MVA2015 IAPR International Conference on Machine Vision Applications, May 18-22, 2015, Tokyo, JAPAN

Research Opportunities in Creating Medical Images

Ramin Zabih
Cornell University
rdz@cs.cornell.edu

Abstract

Computer vision is traditionally defined as the analysis of the content of images. In the context of medical imaging, this naturally leads to a focus on segmentation and registration. I argue that there is an underappreciated opportunity for computer vision researchers in the domain of image creation. I illustrate the technical challenges involved, with an emphasis on magnetic resonance (MR) imaging. These problems have unusual features that make them intellectually interesting, beyond their obvious practical importance.

1 Medical image creation

Most research in medical imaging is driven, explicitly or implicitly, by the long-term goal of automatic diagnosis. As a consequence, the vast majority of work focuses on registration or segmentation. This is a natural course for researchers to take, since registration and segmentation problems arise in a wide range of computer vision problems.

However, a look at the radiology literature is quite revealing. Most issues of the main journal, *Radiology*, devote approximately 25% of their coverage to an important but underappreciated problem, namely the creation of diagnostically informative images in the first place. Most imaging devices (especially CT/PET/MR) collect raw data that does not directly correlate with pixel values; instead, the data must be somehow transformed to create an image. This problem is sometimes referred to as acquisition (which focuses on the collection process itself), and sometimes as reconstruction (which focuses on the transformation process). For simplicity, I will refer to this problem as image creation.

An illustrative example comes from CT imaging. Within the space of a few months, *Radiology* published papers that proposed new CT techniques for creating diagnostic images for ischemic stroke, prostate cancer and coronary artery disease.[1] Each of these papers provides evidence that their new technique will improve diagnostic accuracy.

1.1 Is image creation a computer vision problem?

The importance of creating a quality image for diagnostic purposes is entirely obvious, and is not affected by whether the diagnosis itself is performed by a computer, a human, or a human aided by sophisticated computational tools. However, it is not immediately obvious why image creation would be of interest to computer vision researchers. The traditional relationship between computer vision and computer graphics is that computer vision analyzes images, while computer graphics creates images. Under this definition, image creation is a graphics problem and not a vision problem at all.

However, it is clear that the traditional boundary between graphics and vision has been in flux for a number of years. SIGGRAPH regularly publishes papers that use computer vision techniques; the most prominent such examples come from the creation of photo montages, but there are many others. Such papers also frequently appear in the computer vision literature.

To put it differently, on a regular basis the main vision conferences and journals publish algorithms whose output is another image. This, of course, is very much in the tradition of image processing, and in fact the line between image processing and computer vision has always been a very blurry one. (In my personal opinion, it is primarily a philosophical distinction: image processing researchers do not generally think about the 3D scene that gave rise to an image, while this is of great importance to researchers in computer vision.)

1.2 Deconvolution with priors

A particularly interesting and relevant example comes from image deblurring or deconvolution, which is a core image procesing topic that has gathered great interest from computer vision researchers. While there are many papers in the field, I wish to particularly point to the work of Anat Levin and her co-authors (see for example [16]) and the research from Bill Freeman's group on motion magnification [4].

In deconvolution, the goal is to create a fast algorithm to solve an ill-posed problem by imposing realistic priors. The different methods for this problem are distinguished primarily by their choice of priors and of algorithms. These choices are not independent; an unrealistic prior can sometimes be justified on computational grounds, while a fast algorithm can make previously intractable priors suddenly practical. In general, more specific algorithms perform better than more general algorithms. This is a fundamental tradeoff in computer science, and suggests that the best optimization technique to choose is the most specific one that can be adapted to solve a particular problem, perhaps by changing the problem definition slightly.

2 Image creation is deconvolution with a prior

The creation of CT and MR images from the raw data acquired by a scanner turns out to involve a particular form of deconvolution. In many respects this fact is retrospectively obvious, though it seems to be

[1]The papers mentioned have PubMed ID 21062925, 20663968 and 20093513, respectively. I have not cited them explicitly in the bibliography since they are examples purely for purposes of illustration.

underappreciated in the computer vision and medical imaging community. Scanners, like most sensors, are linear within their intended operating ranges. The scanning process inherently loses information, and so one is faced with the problem of creating an image that is consistent with the observed data but also with prior expectations about the general appearance of images.

Imaging problems that are inherently ambiguous, and which require fast algorithms for realistic priors, are a major topic in computer vision and lie at the core of my own research interests. In addition, as mentioned above this is a problem of great practical important to clinicians; a substantially better method for creating CT or MR images would almost certainly save a considerable number of lives.

Within the MR community, there has been slow movement towards thinking of the image creation problem as a deconvolution problem with a prior. Some of the most notable work in this direction was done by Z-P Liang and colleagues [18, 17].

2.1 Can we do better via post-processing?

For researchers who are not excited by the technical challenges of deconvolution problems, a natural approach is to take the output of current image creation algorithms and try to improve it as a post-processing step. While tempting, in my view this approach is not likely to lead to success. The artifacts created have structures that reflect the image aquisition process, and can be quite strong; for example, CT images often exhibit streaking artifact, while MR tends to have ghosting. Without modifying the underlying image creation process I do not believe that it is possible to correct such artifacts without damaging the underlying images. I take as evidence the experience with fMRI images, which are significantly distorted during the acquisition phase; attempts to undo this distortion during post-processing have had limited success [10].

3 Parallel imaging in MRI

A particularly interesting image creation problem arises from a technique to accelerate MRI called parallel imaging. Reducing scan time is of great importance for MR; it is not just a matter of patient comfort, important though that is. MR imaging essentially generates a limited amount of information per unit time, and this information can be deployed to either improve spatial or temporal resolution.

As mentioned above, fMRI images are severely distorted and have poor spatial resolution; this largely arises from the need for speedy imaging of neural activity. Better image creation techniques would lead to higher spatial or temporal resolution, which would be of great value. Beyond fMRI, the most obvious target for MR is cardiac imaging, where motion is an enormous problem. Part of the success of MR for neuroimaging arises from the simple fact that the head is one of the few body parts that can be immbolized for a prolonged period.

Parallel imaging [20] uses multiple receiver coils to substantially accelerate MR imaging, often by a factor of 2 or 3. The imaging problem is explicitly formulated as a linear inverse system and solved with least squares. The standard Tikhonov regularizer used [20], however, imposes an unnnatural prior that prefers images with the lowest possible intensity (i.e., a perfect MRI is completely black). None the less, the resulting inference problem is convex and can be solved quite efficiently.

3.1 Parallel imaging with graph cuts

My own interest in this problem was due to my student Ashish Raj, who realized that graph cut methods could be quite helpful for this parallel imaging problem. Graph cut techniques, which were popularized in vision by my research group [3, 14], can perform fast inference with realistic image priors. These methods are widely used, and successful on a range of benchmarks [25]. [24] contains a good discussion of graph cuts and their applications in vision, together with an overview of some competing approaches. A review of discrete optimization methods for vision is provided by [5].

Ashish's initial observation was that the linear inverse system that arises in parallel imaging poses some serious difficulties for graph cut methods [22]. Shortly afterwards, however, Kolmogorov and Rother introduced into computer vision some much more general graph cut techniques [13], based on the work of optimization researchers including Boros and Hammer [2]. We went on to apply these improved graph cut methods to MR parallel imaging [21]. A number of improvements to the basic method have been proposed, notably a cardiac imaging technique [23]. I will note in passsing that my group's work on this problem was more or less contemporaneous with the well-known work of Lustig and collaborators on the use of compressed sensing to solve similar problems [19].

3.2 The research frontier: higher order priors

The main challenge with graph cut methods for MR reconstruction is that the priors that we support with fast inference, while realistic, were simply too weak. The graph cut priors involve pairs of adjacent pixels; while we can support a realistic assumption that images are piecewise smooth, the restriction to pairs of pixels at a time is a major limitation.

Many groups are working on higher order priors; a notable early success was [26], which showed their effectiveness for certain stereo matching problems. A very incomplete list of researchers who have done important work on higher order priors would need to include Hiroshi Ishikawa, Nikos Paragios, Phil Torr, as well as many others; in fact, the *IEEE Transactions on Pattern Analysis and Machine Intelligence* has a special issue on higher order priors in press.

My own work on this topic has focused on graph cuts. One line of work reduces higher-order priors to first-order; the first paper to make this practical was [11, 12]. Working with Endre Boros, my student Alex Fix proposed an improved reduction method [6, 7]. Another approach, developed by Alex Fix and Chen Wang, is to generalize the powerful primal-dual approach of Komodakis and Tziritas [15] to handle higher order terms directly [8].

4 An alternative approach

A closely related way to think about this problem, that does not directly involve images, is the selection of the appropriate parameters for an imaging exam. Modern CT and MRI consoles are essentially programmed with templates, that provide default parameter settings (called protocols) for a variety of different imaging exams.

The selection of the appropriate protocol is often done by technologists rather than physicians, especially at smaller hospitals or when there is no radiologist on duty (see, e.g., [9]). There is considerable evidence that mistakes in protocol selection are costly, either in terms of repeated exams (and, for CT, additional ionizing radiation), or worse still in terms of reducing diagnostic accuracy. My group has looked at this problem recently from the perspective of machine learning [1], and obtained some interesting preliminary results.

5 Conclusions

In this paper I have articulated a case for researchers in computer vision and medical imaging to focus on the image creation problem. The problem combines interesting theoretical aspects with the potential for substantial impact.

My own work has focused on graph cut methods, and their extensions to higher order. The higher-order graph cut techniques we have developed in [7, 8] are very powerful, and have interesting theoretical guarantees. However, quite a bit of engineering work remains to be done to close the loop and show that these methods generate improved performance on MR reconstruction problems.

Acknowledgements: My interest in this topic was spurred by Yi Wang and Martin Prince, who introduced me to this topic. Whatever success my research has garnered is due primarily to a string of excellent students whom I have had the good fortune to advise. Ashish Raj was my first student to focus in this area, and has greatly helped me understand the range of technical issues involved; Gurmeet Singh in turn extended Ashish's results in interesting ways. Alex Fix and Chen Wang are primarily responsible for the work on higher-order inference that I believe is the most promising avenue for addressing these problems.

The Radiology department at Weill Cornell Medical College, under the leadership of Dirk Sostman and Robert Min, has supported me for over a decade; I particularly thank Gary Dorfman for finding a way to balance the department's needs and my own research interests. This research has also been supported by the National Science Foundation, under NSF grant IIS-1161282.

References

[1] A. Bhat, G. Shih, and R. Zabih. Automatic selection of radiological protocols using machine learning. In *Proceedings of the KDD workshop on Data mining for medicine and healthcare*, pages 52–55. ACM, 2011.

[2] E. Boros and P. L. Hammer. Pseudo-boolean optimization. *Discrete Applied Mathematics*, 123(1):155–225, 2002.

[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *TPAMI*, 23(11):1222–1239, 2001.

[4] F. Durand, W. T. Freeman, and M. Rubinstein. A world-of-movement. *Scientific American*, 312(1):46–51, 2015.

[5] P. F. Felzenszwalb and R. Zabih. Dynamic programming and graph algorithms in computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(4):721–740, 2011.

[6] A. Fix, A. Gruber, E. Boros, and R. Zabih. A graph cut algorithm for higher-order Markov Random Fields. In *ICCV*, 2011.

[7] A. Fix, A. Gruber, E. Boros, and R. Zabih. A hypergraph-based reduction for higher-order binary Markov Random Fields. *TPAMI*, 2015.

[8] A. Fix, C. Wang, and R. Zabih. A primal-dual algorithm for higher-order multilabel markov random fields. In *CVPR*, 2014.

[9] R. Gupta, S. Greer, and E. Martin. Inefficiencies in a rural trauma system: the burden of repeat imaging in interfacility transfers. *J. Trauma*, 69(2):253–5, 2010. PubMed PMID:20699732.

[10] C. Hutton, A. Bork, O. Josephs, R. Deichmann, J. Ashburner, and R. Turner. Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage*, 16(1):217–40, 2002.

[11] H. Ishikawa. Higher-order clique reduction in binary graph cut. In *CVPR*, 2009.

[12] H. Ishikawa. Transformation of general binary MRF minimization to the first order case. *TPAMI*, 33(6):1234–1249, 2010.

[13] V. Kolmogorov and C. Rother. Minimizing nonsubmodular functions with graph cuts-a review. *TPAMI*, 29(7):1274–1279, 2007. Earlier version appears as technical report MSR-TR-2006-100.

[14] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *TPAMI*, 26(2):147–59, 2004.

[15] N. Komodakis and G. Tziritas. Approximate labeling via graph cuts based on linear programming. *TPAMI*, 29(8):1436–1453, 2007.

[16] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, pages 1964–1971. IEEE, 2009.

[17] Z. P. Liang, R. Bammer, J. Ji, N. Pelc, and G. Glover. Making better SENSE: Wavelet denoising,Tikhonov regularization, and total least squares. In *ISMRM*, page 2388, 2002.

[18] Z.-P. Liang and P. C. Lauterbur. *Principles of Magnetic Resonance Imaging: A Signal Processing Perspective*. IEEE press, 1999.

[19] L. M, D. D, and P. J. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 65(1):184–189, 2011.

[20] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and B. Peter. SENSE: sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 2001.

[21] A. Raj, G. Singh, R. Zabih, B. Kressler, Y. Wang, N. Schuff, and M. Weiner. Bayesian parallel imaging with edge-preserving priors. *Magnetic Resonance in Medicine*, 57(1):8–21, 2007.

[22] A. Raj and R. Zabih. A graph cut algorithm for generalized image deconvolution. In *International Conference on Computer Vision (ICCV)*, 2005.

[23] G. Singh, A. Raj, B. Kressler, T. Nguyen, P. Spincemaille, R. Zabih, and Y. Wang. A fast edge preserving

Bayesian reconstruction method for parallel imaging applications in cardiac MRI. *Magnetic Resonance in Medicine*, 65(1):184–189, 2011.

[24] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2010.

[25] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for Markov Random Fields. *TPAMI*, 30(6):1068–1080, 2008.

[26] O. Woodford, P. Torr, I. Reid, and A. Fitzgibbon. Global stereo reconstruction under second-order smoothness priors. *TPAMI*, 31(12):2115–2128, 2009.