Pattern Recognition Letters 43 (2014) 39–46

Contents lists available at SciVerse ScienceDirect

Pattern Recognition Letters

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

Revisioning the unification of syntax, semantics and statistics in shape analysis

Anand Rangarajan [1]

Department of Computer and Information Science and Engineering (CISE), University of Florida, Gainesville, FL 32611-6120, United States

ARTICLE INFO

Article history:
Available online 9 July 2013

Communicated by H. Bunke

Keywords:
Uncertainty
Interaction
Hamilton–Jacobi
Schrödinger
Intentionality
Compositionality

ABSTRACT

Hidden patterns, rendered visible by hindsight, often stand revealed as strong influences. Rama Chellappa's lab in the mid to late '80s molded our character and scholarship in more ways than one. Rama ably handled the transition from image processing to computer vision and established an applied math and computing infrastructure from which we continue to benefit. In particular, the themes important to King-Sun Fu—*syntax, semantics and statistics*—were all debated in Rama's lab at that time. We argue that this triad remains important. With syntactic representations losing mindshare to statistics, we remain in the hunt for unification. And with the syntax versus semantics debate unresolved, it deserves a hearing as well. We offer two themes—uncertainty and interaction—to aid in the process of unification. First, we show that complex wave functions carry probabilistic location information in their magnitude and syntactic (curve) information in their phase while representing uncertainty at a fundamental level. Next, after reviewing work in analytic philosophy, we connect semantics to intentional, mental content. Analytic philosophy reminds us to take human experience seriously but remaining physicalist if possible. To this end, we introduce a nondualist interactionist model of experience, wherein compositional (physical) subjects are constantly shaping and being shaped by a physical world. We then demonstrate that wave functions can accommodate interaction, closely tracking previous work in physics on the measurement problem. The linearity and superposition properties of wave functions allow for literal addition of waves created by human interaction with shapes. Finally, we briefly survey the current situation in the human–computer interaction (HCI) field and argue that mathematical models of interaction akin to those in pattern recognition can aid HCI. We close by arguing that we can follow in Fu's footsteps and incorporate the mathematical modeling of human interaction into pattern recognition.

1. Invocation

Hold the flame 'til the dream ignites

A spirit with a vision is a dream with a mission

Rush, Mission

Luck is often described as "being in the right place at the right time". Looking back, it seems clear that we were very fortunate indeed to inhabit Prof. Rama Chellappa's lab during the mid '80s, in the Signal and Image Processing Institute (SIPI) at the University of Southern California (USC) set amidst the backdrop of a sophisticated, urban Los Angeles. Computer vision was in its infancy, struggling to emerge from the shadow of an established artificial intelligence (AI). Pattern recognition was seeking to reinvent itself in the hands of a resurgent neural networks field (Bishop, 1996) and would subsequently find a more stable partner in machine learning (Bishop, 2007). If there was interest in human centered computing, we certainly did not see much evidence at that point.

Rama was and remains the embodiment of that transition from a field taxonomy so stable (in the late '70s) that it verged on taxidermy to a period (the early '80s) when "all that is solid melts into air"[2] (Berman, 1988). Not content with resting on his laurels in signal and image processing, Rama sought to bring the clarity and rigor found in mature (and older) fields that stood on the shoulders of applied mathematics to the nascent (and therefore fertile) area of computer vision. This quest was aided by three significant factors: (i) the brilliance of Marr (1982) in clearly articulating the nature of representation in computer vision, (ii) the integrative genius of Fu (1986) in bringing together syntax, semantics and statistics in pattern recognition, and (iii) the strong reliance of neural networks on statistical mechanics, nonlinear optimization and applied mathematics in general. While we return to these themes frequently in this

[1] This work is partially supported by NSF IIS RI:1143963.
E-mail address: anand@cise.ufl.edu

[2] The first IEEE Computer Vision and Pattern Recognition (CVPR) conference was held in 1983.

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.06.017

work, it is to the specifics of Rama's contributions—and the manner in which they illustrate the deployment of applied math in computer vision—that we now turn.

Since we're using a signal processing to computer vision crossover perspective, we bypass Rama's considerable work on Markov random fields in image processing and analysis (Chellappa and Kashyap, 1982) except to note in passing that this framework has served us well over the past three decades. Instead, we highlight the twin contributions in shape from shading and motion estimation to buttress our crossover points. In shape from shading, the problem of enforcing integrability was a thorn in the variational framework. Frankot and Chellappa (1988) designed an elegant projection algorithm to find the closest, valid surface in a least-squares sense to the non-integrable surface obtained via the calculus of variations approach. The projection method was signal processing inspired through and through using the fast Fourier transform (FFT) and least-squares estimation to full effect. Next up, in motion analysis, Broida and Chellappa (1986) were the first to use a Kalman filter to recursively estimate rigid body motion parameters from noisy images. Once again, the application of signal processing methodologies (filtering, estimation etc.) bore immediate fruit in an important computer vision problem. In Rama's lab, these early successes reinforced in all of us not only the importance of mathematical rigor in vision problems but the impact of ideas drawn from other fields such as signal processing on the fledgling area of computer vision. Furthermore, the specific successes in shape from shading and motion estimation emboldened Rama to emphasize (i) the correspondence problem in object recognition and motion and (ii) the link between variational problems and specific differential equations such as the Poisson equation (Horn, 1986; Simchony et al., 1990). The emphasis on correspondence in the face of an ascendant optical flow paradigm was particularly courageous and key to shaping my views (Gold and Rangarajan, 1996; Chui and Rangarajan, 2000) even though I did not realize it at that time.

SIPI culture in general and Rama's background in MRFs in particular meant that inference and estimation principles were drilled into us usually via statistics courses taken in the math department. Geman and Geman's landmark paper on MRFs (Geman and Geman, 1984), Gibbs sampling and simulated annealing forced us to pick up quite a bit of statistical mechanics as well. Entropy, computational temperature and more broadly, the role of uncertainty became common themes in conversations. The change in worldview accompanying the transition from image processing to computer vision meant that we had to grapple with artificial intelligence for the first time. For many of us with undergraduate engineering backgrounds—with a lack of emphasis on philosophy—this was simultaneously uncomfortable and exciting. In particular, this implied taking the central issue of representation in AI seriously. Syntax versus semantics (Putnam, 1980), the Chinese room (Searle, 1980) and more generally the nooks and crannies of strong AI (Russell and Norvig, 2009) were endlessly debated. It is no coincidence that Fu's seminal work on the unification of syntax and semantics (Fu, 1986) with a strong focus on statistical pattern recognition became a linchpin. Fu forced us to focus on the triad—*syntax, semantics and statistics*—while remaining grounded in the nuts and bolts of pattern recognition and computer vision.

We argue that Fu's triad is even more relevant today. While the past thirty years has seen much progress in the development of robust algorithms (for classification and regression), it is fair to say that the dominant paradigm has been statistics with syntactic representations frequently discarded or sidelined due to perceived (and actual) brittleness in the latter. The importance of uncertainty in representation has also faded and replaced (unsuccessfully as we shall argue) by variance and entropy in parametric and non-parametric probability distributions. The syntax versus semantics debate has not fared any better in mainstream AI circles. While analytic philosophy has made tremendous strides in the past twenty years in delineating the distinction between the computational and the experiential, these advances have not been absorbed in information processing circles. Concomitantly, we have seen the rise of human centered computation and the field of human–computer interaction (HCI), but we are not aware of any serious attempts to educate the (usually) younger denizens of HCI with the significance of Fu's triad in their endeavors. Consequently, HCI and related areas remain unaware of the importance of interaction (between the experiential and the computational) in anchoring semantics while informing syntactic and statistical representations. To this end, we identify the twin themes of uncertainty and interaction as central and attempt to unpack their significance in this essay.

2. Uncertainty

Who can face the knowledge
That the truth is not the truth
Obsolete
Absolute
Rush, Distant Early Warning

In Section 1, we introduced Fu's contributions in the larger context of the entire field of pattern recognition. In this section, we focus on shape analysis in our attempt to discuss uncertainty in syntactic and statistical representations. We will not be concerned with semantics here, which is deferred to Section 3.

Representations inspired by probability theory and statistics—placed under the same rubric here—have thrived and prospered in shape analysis. Especially when shapes are parametrized by point-sets, probabilistic representations have become quite popular due to the relative ease of density estimation in lower dimensions. In the past twenty years, shape correspondence, non-rigid deformable matching, shape dictionaries etc. have all seen considerable progress since the robustness afforded by the representations has allowed for outlier detection, incomplete shape matching and so on. In many cases, shape density functions are first estimated using Parzen windows or related methods. Subsequently, shape densities are matched using entropy minimization or other criteria to obtain shape deformation, shape atlases (Chen et al., 2012) and the like. Note the simplicity of the representation schemes which feature little to no explicit syntax. Point-sets are generally i.i.d. allowing for straightforward density estimation and the Hausdorff topology prevents the use of relational information. In spite of this, the robustness of the representation can account for both point jitter and outliers belying the need for synthesizing relational and statistical information.

When shapes are parsed into sets of non self- or other-intersecting closed planar curves (in 2D), level sets and distance transform representations have become popular in this space (Osher and Fedkiw, 2002). Distance transforms satisfy the eikonal equation $\|\nabla S\| = 1$ (with a constant forcing function) with the zero level sets comprising the shape. In contrast to the point-set representation, distance functions embed curve syntactic information into a scalar field $S(x)$. There is no room for uncertainty in the representation however, for signed and unsigned distance functions are highly constrained geometric objects that leverage the curve (relational) topology information. These constraints (implicit in $\|\nabla S\| = 1$) do not allow distance functions to be added, for example, or facilitate shape atlas computation. When one shape contains two curves and another three curves, their scalar distance transform fields cannot be easily combined. Consequently, while this representation has flourished with active contours and level sets