2008

Contextual Influences in Visual Processing

Tai Sing Lee
Carnegie Mellon University, tai@cnbc.cmu.edu

Recommended Citation

.

Contextual Influences in Visual Processing

TAI SING LEE
Computer Science Department and Center for Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA, USA

Synonyms

Surround influence; Contextual modulation; Local global interaction; Extra-classical receptive field modulation

Definition

Vision is the analysis of patterns in visual images with the view to understanding the objects and the physical processes in the world that generate them. Locally, visual patterns are highly ambiguous and subject to multiple interpretations. Image structures surrounding the pattern being analyzed can provide additional constraints or context to disambiguate the interpretation. The resulting ▶contextual influences are ubiquitous in visual perception and manifest at the neuronal level as the modulation of the activity of neurons by image structures outside their ▶classical receptive fields.

Characteristics

The study of contextual influences in visual processing has a long history in psychology and neuroscience [1]. Investigations of these effects in the visual system have focused on the ▶modulatory effect on the activity of a neuron by image structures outside its localized ▶receptive field. The classical approach employs the simplest stimuli such as bars and sinusoidal gratings to probe the interaction between the stimuli presented inside and outside a neuron's classical receptive field. A prevalent finding is that neurons in both the ▶primary visual cortex (striate cortex, V1) and the ▶extrastriate cortex exhibit ▶feature contrast enhancement, i.e., the cells respond better when the stimulus attributes in the area surrounding their receptive fields, such as bar orientation, are different from those inside their receptive fields (Fig. 1a).

Recent approaches seek to understand the neural basis of the perceptual interpretation of the local receptive field stimulus by changing the global image context (Fig. 1b). With this approach, a number of neural correlates of perception have been revealed, providing insights into the representation of subjective perceptual experience in the brain.

Contextual Influences in the Primary Visual Cortex

Neurons in the primary visual cortex receive converging input from the ▶lateral geniculate nucleus (LGN). A neuron's classical receptive field, also known as the minimum responsive field, is the part of visual space in which the presence of appropriate features can excite the neuron. By definition, stimulating the visual space outside a neuron's classical receptive field cannot evoke a response. Modulation of neuronal activity by surround stimulation can be observed, however, only when the neuron is responding to a stimulus presented to its receptive field. This modulation is called the non-classical or ▶extra-classical receptive field effect. Such effects have been considered neural manifestations of contextual influences in visual perception.

A variety of extra-classical receptive field effects have been identified. A commonly reported phenomenon is called ▶surround suppression: the response of a neuron to an oriented bar or grating within its receptive field is suppressed when stimuli are simultaneously introduced to the surrounding area outside its receptive field. There are several types of surround suppression effects, mediated by a number of ▶local circuits as well as ▶recurrent feedback circuits [2]. The early phase of surround suppression is fast and is not sensitive to the exact parameters of the surround stimuli. However, the later phase of surround suppression is stimulus-specific. Simply put, while the neuron can detect the presence of stimuli in the surround immediately, its sensitivity to the precise nature of the surround stimulus or global context takes time to develop. The onset delay of this sensitivity varies considerably depending on the types of the stimuli and the spatial extent of the contextual stimuli.

One well-known stimulus-specific surround suppression, observed with an onset delay, is called ▶iso-orientation suppression. In this phenomenon, a neuron's response is stronger when the orientation of the surround
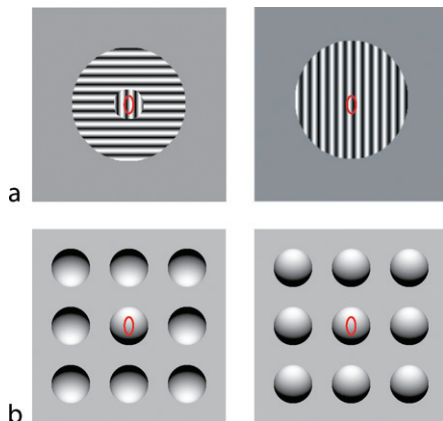
Contextual Influences in Visual Processing.
Figure 1 Stimuli used in contextual modulation studies.
(a) Classic center-surround stimuli that have
been typically used in neurophysiological studies on
iso-orientation surround suppression [3]. Neurons tend
to respond better when the orientations of the center and
surround gratings are different (*left image*) than when
they are the same (*right image*). The red ellipse outlines
the spatial extent of the receptive field of the neuron.
A similar effect observed in a larger center patch with
a significantly longer delay is called figure
enhancement [4]. (b) Surround context can change
the perceptual saliency of the receptive field stimulus.
The receptive field stimulus is said to pop out from the
background on the left image, but not on the right
image. This pop-out phenomenon depends on
3D interpretation of the stimulus elements. Early visual
neurons' activity is correlated with the perceptual
saliency of this pop-out phenomenon [5].

stimulus is different from that of the center receptive field
stimulus than when the orientations are the same. When
the receptive field stimulus is a bar, iso-orientation
suppression emerges at about 10 ms after the onset of
the response to the receptive field stimulus [3]. When the
receptive field stimulus is a part of an oriented texture
region significantly larger than that of the receptive field,
the later part of the neuron's response is inversely
proportional to the size of the region – the larger the
region, the smaller the response. This results in a relative
enhancement of response when the neuron's receptive
field is inside a smaller region than when it is in the larger
background region. Interestingly, the enhancement is
uniform across the surface of a compact region, with
a sudden drop off at the region's border. Hence, it has
been proposed to be a signal that could highlight a figure
against its background and is called the ▶figure
enhancement effect [4]. According to most studies,
the onset delay of this figure enhancement effect is
proportional to the size of the region. When the receptive
field is at the center of a region that is six times larger
than its size, the onset delay is typically 40 ms relative to

response onset on the average. The figure enhancement
effect is more general than iso-orientation suppression as
it has been observed in studies with motion or shape from
shading stimuli without any orientation contrast between
the receptive field stimulus and the surround [4,5].

Functionally, both iso-orientation suppression and
figure enhancement can serve to enhance stimulus
feature contrast, resulting in an increase in ▶perceptual
saliency of the representation of less expected or
surprising visual events to facilitate further processing.
Indeed, it has been demonstrated that this response
enhancement is directly proportional to perceptual
saliency of the visual pattern, as measured in terms of
the reaction time for target detection, and it is disso-
ciable from luminance contrast or orientation contrast in
the stimulus (Fig. 1b) [5]. The broader spatial extent
and the longer onset latency of the figure enhancement
effect suggest that, while iso-orientation suppression
might be mediated primarily by inhibitory ▶local
circuits, the figure enhancement or perceptual saliency
effect likely involves additional long range facilitation
circuits including recurrent ▶feedback from the extra-
striate cortex, as suggested by both anatomical and
deactivation studies.

Surround interaction can be quite complex and can
vary according to the luminance contrast or the spatial
scale of the stimuli. While surround modulation tends to
be suppressive when the luminance contrast of the
stimulus is strong, it can become facilitatory when
the luminance contrast is weak. Neuronal ▶adaptation,
well known in the ▶retina and LGN, is sensitive to the
absolute luminance and luminance contrast levels in
the entire scene. In a dark and low-contrast environ-
ment, retinal and LGN neurons are known to expand
their receptive fields temporally and spatially with
a simultaneous increase in their sensitivity gains. Such
a strategy serves to optimize feature detection in the
presence of noise. The contrast dependence in surround
influence likely results from V1 neurons inheriting and
extending these adaptation or optimization strategies.

Perceptual computations supported by the complex
machinery in V1 likely go beyond feature detection and
feature contrast enhancement. From a computational
perspective, contextual effects reflect the influence of
computational constraints, realized by neuronal con-
nectivity and interaction, necessary for solving visual
inference problems. Surround interaction can bring
in contextual information to improve local estimates
of visual cues, as evident in the observations that
▶orientation tuning curves and ▶disparity tuning
curves tend to sharpen over time during the analysis
of each visual image. The ▶retinotopic organization,
the connection infrastructure, and the tuning properties
of neurons in V1 make it ideally suitable for supporting
a variety of visual computations. One such computation
is the grouping of edges into contours and features into

coherent regions. There is some evidence that V1 plays an important role in this computation to be discussed below.

First, the activity of some V1 neurons is enhanced if the surrounding bars outside their receptive fields line up with the bar presented within their receptive fields to form a longer contour (Fig. 2a).

Moreover, some V1 neurons respond to the ►subjective contour of a ►Kanizsa figure, even when no feature is presented to their classical receptive fields (Fig. 2b). There is also evidence that neurons can interpolate contours across the blind spot or behind an occlusion. Furthermore, collinear contours have been found to induce neuronal synchrony in V1 neurons of the same ►orientation selectivity. Recently, it was also found that neurons with different orientation tunings, when stimulated simultaneously by curved contours, also exhibit an increase in synchrony or ►effective connectivity, as revealed by multi-electrode recordings [6]. This dynamic change in effective connectivity between neurons as a function of stimulus is suggestive of a mechanism for ►contour completion.

In addition, similar changes in effective connectivity have also been observed among spatially disjoint

Contextual Influences in Visual Processing.
Figure 2 Neurophysiological evidence of contour completion in V1. (a) Oriented bars in the surround (*left image*), when aligned with the receptive field stimulus to form a contour, can increase a cell's response to its receptive field stimulus (*right image*) (Kapadia, Westheimer and Gilbert 2000). The *red ellipse* outlines the spatial extent of the receptive field of the neuron. (b) The subjective contour of a Kanizsa's illusory square can evoke response in a V1 neuron even when no stimulus feature is present in its receptive field (*red ellipse*) (Lee and Nguyen 2001). The subtle addition of thin circles on the right image changes the perceptual interpretation of the image from a white square occluding four black circular disks, with a vivid subjective contour over the receptive field (*left image*), to that of a white square in a background visible through four circular windows on a white wall in front (*right image*).

►disparity selective neurons when the 3D depth plane of the random dot stereogram stimulus intersects with the cells' optimal disparity tunings. This process appears to contribute to the gradual sharpening of the neurons' disparity tunings over time, providing a plausible mechanism for improving local estimates of visual cues based on global context. Such cooperative or mutual facilitatory mechanisms might also contribute to *surface association* by increasing the firing rates of the neurons analyzing different parts of the same visual surface simultaneously. The resulting enhanced and correlated activities, partly represented in the figure enhancement effect, can highlight the relevant coincident features in visual input as a group to provide a stronger drive for downstream neurons in the extrastriate cortex to learn explicit representations for higher order features and structures.
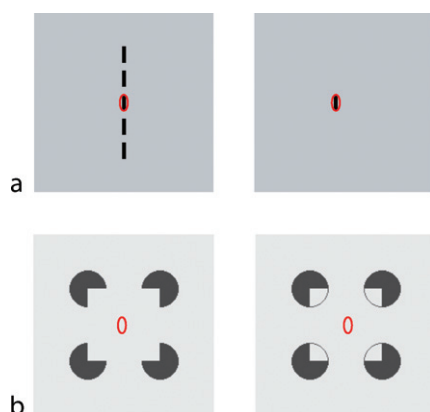
Contextual Influences in the Extrastriate Visual Cortex

The extrastriate cortex, downstream from the striate or primary visual cortex, is partitioned into many different visual areas. The feature contrast enhancement effect observed in V1 is also prevalent in extrastriate visual areas, expressed in the respective feature dimensions that neurons in those areas are tuned to. In area ►MT (medial temporal), for example, the motion of surround stimuli has been shown to significantly modulate the response of a neuron to moving stimuli presented to its receptive field. The response of the neuron is suppressed when the direction of surround motion is the same as the motion detected in the neuron's receptive field. This is analogous to the iso-orientation suppression in V1 but in the motion domain. In addition, the disparity-tuned MT neurons also experience iso-disparity suppression.

The extrastriate cortical areas, however, exhibit some additional contextual effects that are rarely observed in the striate cortex. Many of these *new* contextual effects are concerned with the inference of 3D surfaces, their occlusion and depth ordering relationships, also known as ►figure-ground organization. In MT, it has been shown that the responses of direction-selective neurons to a motion stimulus are sensitive to the figure-ground context defined by the surrounding surface depth structures in a way that is consistent with ►Barber Pole illusion [7].

Several lines of evidence suggest that the computations underlying figure-ground segregation and 3D surface inference might start in visual area V2. First, a significant fraction of V2 neurons (and a small number of V1 neurons) have been shown to signal whether their receptive fields are at the left border or the right border of a figure in an image regardless of the polarity of contrast at the border (Fig. 3a).

A left-border-preferring neuron carries the information that the border within its receptive field belongs to (or is *owned* by) the surface or region to its right [8].

A complementary, right-border-preferring neuron exists at the same location, and both neurons could form a push–pull pair for every border orientation. The activity of a set of such pairs of ►border-ownership neurons in various orientations along the border of each region in an image can encode the depth-order relationship between the different image regions or inferred surfaces. Secondly, it has been found that neurons in V2, but not in V1, are sensitive to the mismatch in features between the images from each eye at visual locations where one surface occludes another [9]. The emergence of sensitivity to this surface occlusion cue in V2, known as the ►Da Vinci stereo, further suggests that 3D surfaces and their occlusions are explicitly represented in V2. The figure-ground context made explicit in V2 could feed back to constrain the computation in V1, resulting in, for example, the figure enhancement effect. However, it should be noted that the figure enhancement effect in V1 has not been conclusively demonstrated to depend solely on figure-ground organization.

The perception of surface attributes such as brightness, shading and color depends very strongly on the interpretation of the underlying 3D surface geometry and the illumination direction in the visual scene. Two observations suggest that these surface attributes might also be inferred and represented in V2 because of the dependence of such inference on 3D surface interpretation. First, the neural correlate of ►shape-from-shading pop-out, a perceptual phenomenon that crucially depends on 3D surface interpretation, is observed in V2 but not in V1 pre-attentively [4]. Second, the neural correlate of the ►Cornsweet-O'Brien illusion, an illusion in perceived brightness induced by edge contrast, which ultimately can be traced back to surface geometry and lighting direction interpretations in natural scenes, is observed in V2 but not V1 [10] (Fig. 3b). There has been, however, some evidence for brightness representation in V1 [1]. It is possible that the construction of brightness representation is a gradual and distributed process, computed first at V1 based on surround luminance contrast, but achieving a more abstract and invariant representation in V2 as the 3D surface representation is made explicit. In general, neuronal activities tend to become progressively more abstract and more correlated with our subjective perceptual experience as one moves up the visual hierarchy.

In addition to global image structures, behavior, task demands and memory are also known to provide strong contextual information to influence visual perception and object recognition. ►Attentional modulation of neuronal responses has been widely observed and studied in the extrastriate cortex (see ►Visual Attention). Attentional effects in V1 are subtle and observable mostly when visual scenes are cluttered or in tasks that demand considerable spatial attention at precise locations
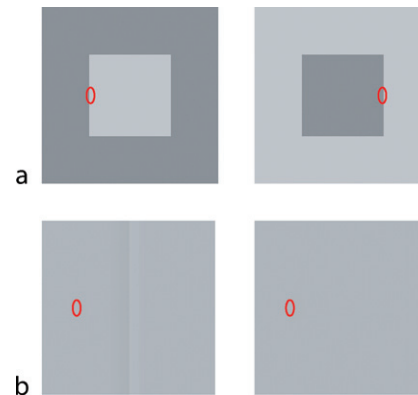
Contextual Influences in Visual Processing.
Figure 3 Neurophysiological evidence of surface inference in V2. (a) A left-border cell will respond more strongly when its receptive field (*red ellipse*) is analyzing the left border of a figure (*left image*) than when it is analyzing the right border of the figure (*right image*), even when the visual pattern on the receptive field and in its immediate surround is identical [8]. This class of cells, observed primarily in V2, is said to convey information about border-ownership or surface occlusion. (b) In the Cornsweet-O'Brien illusion, the presence of a contrast edge can change the perception of the brightness of a region. A V2 neuron that prefers darkness over brightness would respond better to the perceptually darker region (*left image*) than to the perceptually brighter region (*right image*) even though the physical luminance of the receptive field stimulus in the two cases is exactly the same [10].

such as the task of tracing a curve. Beyond V2, extrastriate neurons tend to have large receptive fields. Attentional modulation in neurons of these higher areas typically manifests as the selection of one relevant feature over the others present within their individual receptive fields. Attention can be voluntary, as in selecting a particular spatial location (spatial attention) or a particular feature (feature attention) in the receptive field for further analysis. But it can also be reflexive, driven or captured by the saliency of the stimuli computed automatically in early visual areas. The variety of ►feature contrast and perceptual saliency effects observed in V1 and in the extrastriate cortex likely serves as a part of this reflexive attention mechanism. Recently, higher-order non-spatial contextual effects, such as context familiarity and associative memory, have also been shown to modify the activities of neurons in ►inferotemporal cortex (IT) and medial temporal (MT) respectively.

From the perspective that vision is a process for inferring the various underlying environmental causes of visual patterns such as the 3D geometry of surfaces, the identities of objects and the illumination direction in the scene, the extrastriate areas in the visual hierarchical system might be conceptualized as modules that provide

explicit representation of these decomposable causes. Each extrastriate module furnishes an explanation on some aspect of the visual scene. The inference of the underlying causes involves integration of information across space and over time by neurons in the higher-order visual areas, which in turn provide a variety of context in which visual processing in the earlier visual areas can be refined. V1, with its neurons arranged in a spatially precise ►retinotopic map and endowed with small localized receptive fields capable of representing fine details in images, might serve as a *high resolution buffer* at which all the causes are combined together to synthesize an explanation of the visual input represented explicitly there. These interactive computations can bring about a very rich variety of contextual influences in V1 and the extrastriate cortex. The long latency of many of the contextual effects observed suggests that a substantial amount of recurrent interaction could have taken place. Computations involving such recurrent interaction will predict the *simultaneous* emergence of the perception-related signals in many visual and decision areas in the brain.

References

1. Albright TD, Stoner GR (2002) Contextual influences on visual processing. Annu Rev Neurosci 25:339–379
2. Angelucci A, Levitt JB, Walton EJ, Hupe JM, Bullier J, Lund JS (2002) Circuits for local and global signal integration in primary visual cortex. J Neurosci 22 (19):8633–8646
3. Knierim JJ, van Essen DC (1992) Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol 67:961–980
4. Lamme, VAF (1995) The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 15:1605–1615
5. Lee TS, Yang CF, Romero RD, Mumford D (2002) Neural activity in early visual cortex reflects behavioral experience and higher order perceptual saliency. Nat Neurosci 5:589–597
6. Samonds JM, Zhou Z, Bernard MR, Bonds A (2006) Synchronous activity in cat visual cortex encodes collinear and cocircular contours. J Neurophysiol 95 (4):2602–2616
7. Duncan RO, Albright TD, Stoner GR (2000) Occlusion and the interpretation of visual motion: perceptual and neuronal effects of context. J Neurosci 20:5885–5897
8. Zhou H, Friedman HS, von der Heydt R (2000) Coding of border ownership in monkey visual cortex. J Neurosci 20:6594–6611
9. Bakin JS, Nakayama K, Gilbert CD (2000). Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. J Neurosci 20:8188–8198
10. Hung CP, Ramsden BM, Roe AW (2007) A functional circuitry for edge-induced brightness perception. Nat Neurosci 10(9):1185–1190