UNT at ImageCLEF 2011: Relevance Models and Salient Semantic Analysis for Image Retrieval

Miguel E. Ruiz**
Chee Wee Leong
and Samer Hassan

University of North Texas,
Denton, Texas, USA

**Currently at Nuance Communications Inc.

Goals

- Test corpus based methods to do semantic query expansion
- Identify terms that are more likely to describe an image (“picturable”)
- Explore relevance models and structured queries

Salient Semantic Analysis

- SSA (Hassan & Mihalcea, 2011)finds semantic similarity between words based on salient content links from a corpus such as Wikipedia.
  - "Plants are Life living organisms belonging to the kingdom Plantae. Precise definitions of the kingdom vary, but as the term is used here, plants include familiar organisms such as trees, flowers, herbs, bushes, grasses, vines, ferns, mosses, and green algae."
  - The term "plants" is represented by a weighted vector:

< living organisms 0.85, kingdom 0.6, trees 0.85, flowers 0.85, herbs 0.7, bushes 0.7, grasses 0.8, vines 0.6, ferns 0.8, mosses 0.8, green algae 0.7>

Salient Semantic Analysis

- Formally, given
  - a corpus **C** with m tokens
  - vocabulary of size **N**
  - concept size **W** (no. of unique Wikipedia concepts)
  - A **N** x **W** term-by-concept matrix (**P**) is generated, where

$$P_{ij} = log_2 \frac{f^k(w_i, c_j) \times m}{f^C(w_i) \times f^C(c_j)}$$

  - To calculate the semantic relatedness between two words/texts, A and B,

$$Sim(A,B) = \begin{cases} 1 & Score_{cos}(A,B) > \lambda \\ Score_{cos}(A,B)/\lambda & Score_{cos}(A,B) \leq \lambda \end{cases}$$

where

$$Score_{cos}(A,B) = \frac{\sum_{y=1}^{N}(P_{iy} * P_{jy})^{\gamma}}{\sqrt{\sum_{y=1}^{N} P_{iy}^{2\gamma} * \sum_{y=1}^{N} P_{jy}^{2\gamma}}},$$

γ = weight bias
λ= normalization factor

Salient Semantic Analysis

- Semantic association between two terms or between two pieces of text SSA uses
- Compute a similarity value based on the co-occurrence with in a window of size k in a given corpus (Wikipedia).
- The similarity value is controlled by a parameter (λ)
  - λ is a threshold of the semantic gap between terms that are perfect synonyms (e.g., tiger-tiger) and near synonyms (e.g., tiger-feline) .

Flickr Picturability

- Flickr Picturability (Leong et al., 2001)Based on rewarding terms that match tags assigned to images in Flickr.
- Method:
  - build a corpus with the top Flickr tags most related to the query
  - weights them according to the co-occurrence of the term in the contexts of other query terms

Collection Preparation

- Translated captions and Wikipedia text from French and German into English using Bing.

- Used Indri to index collection using a language model representation:
  - Unigram model
  - Krovetz stemmer
  - Stopwords

Query Expansion

- Step 1: Generate list of candidates
  - Find candidates using SSA
  - Compute Flickr picturability score
  - Compute weights using the formula
    - Weight(wi) = tf (wi) * 1/rank(wi) ∗ flickr(wi)
- Step 2: Ensure semantic focus
  - Select terms with SSA similarity above a threshold (Sim(Q,w) ≥ α)

Retrieval Model

- Inspired on Lavrenko's Relevance models (Pseudo-relevance feedback for Language modes)
- Weighted_query = β Q_original + (1-β) Q_expansion

- We used β=0.5

Table 1. Official results in the Wikipedia Image Retrieval task.

Results

Run name	FB/QE	MAP	P10	P20	Rprec	Bpref
Baseline (unofficial)		0.2621	0.5493	0.4434	0.2900	0.2522
2011_SSA50	QE	0.2143	0.3260	0.2900	0.2438	0.2027
UNTESU_SSA150rf	QEFB	0.2292	0.3120	0.2810	0.2476	0.2050
2011_SSA50_FB	FB	0.2327	0.3160	0.2860	0.2543	0.2113
UNTESU_SSA150W	QE	0.2577	0.4060	0.3510	0.2835	0.2401
UNTESU_SSA50Wrf	QEFB	0.2794	0.4240	0.3630	0.3107	0.2647
UNTESU_SSA150Wrf	FB	**0.2820**	0.4200	0.3610	0.3190	0.2679
UNTESU_BLRF	FB	**0.2866**	0.4220	0.3650	0.3276	0.2821

Conclussions

- The proposed SSA and Flicker picturability prove to be quite good for expanding hard queries.
- Weighted structured queries and relevance feedback are quite important for retrieval performance in query expansion.

- To do:
  - Incorporate word sense disambiguation.
  - Better handling of phrases and compound terms(e.g. “close up”)

Acknowledgement

- Thanks to
  - Language Information Technology group at UNT (Rada Mihalcea)
  - Nuance Communication for supporting my attendance to CLEF.