

# NWU Landscape Visual Search

Zhu Li

## 1 Introduction



Figure 1 Screen shot of TinEye

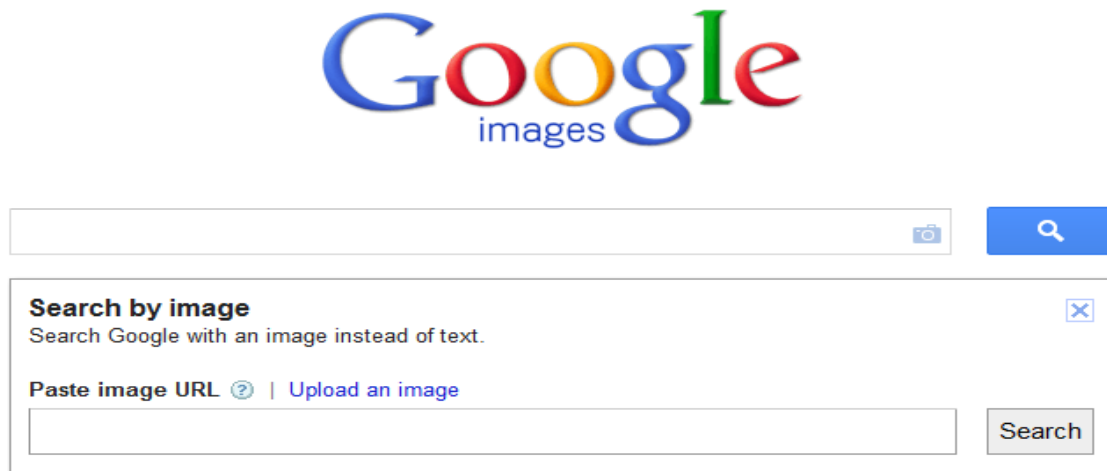


Figure 2 Google Search by uploading

Query by capture has been getting more and more popular since the invention of smart phones equipped with high resolution cameras. Some deployments include Google query by capture and TinEye. Google query by image supports searching captured logos and texts. TinEye supports web search of contents related to queries that is uploaded. The idea of these applications is to support searching pictures to find matches and corresponding contents. Building such a content based visual search system includes two steps: 1) collecting and building an image databases and corresponding contents; 2) search platform supporting live index, searching and uploading new pictures and annotations.

Visual query by capture are made possible with reliable and repeatable visual features (SIFT[1], SURF[2]). N SIFT features are extracted from a query image q. Searching the query image q are actually matching each feature of the query to the databases of image features. Some images in the databases are hit by the feature in query image q and ranked according to the hit count. This process is similar to text search which is counting the word count.

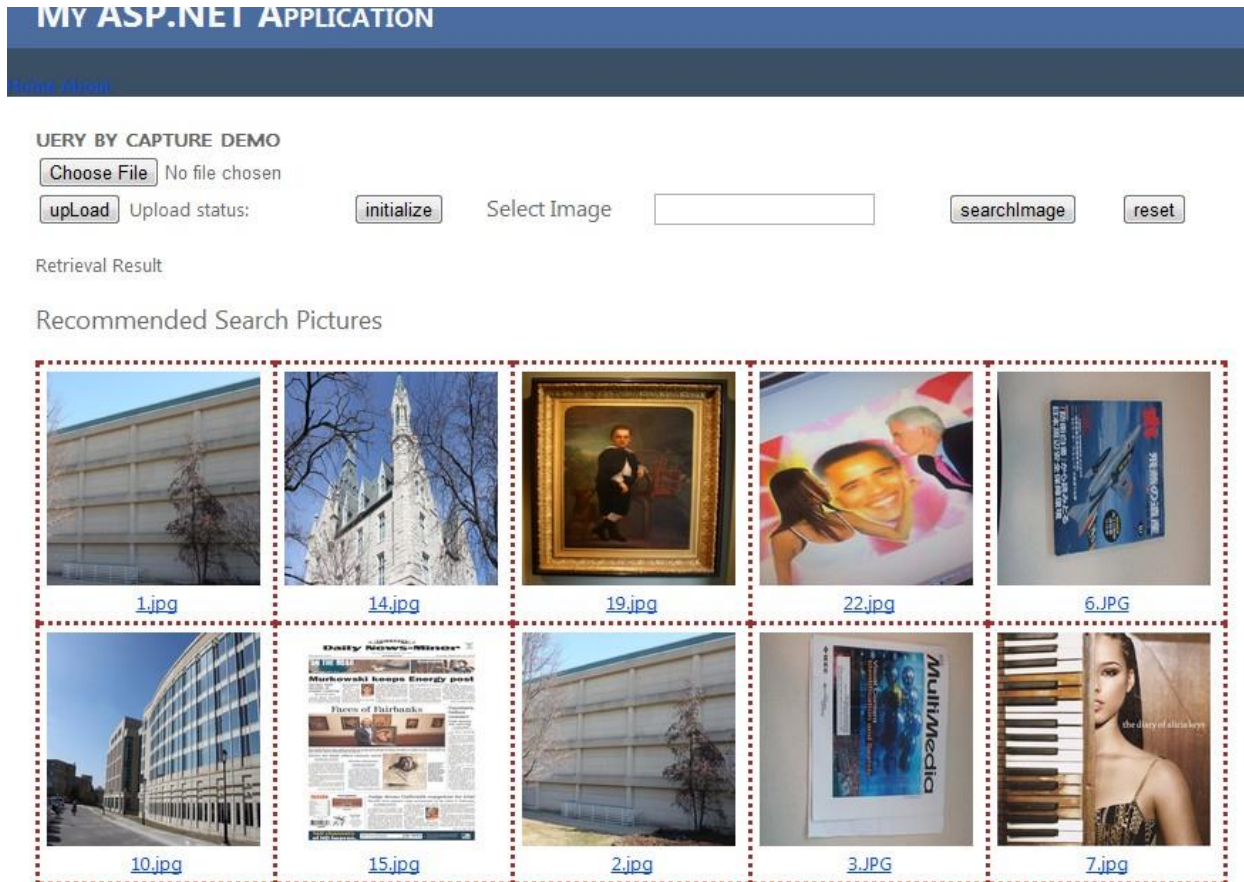


Figure 3 Asp.net Visual Search

In this project, we plan to build a Northwestern University landscape pictures & annotations database and a web-based interface to support landscape picture & annotation upload, live index and search by uploading picture. An existing demo is built with Asp.Net as shown in Figure 3. Some landscape pictures and annotations are available as shown in Figure 4. More pictures and annotations are needed. We can either take pictures by ourselves or sending emails to northwestern university students asking for more participation.

## 2 Technical Validations

Visual search architecture would include the following parts: 1) Feature extraction[3,4,5]; 2) Live indexing of features [6,7,8,9,10,11]; 3) Search algorithm[12,13,14,15,16,17].

These modules are available in Matlab now and we need to convert these codes to Linux executable versions. SIFT features are used as features because SIFT feature is highly discriminative and repeatable among similar images. VL-Feat is used as the SIFT extraction software. Indexing of SIFT features is implemented with kd-tree[3]. Searching features of query images are searching the kd-tree and locating the leaf node and hit count are used as indicator of a match.

We need a web interface to present this application, which supports search landscape images and annotations with mobile phone captured images. Ideally, this web interface is implemented with Php/MySQL [18,19] and served on Apache server in Linux environment. This interface has the following functionality: 1) uploading images and annotations; 2) calling indexing functions and searching functions; 3) presenting results in friendly user interface.

### 3 Collecting Landscape Images

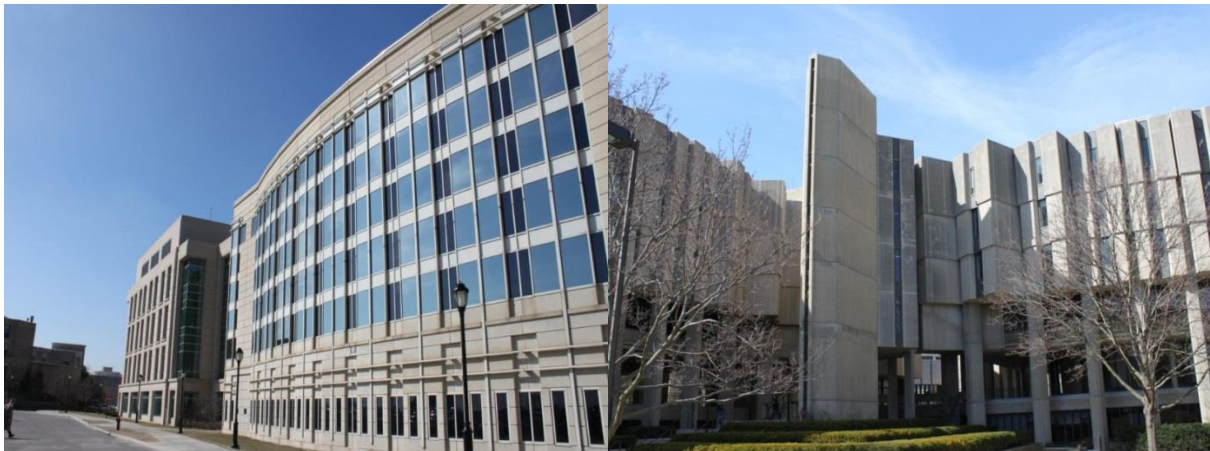


Figure 4 Samples images at Northwestern University campus

Each landscape like library, Gym is organized into a folder containing images and an annotation file containing contents related with the landscapes. I have some of the pictures at hand now and still need more pictures. We can either take more pictures by ourselves or ask students to participate with emails. Images should be taken at resolution of 1024 \* 768. Each building or landscape should be taken from different angles and multiple pictures should be taken for each angle.

[1] Z. Ye, X. Chen, and Z. Li. "Video based mobile location search with large set of SIFT points in cloud", *Proceedings of the 2010 ACM multimedia workshop on Mobile cloud media computing (MCMC '10)*, ACM, New York, USA.

[2] D.G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, 60, pp. 91-110, 2004.

- [3] H. Bay, A. Ess, T. Tuytelaars, L.V. Gool, "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding (CVIU)*, pp. 346--359, 2008
- [4] J. Sivic and A. Zisserman, "Video Google: A Text Retrieval Approach to Object Matching in Videos", *Proceedings of the International Conference on Computer Vision*, 2003.
- [5] X. Xin and A.K. Katsaggelos, "A Novel Image Retrieval Framework Exploring Inter Cluster Distance", *Int. Conf. Image Processing*, September 2010.
- [6] Y. Ke and R. Sukthakar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors", *Computer Vision and Pattern Recognition*, 2004
- [7] B. Girod, V. Chandrasekhar, R. Grzeszczuk, Y. Reznik, "Mobile Visual Search: Architectures, Technologies, and Emerging MPEG Standard", *IEEE Multimedia*.
- [8] M. Iliadis, S. Yoo, X. Xin and A. K. Katsaggelos. "Virtual touring: A content based image retrieval application". In *Multimedia and Expo Workshops (ICMEW)*, 2013 IEEE International Conference on (2013), pp. 1-4.
- [9] Z. Li, X. Xin, and A. K. Katsaggelos. Image topological coding for visual search. In *US Patent* (2012).
- [10] J. Springer, X. Xin Z. Li, J. Watt and A. K. Katsaggelos. "Forest hashing: Expediting large scale image retrieval". In *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on (2013), pp. 1681-1684.
- [11] TJ. Liu, H. Han, X. Xin, Z. Li, and A. K. Katsaggelos. "A robust and lightweight feature system for video fingerprinting". In *Proc. of EUSIPCO* (2010).
- [16] Xin, X., Li, Z., and A.K.Katsaggelos. System and method for randomized point set geometry verification for image identification. In *US Patent* (2013).
- [17] Xin, X., Li, Z., Ma, Z., and Katsaggelos, A. K. Spectral approximation to point set similarity metric. In *Multimedia and Expo Workshops (ICMEW)*, 2013 IEEE International Conference on (2013), pp. 1-4.
- [18] Xin, X., Nagar, A., Srivastava, G., Li, Z., Fernandes, F., and Katsaggelos., A. K. Large visual repository search with hash collision design optimization. *MultiMedia, IEEE 20, 2* (2013), 62-71.
- [19] X. Xin, Li, Z., and A.K.Katsaggelos. Laplacian embedding and key points topology verification for large scale mobile visual identification. *Signal Processing: Image Communication* 28, 4 (2013), 323-333.