What is a Feature?
A Fast and Frugal Approach to the Study of Visual Properties

Dario Taraborelli

Institut Jean Nicod
1bis, avenue de Lowendal
F-75007 Paris
taraborelli@ens.fr
fax : +33 (0)1 53 59 32 99
phone : +33 (0)1 53 59 32 80

March 6, 2003

Abstract

1. What is a 'good' feature?

Perceptual systems, both biological and artificial, are broadly speaking systems that can extract certain features from a sensory array and use these features for categorisation tasks, recognition and motor control. Several characterisations of the notion of 'feature' have been proposed so far in specific fields of vision science. Yet, a *general* account and tentative classification of what counts as a 'good' visual feature is still missing. The 'goodness' of a feature is usually defined according to very local and specific principles: in this paper, I will survey the main criteria used in vision science to define what is a 'good' visual feature and I will suggest a more general account for assessing 'goodness' of a feature in a biologically plausible way.

1.1 A theoretical bias in the study of visual properties: ascribability to individuals

Vision science seems to privilege specific classes of stimulus properties over other classes of properties in studying visual processing. Many studies implicitly assume a number of conditions of *well-formedness* for visual features, thus excluding those properties that do not match such criteria. They are not, generally speaking, necessary and sufficient conditions, but rather criteria that orient the selection of relevant variables chosen in experimental settings and modelling.
A strong privilege in vision science is given to features that are *ascribable* to individuals. Possible candidates for these individual entities range from retinal regions (corresponding to receptive fields of neurones in the primary visual cortex), to regions of perceptual space (in an egocentric or allocentric frame of reference), to bounded material objects etc. 'There is a *vertical* boundary at P<x,y> of the left retinal image', 'I see a *green* spot in front of me', 'a *white* rabbit just crossed the street': in any of these cases, a feature is considered as a property that 'holds of' or 'is instantiated by' one of such individual entities. I will suggest that the privilege given to ascribability depends on strong normative constraints about the alleged aim of perception.

1.2 Localisation

A leading idea in vision science dating back to David Marr is that one of the main tasks of human vision is to correctly localise features: the aim of the visual system is to give a spatial description of the world that specifies *what* is *where* [Marr 1982]. Whatever frame of reference we are dealing with, visual features seem to be by definition properties that refer to spatial locations and whose spatial distribution can be described in terms of maps [Treisman and Gelade 1980]. Investigating the relation between features and places, A. Clark concludes that features are either visual properties used in a feature placing language, phenomenal properties that a subject may use to describe her visual experience or properties eliciting *selective* and *retinotopically organised* activity in the primary visual cortex [Clark 2000]. In each of these cases, localisability – however considered, at a sub-personal, personal or interpersonal level – is a major constraint on the classes of possible candidates for 'good' features.

1.3 Objects

A longstanding tradition has focused on spatial location because of its role in encoding object identity. Recent works in experimental psychology have tried to dissociate spatial location of an object from its identity through spacetime. Given the priority of object recognition in standard theories of high-level vision, these works have provided a *prima facie* argument against the idea that the visual system parses the visual scene into objects *only* (or *mainly*) by specifying what visual features are instantiated at a specific location [Campbell 2002]. Empirical evidence shows that the human visual system is able to track objects and their featural properties across time *independently* of their spatial location. Features do not need to be anchored to places, they can be encoded for entities with complicated spatio-temporal patterns. [Scholl et al., submitted paper]. Yet, the replacement of 'places' with 'visual objects' does not infringe the ascribability principle, in that features are still restricted to properties that can be 'ascribed' to individuals.

Similarly, current debates about the nature of attention [Scholl 2001], whether attention is based on 'objects' instead of 'spatial locations', still assume that featural information can be selected by focusing on either of such kinds of entities, but do not question the ascribability axiom.
The extreme capillarity of ascribability criteria excludes *de facto* from the set of possible candidates for good features all those properties that are not susceptible of 'holding of' or 'being instantiated by' an individual.

1.4 A reconstructivist paradigm: normative constraints on the aim of perception.

The fact that ascribability principles are so widespread in vision science might be due to implicit normative assumptions concerning the alleged goal of perceptual systems. As the most common formulations of feature binding suggest, vision is considered as a process of *correct* specification or *correct* attribution of featural information to an individual or to a spatial region of the visual field. Illusions from feature binding are considered straightforward examples of *incorrect object reconstruction* [Treisman & Schmidt 1982]: objects are considered as bundles of features that need to be correctly bound together in order to be pointed, tracked across our visual field and seized by an adequate motor program [Treisman 1999].
Therefore, features require to be 'bound' to the correct individual in all those tasks that require a specification of a.) what and how many objects are present in the scene and b.) whether an object has or has not a specific property. Feature selection is in this respect a matter of correct ascription of a property to an individual, while feature integration is a matter of correct conjunction of features belonging to the same entity in the real world.
Although consistent with known capacities of human perceptual systems, this reconstructivist account of features in vision need not be the only plausible way of dealing with properties that a biological perceptual system might be designed to process.

2. <u>Features in a bayesian framework</u>

An alternative way to assess 'goodness' of visual properties (independently of their ascribability to individuals) is to consider features as simple environmental measurements serving as 'cues' for inferring complex world properties in structured environments. This approach, though quite widespread in the computational vision community [Knill and Richards 1996], is still far to be acknowledged in other domains of vision science. There is a number of interesting arguments in favour of the idea that a large part of our perceptual system might essentially work as a probabilistic device to draw fast and reliable inferences from low level sensory patterns.

2.1 Fast and frugal heuristics as cognitive shortcuts in a structured environment

The idea that the brain might use hardwired inferential and heuristics-based strategies to retrieve information from the world is certainly a very old and respectable, dating back to Helmholtz and recently resumed by Barlow [Barlow 1974]. The research program on biases and heuristics [Tversky and Kahnemann 1974] has drawn attention on the peculiarity of living beings as individuals endowed with bounded rational resources. What has been so far underestimated is the fact that biases in perceptual system do not necessarily represent a burden, but can serve as powerful tools for fast and reliable reasoning in bounded cognitive systems embedded in highly constrained environments [Todd 1999].

2.2 Features as statistically salient low level cues for high level properties

Reliable statistical correlations between the distribution of low-level features and allegedly high-level properties in a given environmental setting are exploited by current bayesian models of object recognition. Features respecting minimal criteria of likelihood, genericity and informativeness are usually considered as 'good' features for artificial perceptual systems [Jepson 1992]. Simple heuristics that exploit reliable relationships between low-level and easily accessible sensory patterns and much-more-harder-to-compute properties play also an important role in models of fast categorisation [Berretty et al. 1997]. The challenge is to see if these models can apply to biological system as well for describing plausible strategies used by living creatures to process sensory information.

2.3 Empirical evidence from developmental studies

A first open question is whether the idea of perceptual systems as probabilistic devices makes sense from a developmental point of view. Much attention has been devoted to heuristics-based capacities during cognitive development because of the lack of top-down information and the limits of available overall resources [Johnson and Morton 1991; Elman 1993]. F. Keil has argued that specific abilities to categorise in a fast and reliable way complex stimuli might come from the fact that perceptual systems of children usually develop in environments where the distribution of low-level sensory patterns is highly constrained. He calls 'perceptual shunt' such an acquired strategy for retrieving complex properties by solely relying on low-level cues and speculates that adaptive mechanisms of this kind might be much more common in human perception than we actually believe [Keil et al. 1998]. The main challenge in dealing with developmental evidence is to see whether such strategies only play a significant role during learning or if they are still present in adults for solving in a fast way perceptual problems raised by stimulus indeterminacy or for reasoning under heavy computational load.

2.4 Cost and gain issues

The main burden on heuristics-based approaches to the study of perceptual system is the classical problem of inductive generalisation: how informative and reliable is a cue that a system may use it in a heuristic fashion? As we have seen, heuristics are robust tools only in highly constrained environments, where the structure of the environment is providing contextual information thus fixing a number of implicit assumptions. The fact that redness is a good cue for edible food might be valid in specific niches, but ceases to be a reliable correlation as soon as we move to a larger ecological context [Feldman 1999]. Now, what happens if the environment is not enough constrained or too variable to allow reliable probabilistic inferences? What gain do probabilistic models offer for perceptual systems whose routines can vary on a wide range of contexts?

If the bayesian framework for the study of visual features aims to provide not only a set of mathematical tools to model how to solve abstract perceptual problems under uncertainty, but also a plausible model of how biological systems might work, then a crucial question is how such models can cope with the problem of contextual variability and whether context switch does not impose too heavy a burden on effective gain.

2.5 Conclusions: features as reliable cues *vs*. features as attributes for individuals

Although many problems are raised by the heuristics-based approach to perception, a first conclusion can be drawn from the comparison of this approach with more standard approaches in psychology of vision.

'Goodness' of featural information in traditional models of visual perception is related to the extent to which this information is effective in allowing correct binding or parsing of the visual scene into objects, correct attribution of feature to individual entities and correct object recognition.

'Goodness' of featural information in heuristics-based models of perception, on the contrary, cannot be assessed *independently* of the specific environment in which the perceptual system is embedded: 'goodness' of a feature, in this view, is rather a function of how reliable, informative and generic a cue is for retrieving complex properties, *given* a specific environmental setting.

Beside ecological constraints, feature 'goodness' can also depend from specific kinds of routines in which the visual system is engaged [Rensink 2000], so that reliability, informativeness and genericity of a cue can be task-related factors.

The main challenge of probabilistic models of perception is to see not only what are the possible uses of inferential rules that might be applied by perceptual systems, but what are the *actual* features that the human perceptual systems use, *given* their developmental, ecological and cognitive constraints.

3. Discussion and further questions

3.1 What is 'early processing'?

An interesting side-issue to this approach concerns the criteria that are used for defining *early* as opposed to *late* processing. Criteria for distinguishing what makes early vision distinct from high-level vision range from neurophysiological and psychophysical parameters (such as selective activity in primary visual areas or retinotopical organisation of receptive fields) to more behavioural parameters (such as lack of conscious access, encapsulation with respect to knowledge, rapid reaction times, use of preferred rules under high attentional load etc.).

Some authors have challenged the traditional taxonomy of low-level vs. high-level features by providing experimental evidence that many of these criteria, which are used to defined early processing, apply also to allegedly 'high level' properties [Rensink and Enns 1995]. These results provide further support to the idea of direct heuristic strategies causing rapid and effective reaction to such properties. How then can the early vs. late processing distinction be reformulated in order to account for early sensitivity to complex properties?

3.2 Doing without individuals?

In the first part of this paper, I argued that the main trend of vision science considers scene parsing and object recognition as the main aim of perceptual systems and that this assumption imposes strong normative constraints on the set of 'good' features defined at the input level. A speculative consequence of the alternative approach is that, since features are supposedly task-related and bounded to informativeness and reliability parameters given a specific environmental setting, the 'ascribability principle', according to which features should be limited to visual properties that can be bound together to yield individuals or that can be anchored to spatial locations in a given frame of reference, might turn out to be too strong and restrict too much the set of interesting properties. A possible empirical question would be then to investigate how many *high-level* visual routines do not rely on object construction, feature integration or binding of features to spatial locations [see for instance Li et al. 2002].

3.3 Phylogenetic considerations

There is strong disagreement on what should count as a 'good' feature in standard vision science on the one hand and ethology or evolutionary psychology on the other hand. Main features selected in recognition-oriented computational vision and psychology are usually considered quite irrelevant for studies concerned with cues orienting selective and rapid reaction in animal behaviour or visual properties likely to have been selected during phylogenetic development of a species in its adaptive environment [Cosmides and Tooby 1994]. The heuristics-based approach is promising in providing a framework for dealing with cognitive capacities that are relevant for specific environmental settings. It can also be used to predict specific biases or systematic perceptual errors that might derive from evolutionary internalised heuristics that do not lead to optimal performance in current environmental settings.

References

Barlow, H.B. (1974). Inductive inference, coding, perception and language, *Perception* 3,123-134.

Berretty, P.M., Todd, P.M. and Blythe, P.W. (1997)

Categorization by elimination: A fast and frugal approach to categorization, *Proc. 19th Ann. Conf. of the Cognitive ScienceSociety*.

Campbell, J. (2002) *Reference and Consciousness*, Oxford University Press, Oxford.

Clark, A. (2000) *A Theory of Sentience*, Oxford University Press, Oxford.

Cosmides, L. and Tooby, J. (1994)

Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition*, 50, 41-77.

Elman, J.L. (1993) Learning and development in neural networks: The importance of starting small, *Cognition*, 48, 71-99.

Feldman, J. (1999) Does Vision Work? Towards a Semantic of Perception, in: E. Lepore and Z. Pylyshyn, *What is Cognitive Science,* Blackwell, 208-229.

Jepson, A. & Richards, W. (1992)

What Makes a Good Feature?, in: L. Harris, M. Jenkin (eds.), *Spatial Vision in Humans and Robots,* Cambridge University Press, 1992.

Johnson, M. H., and Morton, J. (1991)

Biology and cognitive development: The case of face recognition. Cambridge, MA: Blackwell.

Keil, F., Smith, C., Simons, D., & Levin, D. (1998)

Two dogmas of conceptual empiricism, *Cognition*, 65, 103-135.

Knill, D.C., Richards, W. (1996)

Perception as Bayesian Inference, Cambridge University Press, Cambridge.

Li, F-F. , VanRullen, R., Koch, C. & Perona, P. (2002) .

Rapid natural scene categorization in the near absence of attention. *Proc Natl Acad Sci USA*, 99(14), 9596-9601.

Marr, D. (1982) *Vision*, Freeman Publishers.

Prinzmetal, W. (1995) Visual feature integration in a world of objects, *Current Directions in Psychological Science*, 4, 90-94.

Rensink, R.A. (2000) Seeing, sensing and scrutinizing, in: *Vision Research* 40, 1469-1487.

Rensink, R. A., & Enns, J. T. (1995)

Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review*, 102, 101-130.

Scholl, B. (2001) Objects and attention: the state of the art, *Cognition*, 80 (1-2),159-177.

Scholl, B., Pylyshyn, Z.W., Franconeri, S.L. (submitted)

The relationship between property-encoding and object basedattention: Evidence from multiple object tracking, *Perception and Psychophysics.*

Todd, P.M. (1999a) Simple inference heuristics versus complex decision machines, in: *Mind and Machines*, 1999.

Todd, P.M. (1999b) Fast and Frugal Heuristics for Environmentally Bounded Minds, in: Gigerentzer & Selten (eds.), *Bounded Rationality*, MIT Press, 1999.

Treisman, A., Gelade, G. (1980)

A feature integrated theory of attention, *Cognitive Psychology*, 12, 97-136.

Treisman, A., Schmidt, H. (1982)
Illusory conjunctions in the perception of objects, *Cognitive Psychology*, 14, 107-141.

Treisman, A. (1999) Solutions to the binding problem : progress through controversy and convergence, *Neuron*, 24, 105-110.

Tversky, A., and Kahneman, D. (1974)

Judgement under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131.