THE AUSTRALIAN NATIONAL UNIVERSITY

FACULTY OF ENGINEERING AND INFORMATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE

Project Plan

Structured Object Recognition for Content Based Image Retrieval

Supervisors: Dr. Antonio Robles Kelly

Dr. Jun Zhou

Student: Ali Zia (u4475570)

1. Background

Due to the success of digital cameras both in technology and in the market, a great number of digital images can be produced in no-time. As nowadays we have very effective text search engines available, such as Google and Yahoo which can also used to search for images based on text-matching, such as name or the content of the page the image resides in, people are now more interesting in searching images by using a query image only without the need of providing further information about the image.

For purposes of content-based image retrieval, a number of methods have been proposed so as to achieve robust and efficient systems. These methods often represent images as a bag of features. A query image is then matched with images in the database by computing the distances between images. As a result, in general, object and image retrieval and classification techniques [13,19,2] are based upon the summarization of the image dataset using a codebook of visual words [5,17,12] , which are used to query the dataset so as to retrieve images that best match the query. When a query image is provided by the user, the features in the image are compared with those on the codebook. Then a measure of similarity between the images in the dataset is computed so as to retrieve the closest match.

As a result, the design of architecture for image retrieval requires both, an image representation suitable for search and a similarity measure that can be employed to rank the images with respect to the relevance to the query [24]. The main challenges in existing algorithms remain efficiency (in terms of speed and memory consumption), accuracy and simplicity. By efficiency we mean how quickly the result can be retrieved and how computationally costly is the image representation used. By accuracy we refer to the correctness of the images retrieved by the system provided a query image. Simplicity applies to the degree of ease of deployment and use of the algorithm.

2. Problem/Task description

Problem:

The design of architecture for image retrieval requires both, an image representation suitable for search and a similarity measure that can be employed to rank the images with respect to the relevance to the query. In this work, we will examine the use of a structured learning approach to the image classification problem. In this manner, we aim at investigating the inclusion of structural and feature information in the image into a representation suitable for content-based image retrieval. The main aim is to achieve 50 to 70 percent accuracy in newly designed Algorithm.

Main Tasks

- Finding appropriate Features that can be used to recognize objects.
- Understanding and integrating Saliency detection in system to identify objects.
- Finding appropriate Representation of Objects so that each image can be manipulated efficiently and that the system can match objects from each image more accurately.
- Integration of Learning Algorithm to get appropriate model, which can be used by classification Algorithm to get the best match.
- Integration of classification Algorithms like SVM.

3. Plan / Schedule

As nature of Project is basically Research oriented so Iterative Model will best suit its needs. As the design throughout project will be changing considerably so each time design changes a small SDLC (Software Development Lifecycle) has to be performed.

The following table contains the estimated schedule of the project as it goes through the semester. This schedule is there to help and observe the status of project at particular time instance. As research projects have considerable amount of uncertainty in them to achieve goal so the following estimates can change little bit with the course of time.

Week	Date	Milestone
1-3	21 July – 10 August	Understand Domain and Requirements
4-6	11 August – 31 August	Finalize the major part of Design
7-11	1 September – 5 October	Implement and Verify Design and make changes Accordingly
12-14	6 October – 26 October	Finalize Documentation and prepare for presentation

4. References

[1] I. Borg and P. Groenen. *Modern Multidimensional Scaling, Theory and Applications*. Springer Series in Statistics. Springer, 1997.

[2] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman. Total recall: Automatic query expansion with a generative feature model for object retrieval. In *Int. Conference on Computer Vision*, 2007.

[3] S. E. El-Khamy, M. Lotfy, and N. El-Yamany. A modified fuzzy sobel edge detector. In *Proceedings of The 17th National Radio Science Conference*, pages C32/1–C32/9, 2000.

[4] S. A. M. Farzin and J. Kittler. Robust and efficient shape indexing through curvature scale space. In *Proceedings of the 7th British Machine Vision Conference*, volume 1, pages 53–62, 1996.

[5] L. Fei-Fei and P. Perona. A bayesian hierarchical model for learning natural scene categories. In *Comp. Vision and Pattern Recognition*, pages II:524–531, 2005.

[6] C. Harris and M. Stephens. On measuring low-level saliency in photographic images. In *Proceedings of The IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 84–89, 2000.

[7] C. J. Harris and M. Stephens. A combined corner and edge detector. In *Proc. 4th Alvey Vision Conference*, pages 147–151, 1988.

[8] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.

[9] X.-B. Li, J.-Y. Li, and R.-H. Wang. Dimensionality reduction using mce-optimized lda transformation. In *Proceedings of the IEEE Conference on Acoustics, Speech, and Signal Processing*, pages 37–40, 2004.

[10] D. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[11] W. Niblack et al. The qbic project: Querying images by content using color, texture and shape. In *Proc. SPIE Conference on Storage and Retrieval of Image and Video*

Databases, 1908, pages 173–187, 1993.

[12] M.-E. Nilsback and A. Zisserman. A visual vocabulary for flower classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1447–1454, 2006.

[13] D. Nister and H. Stewenius. Scalable recognition with a vocabulary tree. In *Comp. Vision and Pattern Recognition*, pages II:2161–2168, 2006.

[14] N. Otsu. A thresholding selection method from gray-level histobrams. *IEEE Trans. on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[15] R. W. Picard. Light-years from lena: Video and image libraries ind the furture. In *International Conference on Image Processing*, volume 1, pages 310–313, 1995.

[16] Y.-D. Qu, C.-S. Cui, S.-B. Chen, and Jin-QuanLi. A fast subpixel edge detection method using sobel-zernike moments operator. *Image Vision Computing*, 23(1):11–17, 2005.

[17] P. Quelhas, F. Monay, J. Odobez, D. Gatica-Perez, T. Tuytelaars, and L. V. Gool. Modelling scenes with local descriptors and latent aspects. In *Int. Conference on Computer Vision*, pages I:883–890, 2005.

[18] P. L. Rosin and G. A. W. West. Salience distance transforms. *CVGIP: Graphical Model and Image Processing*, 57(6):483–521, 1995.

[19] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In *Int. Conference on Computer Vision*, pages II:1470–1477, 2003.

[20] Y. Sun and S. T. Bow. Fast wavelet transform for color image compression. In *Proceedings of The International Conference on Image Processing*, volume 2, pages 541–544, 1996.

[21] F. Tang and H. Tao. Fast linear discriminant analysis using binary bases. In *Proceedings of the 18th International Conference on Pattern Recognition*, volume 2, pages 52–55, 2006.

[22] M. Tkalcic and J. F. Tasic. Colour spaces - perceptual, historical and applicational background. In *Proceedings of The IEEE Region 8 Conference on Computer as a Tool*, pages 304–308, 2003.

[23] H. Tong, M. Li, H. Zhang, and C. Zhang. Blur detection for digital images using wavelet transform. In *IEEE InternationalConference on Multimedia and Expo*, pages 17–20, 2004.

[24] N. Vasconcelos. On the efficient evaluation of probabilistic similarity functions for image retrieval. *IEEE Trans. On Informaiton Theory*, 50(7):1482–1496, 2004.

[25] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. of CVPR*, pages 511–518, 2001.

[26] J. Ye. Least squares linear discriminant analysis. In *Proceedings of the International Conference on Machine Learning*, pages 1087–1094, 2007.

[27] R. B. Yossi Cohen. Inferring region salience from binary and gray-level images. *Pattern Recognition*, 36(10):2349–2362, 2003.