Week 8: BOUNDARY DETECTION, COMPLETION, AND SHARPENING

1) Cooperative-Competitive (CC) Loop
2) Bipole cells, then and now
5) von der Heydt, Peterhans, & Baumgartner, 1984 
   and other physiology/anatomy

Guest appearance: Chaitanya Sai -- “Building a better BCS”

3) Spatial impenetrability
4) Boundary webs
6) Free association
7) Autonomy of perception; cognitive impenetrability

Note: Item 7 is the most important unit of CN 530.

COOPERATIVE-COMPETITIVE NONLINEAR FEEDBACK

G & M, 1985: Use cooperative-competitive nonlinear feedback 
(CC Loop) to complete and sharpen boundaries.

Recall: Perpendicular induction at line ends:

Long-range cooperation wins over locally preferred orientations 
if the latter cannot support a coherent grouping.

(After Kennedy, 1979)

BOUNDARY GROUPING

Each line end induces a “fuzzy band” 
of “almost perpendicular” 
candidate directions for grouping: 
(Q: Where is this fuzzy band in vivo?)

When such activations are aligned across perceptual space, 
cooperative completion of boundaries can occur

FROM FUZZY TO SHARP

Why do we not always perceive multiple, or fuzzy, 
(illusory) contours?

Hierarchical resolution of uncertainty:
1) **Need fuzziness** to initiate grouping.
2) Risk loss of acuity.

**CC LOOP** is a decision process.

CHOOSE: The contextually best orientation

SUPPRESS: Other local orientations

before choice (transient)

after choice (“equilibrium”)
THE BIPOLE PROPERTY

“Completable” perceptual gap can be bridged in one cycle -- no “iteration” is required to determine WHETHER to complete, (though there may be additional “shaping” of completed contours by a cascade or resonance that takes some time to establish.)

Completion via long-range cooperative units

fuzzy “AND” gate

BIPOLES THROUGH THE AGES

Grossberg & Mingolla, 1985
Field, Hayes, & Hess, 1993

Heitger & von der Heydt, 1993
Williams and Jacobs, 1997

WANTED: INTERPOLATION WITHOUT EXTRAPOLATION

YES

NO


Cf. “relatability” -- geometric constraints on which contours get to group with which -- Kellman & Shipley, 1991
Also, Ullman, Zucker, Mumford, ...
IN AND OUT

Bipole “logic” is inward -- Inputs needed at A and B to “activate” in the center.

Candidate cells (V1, V2) send signals outward via long-range, horizontal (axonal) connections.

Q: Is this business about outward propagation really a problem?

Williams: No, because “inward” and “outward” perspectives are equivalent geometrically (as noted in panels 9 and 10).

Mingolla: Yes, but nonlinearities of perceptual selection break the symmetry, so there is work to do . . .

From KSJ3

Gilbert et al.
**HOW TO INTERPOLATE PERCEPTUALLY**

How to get from this... to that?

Plan A: NO*

Plan B: YES

* Functional vs. mechanistic account

**REMEMBER -- MODELERS AND PHYSIOLOGISTS:**

“DIVIDED BY A COMMON LANGUAGE”

**RECEPTIVE FIELD** -- functional

Where on the retina will stimulation yield a response at this (cortical) cell?

**KERNEL** -- structural

Which network cells send inputs directly to this cell?

Kernels are trivial for a modeler to specify, but are generally not observable for a physiologist!

**IMPLEMENTATION OF BIPOLE PROPERTY (1985*)**

1) Compute separate sums, left and right

2) Separate saturation, left and right

3) Threshold the total of both sides above the saturation level for one side

RESULT: Completion occurs only when inputs come from both sides.

GOOD NEWS: Bipole functions as “statistical AND gate.”

* The idea was first published in 1985; the phrase “bipole property” in 1987.

**BASIC BOUNDARY COMPLETION**

In this simulation a single active node on each side of a bipole suffices to initiate feedback. Input nodes are at 15 and 25.

What’s going on here?

Fig. 20 of G & M, 1985a. Note that caption refers to “filling-in,” -- a phrase now reserved for diffusion of featural activity in regions -- as opposed to boundary completion!
DYNAMICS OF BOUNDARY COMPLETION, 1985

Close inspection of this unpublished two-orientation simulation shows:

1) **bipoles** complete **mid-point** of gap **first** (as opposed to growing a boundary out from regions of greatest support.)

2) interior **end cuts** are first generated and then suppressed.

BAD NEWS: LOSS OF ANALOG SENSITIVITY

Likewise for other variables that affect perceptual completion **strength**, such as distance, alignment, etc.

Reason? -- **FEEDBACK**

“DIVIDED BY A COMMON LANGUAGE” REDUX

Any topological “closed loop” in network connectivity yields **FEEDBACK**. (modeler’s usage)

\[ \frac{d}{dt}(blah) = -yaddayadda + somethingelse \pm f(blah) \]

Flavors of cortical feedback:

1) **horizontal connections** within lamina (reciprocal)

2) closed **local circuits**, between laminae, within cortical area

3) “descending” **connections between cortical areas** (most common physiologist’s usage)

WHY FEEDBACK?

Possible computational benefits:
- Smoothing of curves
- Sharpening of completed contours
- Amplification of consistent information
- Synchronization
- Resolution of real/illusory conflicts

Data:
*Visual persistence is longer for illusory contours* than for real contours
  Ming & Meyer, 1988

*Two masking regimes for illusory contours*
  1st: at \( \approx 120 \) msec, 2nd: at “additional 140-200 msec”
  Ringach & Shapley 1996
WHEN LESS IS MORE

More “real” contour . . . weaker illusory contour

after Kennedy, 1979

ILLUSORY CONTOUR STRENGTH

Lesher & M, 1993

GREG LESHER, CNS ‘93

Known aliases:
  Fingers
  The Immortal

First CNS student to contribute to the empirical base of human knowledge via psychophysics

Quotation: “Pop tarts! YESSSS!!!”

Cofounder:

ANALOG-SENSITIVE COMPLETION

proportional to “support ratio” vs. inverted-U

Shipley & Kellman, 1992

Lesher & Mingolla, 1993

cf. Soriano, Spillmann, & Bach, 1994 (shifted gratings)
BCS: COOPERATION AND COMPETITION

few lines, wide spacing

more lines overcome slight inhibition from neighbors

crowding lowers overall effective input to cooperation

cf. hyperacuity

Excitation: monosynaptic
Inhibition: disynaptic, via interneuron

RESULT: “Bipole property”
a single input on one side of a cell does not activate it, but two inputs do (one on each side).

WHY? Inhibitory interneuron saturates before cell that receives excitation along apical dendrite.

G, M & Ross, 1997

DYNAMICS OF COMPLETION, 1997

Initial “outward” activation subsides.

Initial state:

- Time 1
- Time 2
- Time 3
- Time 4

Note: Local connectivity ALSO needed in 1997 version:
Bipole cells need input from both nearest neighbors OR direct bottom-up input to remain active.

If the claim on the previous slide seems a little “fishy” to you, congratulations on your intuition.

Please check the “fine print” of the next panel, which describes an additional condition that needs to be satisfied to get this scheme to work.

Note also that Grossberg and Raizada’s method of implementing the bipole property -- mentioned but not described in detail in following panels -- is somewhat different still. For that one, you simply have to read the papers.
In earlier BCS (Gove et al., 1995) long-range feedback was sharpened in position and orientation by competition stages that were separate from the bottom-up ones.

Result (of this and other model stages): **No analog sensitivity in grouping.**

Model V2 circuits: similar to V1, but with larger receptive fields

**Bipole cells** of layers 2/3 send feedback via layer 6 cells...

...to the same "center-surround" circuit of layer 4 that receives inputs from "earlier" brain areas.

**Results:**

- Analog-sensitive completion
- Context-sensitive criteria for grouping “scaled against” bottom-up inputs

Recent CNS graduate Rajeev Raizada extended the bipole circuit idea in his dissertation research, helping to develop the **LAMINART** model depiction shown in the previous panel.

“**FOLDED FEEDBACK**”:

LAMINART gives an account of how perceptual grouping and attentional mechanisms can share common circuits despite having different computational constraints.

Details (including paper downloads) are available at: [http://www.cns.bu.edu/~rajeev](http://www.cns.bu.edu/~rajeev)

The “new” bipole circuit has several variants; most are still “finicky” w/r/t parameters, input sizes, dynamic range, etc.

Note in particular how inhibitory interneurons are handled.

Nonetheless, present research is accelerating the “convergence” of certain results of adaptive resonance theory (ART) and those of specialized models of vision circuits . . .
BIPOLE AS SPECIALIZED ADAPTIVE RESONANCE

“bottom up” information is spatially NEAR

LONG-RANGE information plays the role of "top-down hypothesis" . . . but is not!

GESTALT GROUPING SIMULATIONS

Proximity:
strengthens horizontal grouping
breaks vertical grouping

Fig 6, G, M & Ross, 1997

CONTEXT-SENSITIVE GROUPING, 1985

G & M, 1985b, Figs 23, 24

CONTEXT-SENSITIVE GROUPING

Solid horizontal bars break vertical groupings
Ross, G & M, 2000
ENHANCEMENT OF COHERENT INPUTS

Ross, G & M, 2000

APPLICATION: IMAGE ENHANCEMENT

input  feature

boundary  filling-in

M, Ross, & G, 1999

DETAILS OF IMAGE ENHANCEMENT

Scale: small  medium  large

boundaries before completion

boundaries after completion

filling-in

large scale bipole:

M, Ross, & G, 1999

BIPOLE ANATOMY AND PHYSIOLOGY

Evidence for bipole connectivity in cortex continues to accumulate.

It would be hard to “fake” anatomical data more supportive of the bipole geometry proposed in 1985 (panel 10 of these notes) than what is shown on the next panel.

Also, the physiology described in the subsequent panels has been developed further than we can cover adequately in this class.
HORIZONTAL CONNECTIONS IN STRIATE CORTEX


von der Heydt, Peterhans, & Baumgartner, 1984

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>V1 response</th>
<th>V2 response</th>
<th>Cells in V2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>Strong</td>
<td>Strong</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>Weak</td>
<td>Strong</td>
<td>(more contrast)</td>
</tr>
<tr>
<td></td>
<td>None</td>
<td>Stronger, with orientationally SHARPER receptive field</td>
<td>YES</td>
</tr>
</tbody>
</table>

Probe location: ●

Response?

- YES
- NO
- NO
- YES
- NO
- YES

Evidence for receptive field:

Also Peterhans & von der Heydt (1988).

MORE ON vdh, P, & B

Horizontally tuned cells:

Probe location: ●

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>V1 response</th>
<th>V2 response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
</tr>
<tr>
<td>None</td>
<td>Weak, with orientationally FUZZY receptive field</td>
<td>Stronger, with orientationally SHARPER receptive field (same cell as above)</td>
</tr>
</tbody>
</table>

Evidence for:

1) orientationally fuzzy end cuts
2) oriented, long-range cooperation.

NOTE: There has been much controversy re: completion in V1 and whether shifted grating contours are “illusory.”

CONTRAST-SENSITIVE GROUPING

Collinear flankers

**enhance** response to near-threshold target

**suppress** response to high contrast target

Polat et al. 1998
Cat Area 17 (V1)

SIMULATION OF POLAT *et al.* DATA

Details: Rajeev Raizada: [http://www.nmr.mgh.harvard.edu/~raj/](http://www.nmr.mgh.harvard.edu/~raj/)

PARALLEL STUDIES

Psychophysics

Physiology

Kapadia, Ito, Gilbert, and Westheimer 1995

Figure on previous panel: Left side -- This plot summarizes psychophysical data from trials of one observer in a tilt illusion experiment The observer controlled the orientation of a central line segment displayed in juxtaposition with two symmetrically flanking lines whose actual orientation was 5 degrees from the vertical The observers task was to make the central line appear vertical The data show regions for which attractive (blue) and repulsive (red) tilt illusions were obtained. Right side -- Summary of excitatory and inhibitory zones of influence for flanking stimuli on single cells of vertical orientational preference whose receptive fields are centered in the diagram.
NATURAL IMAGE STATISTICS

Probability of edge co-occurrence:
(not to scale)

Geisler, Perry, Super, and Gallogly 2001

BRIGHTNESS BUTTONS

Day & Jory; 1975 Kennedy, 1979

STRANGE SUMMATION

I'd like to thank the guy who wrote the song
That made my baby fall in love with me
Who put the bomp In the bomp bah bomp bah bomp
Who put the ram In the rama lama ding dong
Who put the bop In the bop shoo bop shoo bop
Who put the dip In the dip da dip da dip
Who was that man, I'd like to shake his hand
He made my baby fall in love with me [Yeah]

Darling, bomp bah bah bomp, bah bomp bah bomp bomp
And my honey, rama lama ding dong forever
And when I say, dip da dip da dip da dip
You know I mean it from the bottom of my boogity boogity boogity shoo

Ooh...ooh...ooh...ooh...ooh...ooh...ooh...ooh...ooh...ooh...Ooh...ooh...ooh...ooh...ooh...ooh...ooh...ooh...ooh...ooh...

Song: Who Put The Bomp (In The Bomp Bah Bomp Bah Bomp?) ~ 1961 ~ 2:43 Artist: Barry Mann LP/CD: Not available ~ ABC Writers: Barry Mann ~ Gerry Goffin

Transcribed by: Char Star
CONTRAST AT LINE ENDS

OFF-center cells: Best response at line ends
ON-center cells: Best response at line flanks

Cf. Minguzzi 1985

Note that once again “lateral inhibition” fails to explain . . .
So: Who put the induced brightness at the end of the line?

AS SUGGESTED BY . . .

Murphy & Sillito, 1987

Some cells of LGN exhibit something that looks pretty darned like endstopping -- firing only if a line ends in a receptive field’s excitatory center.

The “experimental side” below refers to a cell whose length tuning looked as above before its corresponding V1 hemisphere was aspirated.
SPATIAL IMPENETRABILITY

Why do not all collinear contour fragments complete across intervening contours?

Mechanism for “spatial impenetrability:”

*Yes, dipoles AND bipoles

* You say “TO - MAY - TO,” I say “TO - MAH - TO,”
you say “DI - POLE,” I say “BI-POLE.”
Let’s call the whole thing off!

MECHANISM FOR SPATIAL IMPENETRABILITY

Horizontal signals are everywhere, but form a majority only at line ends.

The left receptive fields of two horizontally tuned bipole cells are stimulated by vertical and horizontal (end cut) signals.

Although horizontal signals can excite the upper bipole, their effect is cancelled by the greater number of inhibitory vertical signals.

Within the lower bipole’s field, excitatory signals from end cuts prevail.

Note: spatial impenetrability includes options on: “occlusion,” “transparency,” “figure-ground.” “border ownership,” “T-junctions,” “amodal completion,” extrinsic and intrinsic line ends ...
When does a discrete 2-D texture result in a percept of a smooth surface?

Gibson, 1950

Compression of elements as a function of slant -- angle of surface to line-of-sight:

Human judgments of slant-from-texture in images are not so good.

Might curvature-from-texture judgments be better? Why?

viewing distance (circle diameters): 1 2 3 4

Todd & Akerstrom, 1987, Fig 1
**CURVATURE-FROM-TEXTURE**

Variations of size and compression of texture elements can give an impression of curvature in depth:

![Curvature-from-texture example](image)

**THE UNITS OF CURVATURE FROM TEXTURE**

Note: This effect is reduced by reproduction artifacts.

*Todd & Akerstrom, 1987, Fig 4*

---

**OC FILTER RESPONSE**

Complex cells of BCS “front end” sample local image orientations:

![OC filter response](image)

**BOUNDARY WEBS**

*Boundary webs* are formed as *CC Loop* enhances coherent orientation signals:

Boundary signals can be “fat,” in the sense of covering a region that is wide in the direction perpendicular to the orientational preference of active nodes.
**BCS-DERIVED PREDICTION**

Todd & Akerstrom, (1987) devised a psychophysical test of a prediction they derived from BCS theory:

Boundary webs are generated at *several scales*, whenever elements vary in size and are aligned in orientations predominantly perpendicular to radii from center of displays. More *variation in scale* implies more *perceived curvature*.

**Test:** Create displays with optically incorrect depictions of textured surfaces, but that engage the same BCS processes as those with proper textures.

![Constant compression square](image1.png) ![Constant compression elongated](image2.png)

Correlation of a measure based on BCS model prediction with observers' curvature judgments = .99.

---

**MORE ON BOUNDARY WEBS**

Note that Pessoa et al. 1995 needed a “1-D” implementation of boundary webs for their brightness modeling.

Levin Kuhlman, current CNS student, is working on extensions of the boundary web idea for shape from texture.

---

**RANDOM ACCESS**

The next several panels contain some “unclassifiable” fragments re: BCS.

---

**COMPLETION GENERATORS**

What are the generators for long-range completion?

![Excluded shapes](image3.png)

Cf. BCS account of “end cuts” -- why the terminators of thin lines can act as if an approximately perpendicular edge is attached.
**UNITS OF FILTER ACTIVITY**

- **Units** of early physiological activity:
- **Packets** of nearby orientations

Everything important about BCS design follows from a commitment to treat the inherent early uncertainty of packets of orientations as a “feature” rather than a “bug,” with respect to network processing.

Among the implications are that units like *edges* or *lines* emerge as the “end result” of autonomous “middle vision” processing, not as the *inputs* to “higher level” processes that were found by “low level” filters and thresholds.

Lines and edges are just a few of the possible “outputs” of BCS processing. Others include textures and smooth shading (boundary webs). (Binocular generalizations of BCS are needed for disparity, depth, occlusion, transparency, etc.)

Lines, edges, textures, and shading thus have the same “ontological status” as possible outcomes of BCS dynamics; they are not handled by independent modules.

---

**HIGHLY NONLINEAR PROCESSING**

Consider “isolated” dots

- Small scales build a boundary:
- Large scale activity disappears via *cross-orientation inhibition*

Qualitatively different results from “form-sensitive” processing

---

**Non squitur Q1:** Where and when (and why?!) in the visual system do representations of perpendicular orientations compete?

**Q2:** Where might they “coexist” in close proximity?
DIRTY LITTLE SECRETS

A number of facts about illusory contours do NOT fit into the BCS story told so far. Some can be handled by FACADE (a successor of BCS/FCS), others not.

A number of fundamental challenges shown in subsequent panels and were published in:


GILLAM, 1987: FIGURE 30.5

“Subjective contours are stronger when the presence of a perceived occluding surface results in closure of the rectangles.”

GILLAM, 1987: FIGURE 30.6

“Subjective contours are stronger when the lines interrupted by collinear gaps are unrelated otherwise.”

“Subjective contours are stronger when the orientation of the gap sequence is unrelated to the principal axes of the lines in which the gaps occur.”
TSE STRIKES AGAIN

Peter Tse has recently devised a number of counterexamples in 3-D scenes to the idea that contours that are “relatable” in the image plane can group.

Soon thereafter, Grossberg and Swaminathan generalized bipole cells to cells that are sensitive to particular gradients of depth (“near to far” vs “far to near” as well as collinearity.)

Note: Their paper is not for the faint of heart.

“COMPUTATIONAL THEORY” OF BCS

Core idea of BCS:

BCS is an architecture for getting a fixed-weight network to discover coherence and departures from coherence of whatever sort relative to the degree of variability “specified” by each scene itself.

The units of computation are therefore built “on the fly” — out of the interaction of the structure of light input and physiology.

Neither the world* nor the brain “contains” the units.

* Hence, there’s no point in building an internal model of the world.

AND MOREOVER . . .

Yes, many, perhaps even most of what early and middle primate vision needs to do is relatively “straight[feed]forward” for many scenes, but:

— even simple-appearing tasks may have more contextual modulation than has often been assumed

(See: local recurrent circuits and Bullier lab work)

— the “high end” of human competency in form perception is clearly highly “heterarchical,” “reentrant,” “contextual” or ________ [insert your favorite jargon for “gestalt-driven” here].

BRAIN TEASERS

1) The box to the right contains six symbols, each formed by a connected set of lines. Each symbol is free of any kind of noise and stands for a letter of the English alphabet. What two word phrase do the symbols depict? (The fate of human civilization depends on your answer.)

Yes! No!

2) Does the “1-D” luminance profile depicted on the right contain an edge? ____________

The cat

Yes! No!
Don’t be put off by the cheesy trick of problem 2 on the previous panel. It points to the difficulties inherent in quantization and context-dependent thresholds, and if you could solve the problem that it points to -- i.e., by inventing a machine that would autonomously generate the correct answer in each case -- you would have made a stride that has eluded computational researchers for several decades.

Grossberg’s (1973) “quenching threshold” notion is suggestive, because the value of the threshold depends on particular inputs, as well as on model parameters, but getting this idea to “work” in form vision is not straightforward.

AUTONOMY OF PERCEPTION:
COGNITIVE IMPENETRABILITY

Let’s talk about the effects of cognitive expectancies on perception.

Sure, they exist.

Sure, they profoundly affect what we experience.

RECALL: Our earliest model macrocircuit (from Week 1!) contains down-pointing arrows from ORS to BCS and FCS.

BUT . . .

Rod appears behind sail notwithstanding! Fig 18, Kanizsa & Vicario, 1968
IN THEIR OWN WORDS, I

The following panels are an extensive quotation (all of Section 4) from: Kanizsa, G. and Luccio, R. (1987). *Formation and categorization of visual objects: Höffding’s never refuted but always forgotten argument.* *Gestalt Theory,* 9(2), 111-127.

“The main point” of CN 530 is presented here more succinctly and elegantly than in any of the material I have presented or will present. Inevitably, there are a few passages in what follows that will be obscure, as the authors assume familiarity with a certain literature, to say nothing of the earlier sections of their paper (!), but try to read past these glitches to the core idea.

This discussion brings us directly to what seems to us to be theoretically the most important objection made about the content of our article. This regards the legitimacy of the basic assumption underlying our argument; in other words, the distinction we make between two different moments,

IN THEIR OWN WORDS, II

or aspects, in the perceptual process: the moment of segmentation of the visual field, which we called “pre-categorical” or primary, and the secondary aspect of cognitive processing of the autonomously segregated perceptual units. By “primary process” we mean the silent phase of organization which, beginning with the proximal stimulation, leads to the articulation of the visual field into those objects we find around us when we look about. But in our opinion the result of the primary process is not limited (as many cognitive theorists suppose) to the extraction of elementary features (lines, angles, contours, orientation, and so on) which must then be assembled and ordered by inferential or probabilistic procedures. In our opinion, what the primary process produces is rather the complete organization of the visual objects, endowing them not only with colour, form, and size, but also with 3-dimensionality, kinetic properties, and expressive qualities. It is to these perceptual units (understood as visually self-sufficient Gestalten), that the operations of categorization, recognition, inference, judgment and so on, are applied,

IN THEIR OWN WORDS, III

and these operations are obviously subsequent, from a logical point of view. The pre-categorical moment, in our opinion, is equivalent to seeing in the strict sense of the word, while it is through cognitive operations that visual objects acquire their meaning.

In previous articles, we have repeatedly stated that such a distinction is logically necessary, and scientifically productive; we believe it to be one of the crucial problems in experimental phenomenology, and we have continually warned against the persistent widespread temptation to ignore it, or to explain it away (KANIZSA 1969, 1975, 1985; KANIZSA & GERBINO 1982). In fact, since the two moments co-exist and are practically indistinguishable in a concrete perceptual act, there are some people who claim that the two processes lie along the same continuum, making it not only impossible, but pointless, to try to distinguish them clearly, one from the other. It was in order to refute this very argument that we offered numerous examples in our work to demonstrate that this
IN THEIR OWN WORDS, IV

distinction can, and therefore must, be made. These demonstrations show the autonomy of the primary process, and its resistance to modification through cognitive operations, whether these take the form of hypothesis, reasoning, inference, interpretations or knowledge derived from past experience. Here is the crux of the matter. It is difficult to find anybody prepared to deny that seeing and thinking are different activities, both logically and phenomenally speaking, but there is no agreement about the nature of their relationship. To be more precise, there is no agreement about the answer to the following questions: do thought activities influence the processes which leads to our seeing? If so, to what extent? Do the processes in the visual system which produce our articulated visual world take place autonomously, according to their own laws, or are they concept-driven. In other words, is seeing conditioned and directed from above by the hypotheses, anticipatory schemata and judgments of the perceiver or even (as it was fashionable to think in the days of the now-discredited New Look) by his expectations, desires, emotions, and needs? Is seeing as such, independent from thinking, or is it influenced by it in some way? On attempting to understand what our fellow scholars think about the problem, a strange fact emerges. On the one hand, most scholars admit that in order for any identification or interpretation of perceptual data to take place at all, there must first be a percept to identify. On the other hand, many of these authors, after having explicitly stated that they accept the cogency of Höllding’s argument, seem then to forget about it, for no clear reason. Sooner or later they re-introduce the idea of a guidance from above by the cognitive processes (in the form of schemata, sets, hypotheses, and suchlike) of the process by which the percept is formed. Thus, the cognitive processes are regarded as having a decisive influence on the final articulation of the visual field itself.

We still believe that our line of reasoning is completely valid, not only to avoid a logical contradiction, but above all because there are too many cases in which the perceptual result is in open conflict with the expectations and knowledge of the perceiver. We think that the visual system is a closed system, inaccessible to influence by other psychological systems, and that the perceptual processes are, as WALLACH (1949) says, protected from interference from intellective sources.

IN THEIR OWN WORDS, V

AND, MOREOVER

On a related note, a festschrift for Kanizsa published in the Giornale Italiano di Psicologia, Vol. 20, No. 5, 1993 contains an article by William Epstein entitled: On seeing that thinking is separate and on thinking that seeing is the same.

That article’s abstract reads:

I examine the thesis that the process underlying seeing is a process of inductive inference or thinking. I contend that the thesis is not convincing. My argument is advanced through consideration of a diverse set of phenomena ranging from the relatively primitive such as Mach bands to the putatively more elaborate attainments such as perceiving of affordances. Along the way I give special attention to the implications of the perceptual constancies and putative demonstrations of cognitive penetrability. I conclude that the [sic] seeing and thinking are served by different processes and I endorse a characterization of perception that is inspired by J.J. Gibson’s stance toward information and the connectionist (neural network) stance toward process.
While sometimes, and in some ways, expectations do influence what you “see” (in the casual, everyday sense of the word) there is a sense in which you “see” things (in a technical sense of the word, as used by scientists who study visual perception) DESPITE, not because of, your expectations, and notwithstanding all the “top-down,” inter-area feedback in your brain. This kind of seeing involves a process of autonomous organization that is neither “bottom up” [heirarchy of detectors] nor “top-down”.

Moreover, the study of the “gestalt” processes that underlie such “seeing” (from the point of view of nonlinear dynamical networks) provides an alternative style of explanation to three classical extremes -- which remain very much the only ones considered by many people in machine vision to the present day:

1) This visual competency can be understood by the application of sound engineering principles and clever algorithms (filtering, measurement theory, etc.), or . . .

2) This visual competency is entirely a matter of inference (“unconscious” or Bayesian) or categorization (learned associations, “thinking”, “problem solving”), or . . .

3) This visual competency is a “high level” process that we will never understand in our lifetimes, so let’s confine our attention to “vision” in controlled environments (known illuminants and reflectances, etc.)

From a perspective N years after Kanizsa and Luccio, we can make more refined statements about the interplay of “sensory” factors, “expectations,” and organization in vision.

See Grossberg et al.’s recent body of work on LAMINART with specific reference to interplay of developmental, attentional, and contextual factors in V1-V2 circuits.

For a different, but related -- and provocative -- view of a recent reformulation of “preattentive” and “attentive” vision, see:


View from the top: hierarchies and reverse hierarchies in the visual system.

Hochstein S, Ahissar M.