Pattern Recognition 43 (2010) 431 – 433

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Editorial

Interactive imaging and vision—Ideas, algorithms and applications

For many problems in computational imaging and vision, it is often very difficult or maybe even impossible to develop completely automatic solutions. For example, despite much research effort, a fully automatic solution to the longstanding image segmentation problem remains unattainable. However, it is well known that humans have remarkable abilities in distinguishing different image regions or separating different classes of objects and can perform such tasks effortlessly. Similarly, automatically retrieving user desired images from large image repositories is also very difficult. This is partly due to the fact that a user's intention may differ under different circumstances; yet, methods for accurately capturing users' intentions remain elusive. It is often the case that a good solution to a problem in one application set-up may not be the one the user is looking for in another situation. For example, using query by image content (QBIC) approach to performing image retrieval, the user can use an example image as query. Suppose that a user has picked an image containing a horse in the foreground, blue sky and green grass in the background as the querying image. What is the user's intention? Is the user looking for images containing horses in the foreground, or images containing blue sky and green grass in the background, or images containing horses in the foreground and blue sky and green grass in the background? The image retrieval algorithms must have mechanisms to capture the user's intention in order to retrieve images that meets the user's expectation. However, automatically capturing user intentions from the querying examples remain to be a very challenging task. Furthermore, in some mission critical applications, such as medical image diagnosis, a fully automatic solution may even be undesirable and human intervention and participation in the decision making process may be necessary. Therefore, in tackling many problems in computational imaging and vision, it is both helpful and sometimes necessary to explicitly incorporate high level knowledge and human intentions. The scientific and technological challenges are how to capture and harness such high level knowledge computationally to solve real world problems.

The idea of interactive imaging and vision is to provide semi-automatic solutions to hard imaging and vision problems. The users are in the computational loop and interactively input their knowledge and intentions which are taken as constraints and priors by the computational algorithms to iteratively refine the models and solutions, as illustrated in Fig. 1. Such approach has the potential of not only exploiting human intelligence to develop powerful machine systems that can incorporate some of human's decision making power into the computational models but also turning computers into more useful and effective tools that can solve problems according to their users intends.

Human-in-the-loop computing is ubiquitous and the prevailing graphical user interface (GUI) in modern computer systems is a primary example. However, in traditional GUI, the computer's response to the user input, e.g., the clicking of an icon in the interface, is often "hard wired", therefore in this sense traditional human in the loop computing such as GUI is not intelligent nor is it adaptive. In contrast, the interactive approaches to imaging and vision problems have intelligent computational models that explicitly incorporate human inputs in the computational models either as constraint conditions to optimization problems or as priors of statistical models. Importantly, such approaches enable users to iteratively interact with the models to incrementally refine the models and solutions.

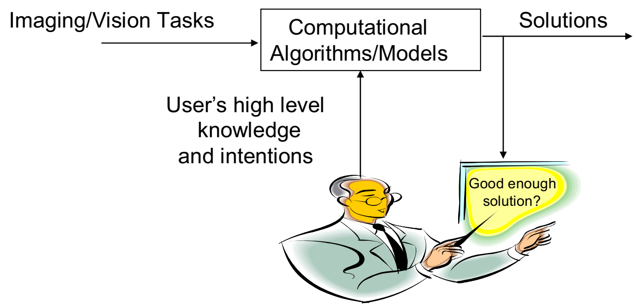

Fig. 1. An illustration of interactive approaches to imaging and vision. For a given imaging and vision task, an initial computational algorithm or model is developed to find a solution. This solution may or may not satisfy the user's expectation. If the solution is what the user is looking for and is good enough, then the problem is solved. However, more often than not, the solution may not match the user's expectation, in which case the user will interact with the computer to provide feedbacks which will contain the user's high level knowledge about the problem and the user's intentions in the particular application. These feedbacks will then be used by the computational algorithms and models as more accurate constraint conditions and stronger priors to refine the model and solution. Such user in the loop computational process can be iterated until a satisfactory solution is found.

Computational algorithms that have been explored in recent years for interactive imaging and vision include those based on graph cut [1,2,8], random walk [3], semi-supervised learning [5], constrained optimization [6,7], belief propagation [9], statistical modeling [7–10] and active learning [11]. Application areas include interactive image segmentation [1–4], interactive image matting [4,9], interactive colorization [6], interactive image retrieval [8,11] and interactive data clustering [10].

This special issue on Interactive Imaging and Vision includes 10 papers selected from over 50 submissions covering a wide range of interactive imaging and vision applications.

In recent years, many algorithms have been developed for interactive image segmentation. However, researchers often find it hard to

0031-3203/$ - see front matter
doi:10.1016/j.patcog.2009.04.010

choose one from the other for their specific applications. In the paper "A Comparative Evaluation of Interactive Segmentation Algorithms", McGuinness and O'Connor presented a comparative evaluation of four popular interactive segmentation algorithms using a series of user-experiments in which participants were tasked with extracting one hundred objects from a common dataset. They used a scribble-driven segmentation tool to enable interactive image segmentation and compared results against a manually segmented ground-truth. They used two benchmarks, the Jaccard index for measuring object accuracy and a new fuzzy metric for measuring boundary accuracy to gain insight into the performances and characteristics of the algorithms. In the paper "Interactive Image Segmentation by Maximal Similarity Based Region Merging", Zhang and co-workers presented a new region merging based interactive image segmentation method. They presented a novel maximal-similarity based region merging mechanism to guide the region merging process with the help of user supplied markers. Their method automatically merges the regions that were initially segmented by mean shift segmentation and then extracts the object contour by labeling all the non-marker regions as either background or object. The region merging process is adaptive and there is no need to set the similarity threshold in advance.

Any design process is often interactive and iterative. Interactive imaging and vision techniques are especially suited for designing image patterns such as texture. In the paper "Interactive Texture Design using Energy Optimization", Shen and his colleagues presented an interactive scheme for designing textures using energy optimization and deformation. With a little help from the user, their interactive texture design technique has the ability to change global and local visual properties of texture elements. Given a small sample texture, their design process starts by applying a set of global deformation operations to the sample texture to obtain a set of deformed textures automatically and then local deformation is applied interactively.

With the rapid rise of social media, billions of photographs are being uploaded onto online photo sharing sites. Making the huge numbers of photographs on the Web searchable is extremely challenging. One solution is to use the Web 2.0 technology to enable users to tag the images from anywhere at any time. However, manually tagging large number of images is labor intensive and automatic or semi-automatic image tagging tools should be developed to improve image tagging efficiency. In the paper "Semi-Automatic Dynamic Auxiliary-Tag-Aided Image Annotation", Zhang, Li and Xue presented a semi-automatic image tagging technique. They formulated image annotation as a multi-label learning problem, and developed a semi-automatic image annotation system. Their system chose proper words from a vocabulary as tags for a given image, and refines the tags with the help of user feedbacks. The refinement is achieved through a novel multi-label learning framework termed semi-automatic dynamic auxiliary-tag-aid (SADATA) in which the classification result for one tag can be boosted by the classification results of a subset of other tags that have strong correlations with the target tag. In the paper "Interactive Localized Content Based Image Retrieval with Multiple Instance Active Learning", Zhang, Wang, Shi and Zhang presented two general multiple instance active learning methods for localized content based image retrieval.

One way to quickly find a particular image from very large image collections is through active browsing, just like browsing through a book or text document for specific textual passages [12]. However, unlike textual documents, browsing through an image repository is much harder to do because visual contents and concepts are often vague and very hard to describe with unambiguous language. In the paper "Interactive unsupervised classification and visualization for browsing an image collection", Bruneau, Picarougne and Gelgon presented an approach to interactively navigating image collections. They argued that structured groups are more appealing to users than flat image collections and proposed an image clustering algorithm that incrementally handles time-varying collections. They also presented a 3D graph based interactive visualization technique that also enables user feedbacks for improving classification.

This special issue also includes the paper "Registration and Interactive Planar Segmentation for Stereo Images of Polyhedral" by Vigueras and Rivera in which the authors introduced a two-step iterative segmentation and registration method to find coplanar surfaces among stereo images of a polyhedral environment. The paper "A Semi-supervised Approach to Space Carving" by Prakash and Robles-Kelly described a semi-supervised approach to space carving in which the authors cast the problem of recovery of volumetric data from multiple views into an evidence combining setting. The method combines the advantages of shape-from-silhouette techniques and statistical space carving approaches. Quantitative results were used to illustrate the utility of the method on real-world imagery.

The usability issue of imaging and vision algorithms has not thus far received the attention that it deserves. As imaging and vision techniques are becoming more stable and mature, they have found increasing applications in other areas of computing such as computer human interaction in which practitioners are often not imaging and vision experts. To facilitate applications of image and vision technology, user-friendliness is very important. The paper "User-Centric Image Segmentation using an Interactive Parameter Adaptation Tool" by Pauplin, Caleb-Solly and Smith studied this issue. They identified that one of the key requirements for any interactive system is a high level of usability, both in terms of effectiveness—being able to build accurate models that meet end-user requirements and efficiency—being able to achieve the required results within a minimal amount of time and undue effort. They presented a system that has been designed with these considerations in mind to ensure a high level of user-experience of the interaction process.

One of the important areas that interactive approaches may be especially useful is medical application. In the paper "Classification and Interactive Segmentation of EEG Synchrony Patterns", Alba, Marroquín, Arce-Santana, and Harmony, presented a methodology for the exploratory analysis of power and synchronization patterns in EEG data obtained from psychophysiological experiments. The methodology was based on the segmentation of the time-frequency plane in regions with relatively homogeneous synchronization patterns. They implemented the method in an interactive application for the study of cognitive experiments.

We would like to thank the Editors in Chief, Professor Robert S. Ledley and Professor C.Y. Suen, for their support for this special issue. We would like to express our sincere gratitude to the many reviewers for their help in the paper selection process. We would like to thank the authors for submitting their works for consideration by this special issue. Finally, we would like to thank the staff at the journal's editorial office for their assistance.

References

[1] Y. Boykov, G. Funka-Lea, Graph cuts and efficient N-D image segmentation, International Journal of Computer Vision (IJCV) 70 (2) (2006) 109–131.
[2] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, IEEE Transactions on Pattern Analysis Machine Intelligence 26 (2004) 1124–1137.
[3] L. Grady, Random walks for image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (2006) 1.
[4] C. Rother, V. Kolmogorov, A. Blake, GrabCut: interactive foreground extraction using iterated graph cuts, ACM Transactions on Graphics 23 (3) (2004) 309–314.
[5] X. Zhu, Semi-supervised learning literature survey, Computer Science TR 1530, University of Wisconsin, Madison, 2008.
[6] A. Levin, D. Lischinski, Y. Weiss, Colorization using optimization, SIGGRAPH, ACM Transactions on Graphics (2004) 689–694.

[7] J. Guan, G. Qiu, Interactive image segmentation using optimization with statistical priors, in: Proceedings of International Workshop on the Representation and use of Prior Knowledge in Vision (in conjunction with ECCV 2006), Graz, Austria (CDROM), May 13, 2006.
[8] J. Guan, G. Qiu, Modeling user feedback using a hierarchical graphical model for interactive image retrieval, in: H.H.S. Ip, et al. (Eds.), Advances in Multimedia Information Processing—PCM 2007, Lecture Notes in Computer Sciences, vol. 4810, 2007, pp. 18–29.
[9] J. Wang, M. Cohen, Image and video matting: a survey, Foundations and Trends in Computer Graphics and Vision 3 (2) (2007) 97–175.
[10] A.K. Jain, P.K. Mallapragada, M. Law, Bayesian feedback in data clustering, in: Proceedings of International Conference on Pattern Recognition (ICPR), vol. 3, Hong Kong, August 2006, pp. 374–378.
[11] S. Tong, E. Chang, Support vector machine active learning for image retrieval, in: Proceedings of ACM Multimedia, 2001, pp. 107–118.
[12] G. Qiu, J. Morris, X. Fan, Visual guided navigation for image retrieval, Pattern Recognition 40 (6) (2007) 1711–1721.

Guest Editors
Guoping Qiu*
School of Computer Science, The University of Nottingham, UK
E-mail address: qiu@cs.nott.ac.uk

Pong C. Yuen
Department of Computer Science, Hong Kong Baptist University, Hong Kong, PR China
E-mail address: pcyuen@comp.hkbu.edu.hk

* Corresponding author. Tel.: +44 115 8466507; fax: +44 115 9514254.