Neural and perceptual responses to a visual stimulus can be suppressed by the addition of both spatially overlapping and spatially adjacent contextual stimuli. We investigated the temporal characteristics of these suppressive interactions in psychophysical contrast masking experiments using Gabor and grating stimuli with a spatial frequency of 4 cycles per degree. We found that the time course of masking strongly depended on mask orientation. Most interestingly, masking by a spatially overlaid, iso-oriented mask was strongest when the target was presented immediately before or immediately after the mask. This masking was transient, presumably caused by the neural responses to mask onset and offset. Adding a surround to the mask modulated the backward masking effect, but only when the target and the central mask were iso-oriented. Our results provide evidence for a surround suppression mechanism that affected the transient responses to the mask onset, but not the responses to the mask offset. Together, these results demonstrate how the effects of spatial context in visual processing critically depend on stimulus timing.

Keywords: temporal processing, surround suppression, contextual modulation, Gabors

faster than surround inhibition (Smith, Bair, & Movshon, 2006).

We were interested in comparing the temporal characteristics of psychophysical iso- and cross-orientation masking. The main questions of the present experiments were as follows: (1) what are the time courses of cross-orientation and iso-orientation masking and how do they differ; (2) what is the timing of surround modulation, and how do its temporal characteristics differ between cross-orientation and iso-orientation masking?

To measure the temporal characteristics of overlay masking and the effect of surround suppression, we combined the two-mask paradigm with a temporal masking approach. As the transient responses to stimulus onset and offset are known to be important for contrast detection (Breitmeyer & Julesz, 1975; Mitov, Vassilev, & Manahilov, 1981; Snowden, 2001; Stromeyer, Zeevi, & Klein, 1979; Tolhurst, 1975), we chose to use a brief target and a longer mask duration. This way, it is possible to separate between the effects of onset and offset responses in masking. We found that the time course of overlay masking is strongly orientation dependent. Surprisingly, iso-orientation masking was maximal when a briefly presented target immediately preceded or followed a spatially overlapping central mask, i.e., with temporally non-overlapping presentation. Addition of a surround greatly reduced iso-orientation masking, but the reduction only occurred in the backward masking condition, where the target preceded the mask and its surround. We link this effect to physiological results on surround suppression and propose that surround suppression (1) affects the transient responses to stimulus onset and (2) can precede the interference of the iso-oriented overlay mask with the target even when the target presentation precedes the mask and its surround.

**Methods**

**Observers**

The first author and three paid students served as observers (aged 21–29 years). The student participants were naïve to the purposes of the experiments. Two of the observers participated in all conditions. All observers had normal, uncorrected vision, reaching a visual acuity of 1 (corresponding to 20/20) or better in at least one eye as tested with the Freiburg visual acuity test (Bach, 1996). Experiments were approved by the local ethical committee, and the student participants signed an informed consent before taking part in the experiments.

**Apparatus**

The stimuli were presented on a Philips 201B4 CRT monitor. The monitor was driven by a RadeOn 9200 SE graphics card with an effective luminance resolution of 8 bits per channel, linearized through look-up tables. The white point of the monitor was set to D65. The screen had a spatial resolution of 1024 × 768 pixels, and it was refreshed at 100 Hz. From the viewing distance of 150 cm used in the experiments, the screen subtended 14.6 × 11.0 degrees of visual angle. The mean luminance of the screen was 45 cd/m.

**Stimuli**

All stimuli were presented as grayscale luminance modulations around the mean luminance of the screen (Figure 1). The target was a 4-cycles-per-degree (cpd) Gabor patch (with $\sigma = \lambda = 0.25$ deg) with a horizontal sinusoidal carrier. The mask carrier was a 4-cpd grating of 40% Michelson contrast, oriented either horizontally or vertically. The mask was either a circular disk or an annulus. In the experiments on the time course of masking, the outer diameter of the circular mask was either 1 degree of visual angle (center mask, 4 cycles of the sinusoidal carrier) or 8 degrees (combined mask, 32 cycles). The outer diameter of the annular (surround) mask was 8 degrees (32 cycles) and inner diameter 1 degree (4 cycles). In the experiment on the spatial properties of masking, the outer diameter of the center mask was varied from 0.25 degrees (1 cycle) to 8 degrees (32 cycles). The outer diameter of the surround mask was fixed at 8 degrees (32 cycles), and

![Figure 1](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932853/)
the inner diameter was varied from 0.25 to 4 degrees (1 to 16 cycles). The target and mask gratings had the same spatial phase. The phase was randomized across trials to prevent adaptation effects.

**Procedure**

The observer viewed the stimuli binocularly from a distance of 150 cm in a dimly lit room. A 2IFC procedure was used to determine the contrast thresholds. A trial consisted of a 250-ms period of a blank gray screen, the first stimulus interval, a 750-ms blank interval, and the second stimulus interval. The mask was presented in both stimulus intervals, the target in only one. Stimulus intervals were marked by auditory tones presented at times corresponding to the target onset. The target interval was chosen randomly for each trial, with equal probabilities for the two intervals.

The observer indicated by pressing one of two response buttons which interval had contained the target. Auditory feedback was given after an incorrect response. After the observer’s response, a fixation dot was presented for 500 ms, followed by the next trial. One block consisted of 80 trials. The target contrast for each trial was determined by an adaptive procedure (Taylor & Creelman, 1967). In each block of trials, one mask type and one stimulus onset asynchrony (SOA) were tested. The order of blocks was randomized for each participant. Each condition was repeated 2–4 times to achieve a good fit of the psychometric function to the data (see Data analysis section).

In the experiments on the time course of masking, target and mask durations were 40 and 100 ms for observers TS and LG and 30 and 150 ms for observer MR, respectively. The SOA between the target and the mask was varied.

In the experiment on the spatial properties of masking, target and mask durations were 40 and 100 ms, and the SOA was −40 ms for two observers (TS and IS). For one observer (MR), target and mask durations were 30 and 150 ms, and the SOA was −30 ms. Negative SOA values indicate that the target onset preceded the mask onset. As a control, detection thresholds were also determined for the target Gabor alone.

**Data analysis**

The data from repetitions of the same condition were pooled. A psychometric function (a cumulative normal) was fitted to the proportion correct-data using a maximum likelihood method. The contrast at which performance reached 75% correct was taken as the empirical threshold. The standard errors (68% confidence intervals) for the thresholds were estimated with a bootstrap procedure. All data analyses were done using the psignifit toolbox version 2.5.6 for Matlab (http://bootstrap-software.org/psignifit; Wichmann & Hill, 2001a, 2001b).

**Results**

**Cross-orientation masking**

Figure 2 shows the contrast detection thresholds as a function of stimulus onset asynchrony (SOA) between the...
target and the cross-oriented (vertical) mask. With the center mask, target detection thresholds were highest when the target onset coincided with either the mask onset (SOA 0 ms) or with the mask offset (SOA 100 ms for observers TS and LG; Figure 2, solid lines). Detection thresholds were at most roughly two times higher than the control (unmasked) level and quickly dropped close to the control level as the SOA was changed. To test whether the first and the second masking “peak” were time-locked to mask onset and offset, respectively, stimulus duration was changed for the third observer (MR). Mask duration was increased from 100 to 150 ms, and the second peak of threshold elevation shifted accordingly to an SOA of 150 ms (Figure 2, bottom panel). Note that this second peak is due to forward masking, because at this SOA the mask immediately preceded the target.

The surround mask did not produce any substantial masking (Figure 2, dashed lines, open circles). In fact, the thresholds were sometimes even lowered by the surround mask (the thresholds are below the horizontal control line). A similar effect was reported by Yu, Klein, and Levi (2002) in simultaneous contrast masking. On the other hand, adding the surround to the center mask did not lower the thresholds relative to the center-mask-only condition (Figure 2, dashed lines, filled circles). That is, the combination of the center and the surround masks yielded thresholds roughly comparable to the center-mask-only condition. With one observer (TS), masking was consistently even slightly stronger in this case.

### Iso-orientation masking

With the iso-oriented (horizontal) center mask, two distinct peaks in masking were again observed (Figure 3, solid lines). Surprisingly, compared to cross-orientation masking, the first masking peak did not occur when the target onset coincided with the mask onset. Instead, masking was now strongest when the target offset temporally coincided with the mask onset (SOA −40 ms for TS and LG, SOA −30 ms for MR). The second peak, on the other hand, again occurred when the target onset coincided with the mask offset, as in cross-orientation masking (SOA 100 ms for TS and LG, SOA 150 ms for MR). Hence, the strongest masking occurred when the target immediately preceded or immediately followed the mask. Masking at the peaks was unexpectedly strong, raising the thresholds from a 3% to 4% control level to near 50% contrast. In comparison, in cross-orientation masking the thresholds were always less than 10% contrast.

To confirm that the first peak occurred precisely when the target offset coincided with the mask onset, two observers participated in an additional experiment where the target duration was varied. The results are shown in Figure 4. For observer TS, target duration was either 20 or 40 ms, and the peaks in masking occurred at SOAs of −20 and −40 ms, respectively. For observer MR, target duration was either 20 or 30 ms, and the peaks in masking occurred at SOAs of −20 and −30 ms, respectively. Thus, the maximum backward masking effect occurred when the mask immediately followed the target, with no inter-stimulus interval in between.

The surround mask again produced very little or no masking (Figure 3, dashed lines, open circles). At most SOAs, masking with the large, combined mask was roughly similar to masking with the center mask only.

![Figure 3](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932853/)

Figure 3. Time course of iso-orientation contrast masking. Conventions as in Figure 2. Negative SOAs again indicate that the target onset preceded the mask onset. Error bars show the standard errors. Mask contrast was 40%. Target contrast was 40%. Target and mask durations were 40 ms and 100 ms for observers TS and LG and 30 ms and 150 ms for observer MR. The dashed horizontal lines show the control detection thresholds, measured with no mask. The standard errors for the control condition are again shown with dotted lines.
However, there was one notable difference: In the backward masking conditions, where the target was presented before the mask, masking was greatly reduced (Figure 3, dashed lines, filled circles, see the negative SOAs). Thus, a surround, which by itself had only a small effect on the thresholds, greatly reduced the effect of the strong center mask. The reduction was large, lowering the thresholds from about 50% contrast to a range between 10% and 20% contrast on average.

Spatial characteristics

Above, strong surround modulation (or dependency on mask size) was only found in one condition: backward masking with an iso-oriented mask. However, only two overlapping mask sizes (the center and the combined mask) and one surround mask size were used. The purpose of this experiment was to provide a more detailed description of the spatial properties of this surround modulation. We used the SOA at which the target immediately preceded the mask and measured masking as a function of size of the center mask and as a function of distance of the surround mask. To measure masking as a function of mask size, the outer diameter of the center mask was varied from 1 to 32 cycles. To measure masking as a function of mask distance, the inner diameter of the surround mask was varied from 1 to 16 cycles while keeping the outer diameter constant at 32 cycles.

When the size of the center mask was increased, detection thresholds first sharply increased, peaking at a mask width of 4 cycles (Figure 5, solid lines). When the mask size was further increased, the thresholds decreased. This decrease seemed to approach an asymptotic level that was still clearly above the control level. With the surround mask and as a function of distance of the surround mask. To measure masking as a function of mask size, the outer diameter of the center mask was varied from 1 to 32 cycles. To measure masking as a function of mask distance, the inner diameter of the surround mask was varied from 1 to 16 cycles while keeping the outer diameter constant at 32 cycles.

When the size of the center mask was increased, detection thresholds first sharply increased, peaking at a mask width of 4 cycles (Figure 5, solid lines). When the mask size was further increased, the thresholds decreased. This decrease seemed to approach an asymptotic level that was still clearly above the control level. With the surround mask and as a function of distance of the surround mask. To measure masking as a function of mask size, the outer diameter of the center mask was varied from 1 to 32 cycles. To measure masking as a function of mask distance, the inner diameter of the surround mask was varied from 1 to 16 cycles while keeping the outer diameter constant at 32 cycles.

When the size of the center mask was increased, detection thresholds first sharply increased, peaking at a mask width of 4 cycles (Figure 5, solid lines). When the mask size was further increased, the thresholds decreased. This decrease seemed to approach an asymptotic level that was still clearly above the control level. With the surround mask and as a function of distance of the surround mask. To measure masking as a function of mask size, the outer diameter of the center mask was varied from 1 to 32 cycles. To measure masking as a function of mask distance, the inner diameter of the surround mask was varied from 1 to 16 cycles while keeping the outer diameter constant at 32 cycles.

When the size of the center mask was increased, detection thresholds first sharply increased, peaking at a mask width of 4 cycles (Figure 5, solid lines). When the mask size was further increased, the thresholds decreased. This decrease seemed to approach an asymptotic level that was still clearly above the control level. With the surround mask and as a function of distance of the surround mask. To measure masking as a function of mask size, the outer diameter of the center mask was varied from 1 to 32 cycles. To measure masking as a function of mask distance, the inner diameter of the surround mask was varied from 1 to 16 cycles while keeping the outer diameter constant at 32 cycles.
mask, the thresholds were elevated when the inner diameter of the mask annulus was very small. This effect quickly disappeared as the target-mask distance was increased (Figure 5, dashed lines, open circles). All three observers who participated in the experiment showed an almost identical pattern of results.

**Discussion**

With a cross-oriented mask, greatest masking of contrast detection occurs when the target onset coincides with the mask onset or offset (Figure 2). With an iso-oriented mask, one of the two peaks in masking occurs when the target onset coincides with the mask offset (Figure 3, the rightmost peaks). These peak masking effects shift with changing mask duration (Figures 2 and 3). The present results are in good agreement with previous studies that have used both gratings (Mitov et al., 1981; Snowden, 2001) and luminance disks and bars (Crawford, 1947; Macknik & Livingstone, 1998; Macknik, Martinez-Conde, & Haglund, 2000) as stimuli. These studies have suggested that masking occurs when the transient responses to the target are suppressed by the transient onset or offset responses to the mask. Our findings are also consistent with those of Meese and Holmes (2007), who reported that cross-orientation masking is strongest at high temporal frequencies, relating it to the transiently responding magnocellular system. Further, we observed practically no surround modulation of overlay masking with cross-oriented masks: An added surround does not change the effect of an overlaid cross-orientation mask (Figure 2). This is in line with earlier psychophysical results in simultaneous masking, showing that large cross-orientation masks produce similar (Meese, 2004; Meese, Summers, Holmes, & Wallis, 2007) or even slightly greater (Petrov et al., 2005) masking as compared to small masks. This result is also consistent with recent neurophysiological data showing that cross-orientation inhibition is faster than surround inhibition (Smith et al., 2006).

However, there is one notable difference between the time courses of cross-orientation and iso-orientation masking. With an iso-oriented overlay mask, strongest masking is not observed when the target onset coincides with the mask onset, but when the target offset coincides with the mask onset (Figure 3, the leftmost peaks). In other words, masking is strongest when the target is presented before the mask. It is surprising that when the target onset coincides with the mask onset (SOA 0 ms), and the two are thus temporally overlapping, detection performance is much better than when the target is presented before the mask. As seen in Figure 4, when the target duration is varied, the strongest masking always occurs precisely when the target offset coincides with mask onset. Therefore, it is the immediate temporal succession that leads to the strongest deterioration of detection in iso-orientation masking.

Iso-orientation masking is also much stronger than cross-orientation masking. This suggests that the sharp edge of the center mask is not a major source of masking because the iso- and cross-oriented masks have the edge at the same spatial location. This argument is also supported by the observation that the surround mask alone does not produce masking with either orientation (Figures 2 and 3, dashed lines, open circles) even though it also has its (inner) edge at the same spatial location as the center masks. We conclude that the masking we observed is produced by interactions between mechanisms responding to the target and mask patterns rather than only the edges.

Some previous contrast detection studies also reported strong masking near mask onset (Wilson & Kim, 1998; Yu & Levi, 1999) or near mask onset and offset (Snowden, 2001) at low spatial frequencies. In these studies, a brief target and a longer mask were always temporally overlapping, so the strong backward and forward masking effects reported here could not be observed. Mitov et al. (1981) used an extended range of SOAs as we did but did not observe such dramatic backward and forward masking, presumably because their target and mask patterns differed from each other in spatial frequency. backward and forward masking of contrast detection has been reported (Georgeson & Georgeson, 1987), but because of very short presentation times used for both the target and the mask, the separate effects of onset and offset transients are difficult to estimate from those data.

Although masking is strongest at the mask onset and offset at low spatial frequencies, some masking also occurs when a brief target is flashed during the mask presentation (Mitov et al., 1981; Wilson & Kim, 1998). The same is true in our results: there is some masking evident at intermediate SOAs, although with one observer (TS, Figure 3), the thresholds are very close to the control level. We expect that with a lower spatial frequency or with a longer mask duration the thresholds at intermediate SOAs could be even lower. This is because the sustained responses should become weaker the lower the spatial frequency and because a longer duration would prevent the effects of onset and offset transients from merging together.

In our experiments, the target and the mask had the same spatial frequency and the same spatial phase. In this case, when the target contrast is low and the mask comes on immediately after the target, there is no physical target offset because the local contrast goes directly up to 40% (mask contrast) without going to 0% first. Therefore, there can be no response to the target offset. The same is true for the target onset when the mask immediately precedes the target. It could be argued that in these cases, when the target can elicit no onset (or offset) response, the target is not detected as a separate temporal event but is integrated with the mask. Consequently, target contrast would have to be increased to a very high level for the detection to become
possible. In fact, it is evident from the data with the iso-oriented center mask (Figure 3, solid lines) that in five out of the six "peaks" (both a backward and a forward masking peak for each of the three observers), the contrast detection threshold for the target is higher than the mask contrast (40%). This implies that the target becomes visible when its onset (or offset) is restored. However, this straightforward explanation based on the absence of transients, although possibly partly true, is insufficient if we consider the data as a whole. Most importantly, backward masking is greatly reduced, and target detection becomes possible at much lower contrasts, when a surround is added to the center mask.

Surround modulation only affects iso-orientation masking (Figure 3), where adding a surround pattern greatly reduces the effect of the center mask, as reported earlier in simultaneous contrast masking (Yu & Levi, 2000). Interestingly, in the present data, surround modulation only occurs when the target temporally precedes the mask (Figure 3, compare the solid lines with the dashed lines, filled circles at negative SOAs). The large reduction in masking can be explained by surround suppression: The center mask alone induces strong masking, but the added surround suppresses the neural responses to the center, causing a reduction in the masking induced by the center mask. Further, because surround modulation is restricted to backward masking near the onset of the mask (Figure 3), it seems that the transient onset responses are susceptible to strong surround modulation. The strong masking near the offset of the mask, on the other hand, is unaffected by the surround, suggesting no surround modulation of the transient offset response. The idea of very fast surround modulation is in line with results on the speed of surround modulation of perceived contrast (Kilpeläinen, Donner, & Laurinen, 2007) and is also supported by neurophysiological observations of fast surround influences mediated by feedback connections to the primary visual cortex (Angelucci & Bullier, 2003; Bair et al., 2003; Hupé et al., 2001). A similar psychophysical effect, where a bigger mask leads to weaker masking, has been observed with very different stimuli and tasks such as orientation (Wehrhahn, Li, & Westheimer, 1996) and vernier (Herzog & Koch, 2001) discrimination and modeled as fast neural inhibition (Herzog, Ernst, Etzold, & Eurič, 2003).

It should be noted, however, that in simultaneous contrast masking the results on increasing mask size are mixed. Yu and Levi (2000), for example, found that adding a surround to an iso-oriented mask reduced masking. Others have found that increasing mask size (Bonneh & Sagi, 1999; Foley, 1994; Foley & Chen, 1999; Meese, 2004), addition of flankers (Chen & Tyler, 2001), or addition of an in-phase, equal-contrast surround (Olizak & Laurinen, 2005) can actually increase masking. Yu, Klein, and Levi (2003) suggested that the effect of a surround depends on the contrast ratio between the pedestal and the surround, higher surround contrasts leading to higher thresholds. Meese, Hess, and Williams (2005) found individual differences in the effect of the surround and showed with a gain control model that the deteriorating effect of the surround may result not only from suppression but also, counter-intuitively, from contrast enhancement by the surround. In our results on temporal contrast masking, only a strong reduction in masking was evident in the backward masking conditions. However, in light of the abovementioned results, it cannot be ruled out that a higher surround contrast could eventually turn this enhancement into a further increase in masking (but note that a change in contrast also affects the timing of surround modulation; see Kilpeläinen et al., 2007).

Yu and Levi (1999) reported that when a target and a mask were presented simultaneously for 150 ms, flanking stimuli only modulated contrast masking when the flanks were also presented simultaneously with the target. However, we consider the data as a whole. Most importantly, backward masking is greatly reduced, and target detection becomes possible at much lower contrasts, when a surround is added to the center mask.

Yu and Levi (1999) reported that when a target and a mask were presented simultaneously for 150 ms, flanking stimuli only modulated contrast masking when the flanks were also presented simultaneously with the target. However, we consider the data as a whole. Most importantly, backward masking is greatly reduced, and target detection becomes possible at much lower contrasts, when a surround is added to the center mask.

Yu and Levi (1999) reported that when a target and a mask were presented simultaneously for 150 ms, flanking stimuli only modulated contrast masking when the flanks were also presented simultaneously with the target. However, we consider the data as a whole. Most importantly, backward masking is greatly reduced, and target detection becomes possible at much lower contrasts, when a surround is added to the center mask.
ing target-surround distance. In our results, the extent of the direct surround influence is so small that it is probably not due to lateral interactions but is more likely due to overlay masking from the near surround. This is in line with the suggestion that surround suppression of detection is weak or absent in the fovea (see above). The size- and distance-dependency functions we measured are strikingly similar in shape to response functions measured from neurons in the primary visual cortex of both cat and monkey (Cavanaugh et al., 2002; Jones et al., 2001; Sengpiel et al., 1997). The shapes of these functions have been explained with a model having antagonistic excitatory and inhibitory fields of different sizes (Cavanaugh et al., 2002). The similarity to our psychophysical results suggests that the observed changes in masking strength could also be explained by a similar mechanism. When the mask is very small, an increase in mask size leads to a large change in the response of the (small) excitatory center mechanism, resulting in a stronger masking effect. As the mask size is further increased, however, the larger inhibitory surround mechanism is more strongly activated, leading to a reduced response and thus to a reduction in masking.

Conclusions

The strength of suppressive interactions in human spatial vision depends jointly on stimulus timing and relative orientation between the target and the mask. The transient iso-orientation masking that is characteristic to low spatial frequency stimuli is strongest when the target immediately precedes (backward masking) or follows (forward masking) the mask. This masking is reduced in the presence of a surround, but only in the backward masking condition. This suggests very fast surround modulation that suppresses the transient onset, but not the transient offset, response to the mask. The spatial characteristics of this modulation suggest it has its neural basis in early cortical surround suppression mechanisms.

Acknowledgments

We would like to thank Marc Repnow for technical support and Bilge Sayim and Maria Olkkonen for useful comments on the manuscript. TS was supported by the Swiss National Science Foundation (SNF).

Commercial relationships: none.
Corresponding author: Toni P. Saarela.
Email: toni.saarela@epfl.ch.
Address: EPFL SV BMI LPSY, Station 15, CH-1015, Lausanne, Switzerland.

References


