated with visual properties of images

Measure "visualness" of word concepts

We propose a new measure: **"image region entropy"**.

Based on it, we can select words for image-annotation.

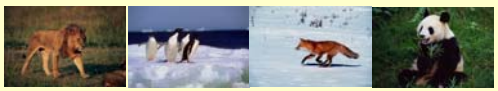
2. Motivation

Corel ID 108041

tiger feline cat mammal animal wildlife grass forest

Corel keywords *[Corel image gallery 1,000,000]*

- "Mammal" is categorized based on the way of their birth, not based on their appearance.
- "Animal" is too broad. There are many kinds of animals.



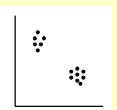
- "Mammal" and "animal" should be inferred from "tiger" rather than be recognized from images directly.

Some words are not appropriate for image recognition. Words related to "visual properties" are suitable for that.

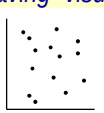
We need to find "visual" words for image-annotation.

3. Image region entropy

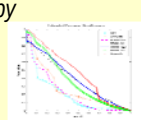
- Measure "visualness" of word concepts
- Represent the property of the distribution of image features
  - Biased / uneven: *having "visualness"* ⇨ **low entropy**
  - Random/uniform: *not having "visualness"* ⇨ **high entropy**



biased
low entropy
visual

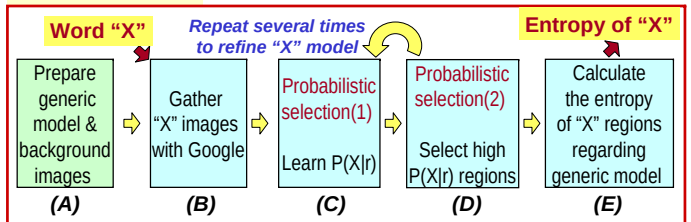


random
high entropy
not visual



- Need no ground truth data unlike r-p diagram
- Use **images on the Web with Google**
  - This enables us examine about any words
- To make "entropy" meaningful, pick up only "X" regions, excluding backgrounds with a probabilistic method.

4. Method



Processing flow to obtain "image region entropy"

A. Prepare the generic model and negative images

- To get generic distribution of region features,
  - Collect a lot of images from the Web randomly.
  - Segment all the collected images and extract region features
  - Model generic distribution of region features with GMM using EM ⇨ **generic base model**
- Use these images also as negative sample images (background images) for selecting "X" regions

Random Web images for building the generic model

B. Gather Web images associated with a given keyword "X" with Google Image Search

- Segment all the images and extract region features.

C. Use EM to learn which regions should be associated with the concept "X". (Learn P(X|r))

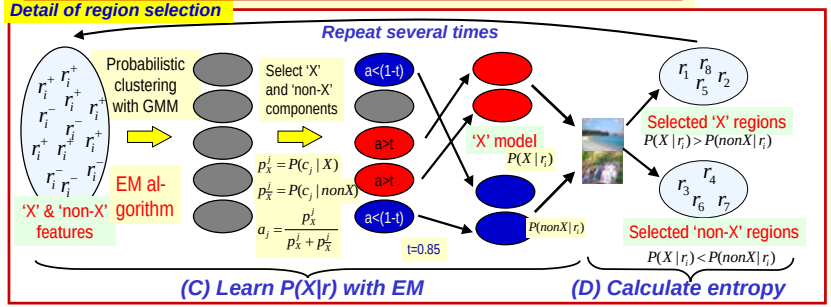
D. Select regions with high P(X|r) as "X" regions

- Base on the modified image-annotation-model [ICCV 01]

E. Calculate the entropy of "X" regions regarding **the generic model.** $E(X) = \sum_{j=1}^{N} -P(X \mid c_j)\log_2 P(X \mid c_j)$

Entropy : how much the distribution of region features is biased compared to the generic distribution of region features

Detail of region selection



5. Experiments

Case study: Finding "visual" adjectives

- Collect 250 images per word for 150 adjectives selected from frequently words of keywords of Hemera image database
  - Color adj.: purple, black, red, blue, white, pink,..
  - Emotional adj.: sad, happy, scary, romantic,…
  - Physical adj.: heavy, light, solid, sharp, old, electric, oval,…
  - Many other kinds of adj.: floral, international, safe, digital,….
- Segmentation by JSEG
- Region features: color (avg. RGB), texture (Gaussian filter), shape

- Our model detects regions related to the concepts of the given word **with no prior knowledge**

"yellow" regions after one iteration

"yellow" regions after five iterations

- **Results for 150 adjectives**

rank	adjective.	entropy
1	dark	0.0118
2	senior	0.0166
3	beautiful	0.0178
4	visual	0.0222
5	rusted	0.0254
6	musical	0.0321
7	purple	0.0412
8	black	0.0443
9	ancient	0.0593
10	cute	0.0607
11	shiny	0.0643
12	scary	0.0653
13	professional	0.0785
14	stationary	0.1201
15	electric	0.1411

rank	adjective.	entropy
135	female	2.4986
136	medical	2.5246
137	assorted	2.5279
138	large	2.5488
139	playful	2.5541
140	acoustic	2.5627
141	elderly	2.5677
142	angry	2.5942
143	sexy	2.6015
144	open	2.6122
145	religious	2.7242
146	dry	2.8531
147	male	2.8835
148	patriotic	3.0840
149	vintage	3.1296
150	mature	3.2265

Results of color adj.

7	purple	0.0412
8	black	0.0443
36	red	0.9762
39	blue	1.1289
46	yellow	1.2827

- Color adjectives which are expected to be "visual " are ranked in the relatively upper ranking.
- "Scary" images have black no-textured regions
  - ⇨ **"visual" word**
- "Famous" images have no prominent tendency.
  - ⇨ **"non-visual" word**

Low entropy: "scary"

High entropy: "famous"

6. Conclusion

- Proposed "image region entropy" to measure "visualness" of words, and applied it to150 adjectives.
- For future work, we will improve and apply it to nouns.
- Incorporate "visual" adjectives with our image-annotation models in addition to nouns.
- Select "visual" words from all nouns and build "generic image corpus" by the Web image-gathering.