

Image Recognition on Mobile Phones

Serdar Sali

Revised Related Work

2. Related Work

While CBIR is a large problem and encompasses a broad area of applications such as semantic retrieval of images, searching based on particular features of an image such as color distribution or frequency, or retrieving pictures with similar textual annotations, the relevant area to our research is retrieval of similar images to a sample input image, which is also called query by example.

Since we use as input images taken by a user, the recognition algorithm employed must be invariant to rotation, translation, scaling and partial occlusion. It is also important that the algorithm is invariant to some degree to noise in the image and differences in illumination.

2.1. Similarity based search

Over the years, various approaches have been proposed to address the problem of matching images based on similarity. These approaches mainly differ in how they form the feature vector for an image. The most widely used sources for feature extraction are color, shape and texture.

One of the first algorithms utilizing color was proposed by Swain et al. [1]. This algorithm computes color histograms for images, followed by an intersection of the histograms to perform the similarity search. Improvements that add spatial information and correlation to color histograms were proposed in [2] and [3]. Since color histograms are very sensitive to noisy images, applying this to our domain is likely to cause problems since cameras on mobile phones tend to take relatively noisy pictures.

In addition to color histograms, various methods look for textures in an image and their spatial placement, and construct a feature vector based on this information. One important work utilizing textures is given by Tamura et al. [4]. In this paper, the authors propose approximations to the following texture features: Coarseness, contrast, roughness, regularity, directionality and linelikeness. These properties are based on how humans actually perceive textures.

Shape can also be used to build a feature vector from an image. Methods relying on shape usually work on the similarity of edges, corners and shapes of the objects in the image. Feature extraction using shapes usually works in a local level: it is concerned with locating points of interest in the image, rather than considering the global distribution of a feature as in color-based methods. One of the most commonly used shape detectors is the corner detector by Harris [5].

SIFT (Scale-invariant feature transform) [6] is a popular method based in detection of key points in an image. The resulting features are invariant to scaling, differences in illumination and rotation, and works good even for 3D images and different points of view. These properties make SIFT ideal for our architecture.

2.2. Object Recognition Systems on Mobile Phones

While many web-based systems exist for content-based image search, applications of these systems to mobile devices are relatively rare. Some research is being conducted on building assistive technologies for the blind and visually impaired using mobile devices and object recognition. One such system is GroZi [7]. GroZi aims to help the visually impaired by recognizing and locating the items on a customer's shopping list among the items on the shelf.

Another set of applications aim to build on-demand mobile tour or museum guides [8]. Some of these systems try to perform location recognition by using photos from the location as input. Yeh et al. [9] propose a system that first queries an image database by using a photo taken by a user. After a match is found, the system then searches the web based on keywords associated with the determined location. PhoneGuide [10] is another application that makes use of the mobile phone as a museum guide. The user can take pictures of the various exhibits in the museum using the digital camera on her/his mobile phone, and then query a database for information on the exhibit using the image as the input.

Seifert et al. [11] propose a system that utilizes object recognition and GPS on a mobile device for taking an inventory of traffic signs. The algorithm for sign recognition takes advantage of the fact that traffic signs have obvious shapes and patterns which are known in advance.

Jia et al. [12] propose a generic architecture called Photo-to-Search that can be used to query the web directly from mobile devices using images taken by the digital camera on the device. The underlying object recognition system uses SIFT to detect key points and perform the matching. There are also systems for outdoor hobbyists, such as systems for flower or fish recognition [13,14].

References

- [1] S. Ballard, "Color Indexing", International Journal of Computer Vision, Vol. 7, No. 1, pp. 11-32, 1991.
- [2] G. Pass, and R. Zabith, "Comparing images using joint histograms," Multimedia Systems, Vol.7, pp.234-240, 1999.
- [3] G. Pass, and R. Zabith, "Histogram refinement for content-based image retrieval," IEEE Workshop on Applications of Computer Vision, pp. 96-102, 1996.
- [4] H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception", IEEE Trans. On Systems, Man, and Cybernetics, Vol. Smc-8, No. 6, 1978.
- [5] C. Harris and M. Stephens, "A combined corner and edge detector", ALVEY Vision Conference, pages 147-151, 1988.
- [6] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, Vol. 60, No. 2, pp. 91-110, 2004.
- [7] M. Merler, C. Galleguillos, and S. Belongie. "Recognizing Groceries in situ Using in vitro Training Data", to appear, SLAM 2007, Minneapolis, MN, retrieved from http://grozi.calit2.net/files/grozi_slam.pdf.
- [8] N. Davies, K. Cheverst, A. Dix, and A. Hesse, "Understanding the role of image recognition in mobile tour guides", Proceedings of the 7th international Conference on Human Computer interaction with Mobile Devices & Services, pp. 191-198, 2005.

- [9] T. Yeh, K. Tollmar, and T. Darrell, "Searching the Web with Mobile Images for Location Recognition", 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), Vol. 2, pp. 76-81, 2004.
- [10] P. Föckler, T. Zeidler, B. Brombach, E. Bruns, and O. Bimber, "PhoneGuide: museum guidance supported by on-device object recognition on mobile phones", Proceedings of the 4th international Conference on Mobile and Ubiquitous Multimedia, MUM '05, Vol. 154, pp. 3-10, 2005.
- [11] C. Seifert, L. Paletta, A. Jeitler, E. Hödl, J. Andreu, P. Luley and A. Almer, "Visual Object Detection for Mobile Road Sign Inventory", Lecture Notes in Computer Science, MobileHCI 2004, pp. 491-495, Springer, 2004.
- [12] M. Jia, X. Fan, X. Xie, M. Li, W. Ma, "Photo-to-Search: Using Camera Phones to Inquire of the Surrounding World", The 7th International Conference on Mobile Data Management (MDM'06), p.46, 2006.
- [13] M. Noda, H. Sonobe, S. Takagi, et al., "Cosmos: Convenient Image Retrieval System of Flowers for Mobile Computing Situations", Proceedings of the IASTED Conf. on Information Systems and Databases 2002, pp. 25-30, 2002.
- [14] H. Sonobe, S. Takagi, and F. Yoshimoto, "Image Retrieval System of Fishes Using a Mobile Device", Proceedings of International Workshop on Advanced Image Technology 2004, pp. 33-37, 2004.