

1. Goal:

The project is going to build an on-line image search system for demo in Grand Challenge of ACM MM 2009 conference. The most important characteristic of this image search system is image search result clustering.

2. Approach:

We propose a three-stage process for image search result clustering.

- Traditional text-based image retrieval and improving research result with re-ranking paradigm
- Canonical image selection from re-ranked search result as cluster center of image clustering
- Grouping images to canonical images and constructing final image clusters based on image linking graph

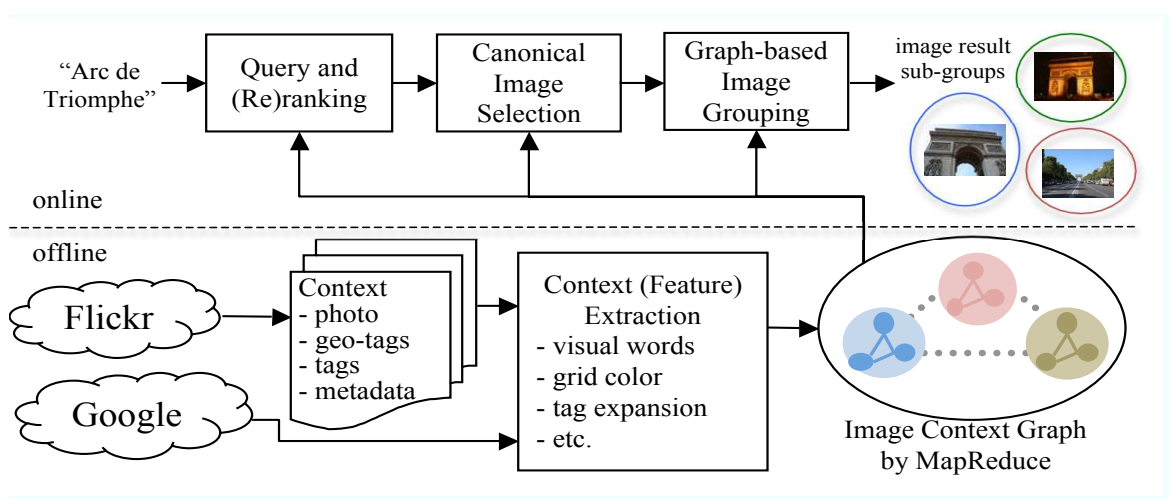


Figure 1. The system architecture of proposed image search clustering.

3. System components:

No.	Component	Description & details
1	Image dataset and features	Flickr550 dataset
		Textual features <ul style="list-style-type: none"> • Title • Tag • Description • Comment Visual features <ul style="list-style-type: none"> • GCM • Gabor • Visual Word
2	Text-based image retrieval	Supported by MySQL database storing text description of images
		On-line searching image database for text query from searcher
		The component for Flickr550 dataset is ready
3	Search result re-ranking	Implementation of ordinal reranking algorithm
		On-line re-ranking text-based search result
4	Canonical image selection	Implementation of canonG algorithm
		On-line selecting canonical images from re-ranked search result
		Off-line pre-computed items: <ul style="list-style-type: none"> • Frequency histogram (FH) • Coverage vector (CV)

No.	Component	Description & details
5	Graph-based image grouping	Nearest-neighbor assignment of cluster
		On-line grouping images in search result to canonical images to form image clusters
		Offline pre-computed items: <ul style="list-style-type: none"> • Image graph computation powered by Hadoop technology
6	Image search result interface	It could be: <ul style="list-style-type: none"> • Animated search result visualization • Static web interface
		On-line representing image search result to searcher

4. Action items:

Task	Detailed items
Project preparation	Stakeholder assignment
	Sketching schedule
Data preparation	Image database
	Image feature
	FH and CV for 3.4
	Image linking graph data
System specification	System architecture
	Data schema
	Component APIs
	System interfaces