Vision Research 105 (2014) 21–28

Contents lists available at ScienceDirect

Vision Research

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

Dissociation between spatial and temporal integration mechanisms in Vernier fusion

Jan Drewes [a,*], Weina Zhu [a,b,c,d], David Melcher [a]

[a] Center for Mind/Brain Sciences (CIMeC), Trento University, Rovereto, Italy
[b] School of Information Science, Yunnan University, Kunming, China
[c] Department of Psychology, Giessen University, Giessen, Germany
[d] Kunming Institute of Zoology, Chinese Academy of Science, Kunming, China

ARTICLE INFO

Article history:
Received 30 April 2014
Received in revised form 29 July 2014
Available online 16 September 2014

Keywords:
Vernier
Spatial integration
Temporal integration

ABSTRACT

The visual system constructs a percept of the world across multiple spatial and temporal scales. This raises the questions of whether different scales involve separate integration mechanisms and whether spatial and temporal factors are linked via spatio-temporal reference frames. We investigated this using Vernier fusion, a phenomenon in which the features of two Vernier stimuli presented in close spatio-temporal proximity are fused into a single percept. With increasing spatial offset, perception changes dramatically from a single percept into apparent motion and later, at larger offsets, into two separately perceived stimuli. We tested the link between spatial and temporal integration by presenting two successive Vernier stimuli presented at varying spatial and temporal offsets. The second Vernier either had the same or the opposite offset as the first. We found that the type of percept depended not only on spatial offset, as reported previously, but interacted with the temporal parameter as well. At temporal separations around 30–40 ms the majority of trials were perceived as motion, while above 70 ms predominantly two separate stimuli were reported. The dominance of the second Vernier varied systematically with temporal offset, peaking around 40 ms ISI. Same-offset conditions showed increasing amounts of perceived separation at large ISIs, but little dependence on spatial offset. As subjects did not always completely fuse stimuli, we separated trials by reported percept (single/fusion, motion, double/segregation). We found systematic indications of spatial fusion even on trials in which subjects perceived temporal segregation. These findings imply that spatial integration/fusion may occur even when the stimuli are perceived as temporally separate entities, suggesting that the mechanisms responsible for temporal segregation and spatial integration may not be mutually exclusive.

1. Introduction

One of the basic goals of visual perception is extract features and bind them spatially and temporally into coherent objects. Feature fusion is a phenomenon in which the binding of features to their parent objects fails, such that the features are either confused, mis-attributed to the wrong object, or collapsed into a single perceptual object. For example, a green disc followed immediately by a red disc may lead to a yellow percept (Efron, 1967). Feature fusion may occur with geometric features, as has been repeatedly demonstrated with Vernier stimuli. A Vernier stimulus is a parallel pair of vertical lines, with a horizontal offset. Human vision is remarkable sensitive to the spatial offset between the two lines in a simple Vernier stimulus, even reaching hyper-acuity – a kind of spatial resolution that cannot be explained with the retinal sensor density alone (Beard, Levi, & Klein, 1997; Fahle, 1991; Westheimer & McKee, 1977; Wülfing, 1893). When two Vernier stimuli with opposite spatial offsets are shown in close spatial and temporal vicinity, their critical feature (the spatial offset) can be fused, resulting in the perception of not two, but one Vernier stimulus with an apparent spatial offset that is compressed between the two actual spatial offsets displayed to the subject (Herzog & Koch, 2001; Herzog, Scharnowski, & Hermens, 2007; Scharnowski, Hermens, & Herzog, 2007a; Scharnowski et al., 2007b).

Several different aspects of Vernier fusion and related phenomena have been studied in detail. Varying the spatial offset, for instance, affects the resulting percept: with increasing offset magnitude, subjects reportedly perceive the two stimuli either as one single Vernier stimulus, as motion from the offset of the first Vernier in the direction of the second, or as two separate, superimposed stimuli (Scharnowski et al., 2007b). The addition of a grating of

---

* Corresponding author.
E-mail address: mail@jandrewes.de (J. Drewes).

http://dx.doi.org/10.1016/j.visres.2014.08.017
0042-6989/© 2014 Elsevier Ltd. All rights reserved.

Vernier-like lines may achieve an effect called “shine-through” (Herzog & Koch, 2001), and multiple Vernier presentations may be fused instead of just two (Herzog et al., 2003).

This type of feature fusion is particularly interesting to study because it allows precise control over the spatial and temporal aspects of the two stimuli. Given the importance of temporal and spatial integration in the creation of coherent objects and in the parsing of the continuous influx of visual information into events over different temporal windows (Wutz & Melcher, 2013; Wutz et al., 2014), Vernier fusion provides a useful methodology to examine the fine grain evolution of spatiotemporal integration.

One factor that all previous studies of Vernier fusion have in common is the methodology of performance measurement. Subjects are always asked for the direction of the perceived spatial offset, realized as a binary choice, usually by means of two buttons (Brand, Kopmann, & Herzog, 2004; Brand et al., 2005; Herzog & Koch, 2001; Herzog, Lesemann, & Eurich, 2006; Herzog, Scharnowski, & Hermens, 2007; Herzog et al., 2003; Rüter, Kammer, & Herzog, 2010; Rüter et al., 2013; Scharnowski, Hermens, & Herzog, 2007a; Scharnowski et al., 2007b). All previous studies therefore measured only the *sign* of the perceived feature; none of the studies known to us attempted to measure the actual *magnitude* of the perceived spatial offset. With different display timings may come different percepts of the offset magnitude, which would remain hidden when only the sign of the percept is recorded.

Although previous studies have demonstrated many of the key properties of Vernier fusion, none have systematically investigated the type of percept on a trial-by-trial basis. Generally, it is assumed that subjects always fuse the two presented Vernier stimuli into a single percept (Herzog, Scharnowski, & Hermens, 2007; Scharnowski et al., 2007b). The only study (briefly) investigating blank ISIs between Vernier presentations reported that subjects see either motion or two consecutively flashed stimuli, depending on the ISI (Rüter, Kammer, & Herzog, 2010); however, no systematic account of the relation between ISI and percept was given – and the possible case of perceiving the two Verniers as just one stimulus was not mentioned, although this type of percept should be possible at least at very short ISIs. Instead, only one fixed, non-symmetric pairing of Vernier offsets was tested.

Summarizing, in the context of Vernier fusion, very similar stimulus presentations may result in very different percepts (single stimulus, motion, or two separate stimuli), depending on both spatial and temporal offset. None of the previous studies described above systematically investigated the type of percept on a trial-by-trial basis. It is therefore unclear whether the magnitude of the perceived spatial offset depends on the type of percept, which in turn may be influenced by the timing between the two Vernier stimuli.

The goal of the present study was to study both the perceived spatial offset and the perceived type of presentation (double, motion or single) on a trial-by-trial basis. Seen from the perspective of temporal integration windows, if the second Vernier follows the first Vernier in very rapid succession, they may fall in the same integration period, or “Perceptual Moment” (Harter, 1967; Purves, Paydarfar, & Andrews, 1996; Stroud, 1967; VanRullen & Koch, 2003; Varela et al., 1981), and thus their features may be confused. The longer the temporal gap between the two stimulus presentations, the more likely it may be for the visual system to separate the stimuli into individual “moments”, or objects, and thus their features would be more likely to be segregated. Based on the previous literature, we hypothesized that there may also be intermediate forms, e. g. the above mentioned percept of motion may represent a case where the two stimuli were neither fused nor segregated fully. We tested whether, for a given pair of rapidly presented Vernier stimuli, the magnitude of the perceived Vernier offset would be dependent on whether a particular stimulus presentation was perceived as a single stimulation, a moving stimulus, or two sequential stimuli. In consequence, the range of timings at which each of the three expected percepts happens most frequently promised to provide insight into the nature of the involved integration windows. Of particular interest was the question of whether spatial and temporal integration are based on independent or linked mechanisms.

2. Methods

2.1. Experimental setup

According to previous reports, Vernier fusion is only found in a narrow range of spatial offsets, commonly 2′ or less (Scharnowski et al., 2007b). To satisfy these requirements, a 20″ CRT was set to run at a resolution of 1600 × 1200 px, resulting in 0.5′ spatial resolution at a viewing distance of 165 cm. Together with a refresh rate of 100 Hz this allowed for a detailed mapping of the characteristics of Vernier fusion in both time and space. Subjects were seated orthogonally in front of the screen, with a combined chin/forehead support to maintain a fixed head position. A standard 2-button computer mouse was used for response input. All relevant stimuli were drawn as white lines on a black background. The experimental chamber was dimly lit, and while not calibrated to any particular level, care was taken to ensure equal lighting across all subjects and sessions.

2.2. Paradigm

A graphical illustration of the paradigm can be seen in Fig. 1. Each trial began with the display of a red fixation dot in the center of the screen. Subjects then signaled that they were ready to commence the trial by pushing a button on the mouse. The red dot then changed to green to announce the beginning of the sequence, and remained on screen for a randomized interval between 1 and 1.5 s. The fixation dot was then followed by a 200 ms gap, after which the first Vernier stimulus was presented for a fixed duration of 30 ms. Following a variable inter-stimulus interval (ISI), a second Vernier stimulus was shown, with an identical fixed presentation time of 30 ms. The second Vernier could either have the same spatial offset, or the opposite spatial offset (sometimes called “Anti-Vernier”) as the first Vernier stimulus. Each of the two vertical lines comprising a Vernier stimulus was 0.5′ in width and 15′ long, with a vertical separation of 1.5′, making for a total stimulus height of 31.5′. This is within the range where best Vernier acuity is expected (Watt, 1984; Westheimer & McKee, 1977). Horizontal spatial offsets (0.5′, 1.0′, 2.0′ or 3.0′) and temporal ISIs (0, 10, 20, 30, 40, 50, 70, 100 or 200 ms) were selected on a trial-by-trial basis in a controlled randomized fashion, resulting in a total of 18 conditions per spatial offset (72 conditions for all 4 spatial offsets). Each condition was displayed once per block.

After each trial, subjects were required to give 2 responses. The first response was realized as a visual scale of Vernier offsets, on which subjects were asked to use the mouse to indicate the offset of the last Vernier stimulus they saw (see Fig. 1, lower left). The ticks of the scale covered an offset range from −4′ to +4′ and were positioned in discreet steps of 0.5′ with a spatial distance of 40′; subjects were encouraged to click anywhere between the ticks to allow for a pseudo-continuous measurement. This mode of response avoids a limitation found with the binary measure of offset direction. In cases when the (spatial) Vernier offset is large enough, subjects may be able to reliably judge the Vernier offset direction, leading to saturation of the direction measurement. Any dynamics of the perceived *magnitude* of the Vernier offset would thus be lost; however, the magnitude-based response based on

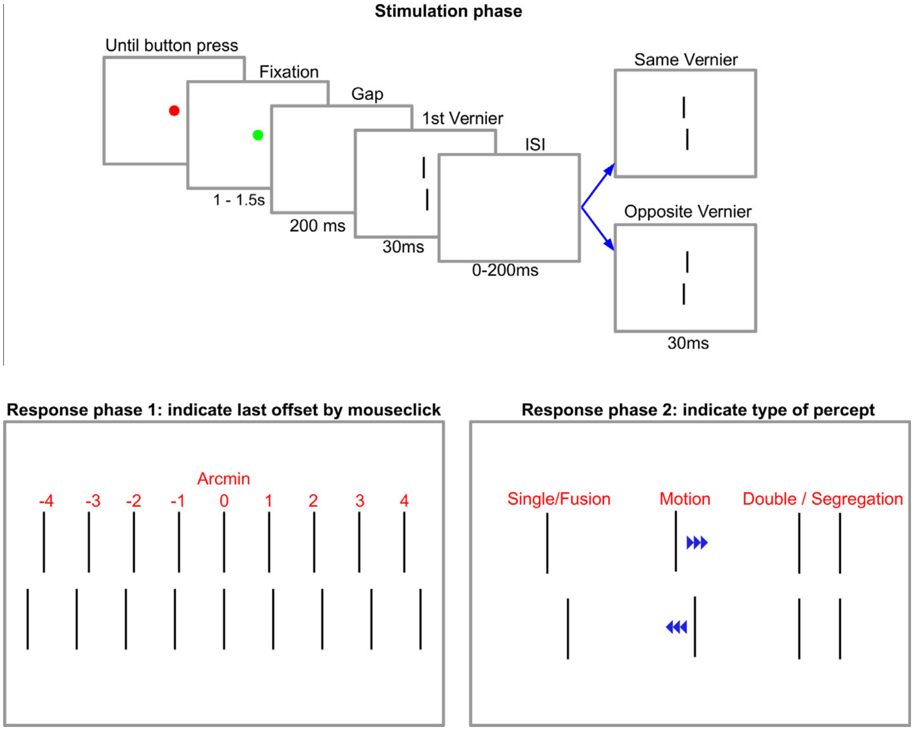

Fig. 1. Experimental paradigm. Top row: stimulation phase. Bottom row: response phases. Subjects used the mouse to direct a cross hair onto the desired screen position. Red text is for illustration only and was not shown during the experiment. The number of tick marks on the scale was reduced for better viewing (original tick marks were spaced at 0.5 arcmin). Contrast has been inverted, and depicted stimulus sizes have been magnified for better viewing. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the continuous scale does not saturate and is therefore well suited to measure the dynamics of the perceived Vernier offset magnitude.

The second response asked subjects to indicate the perceived stimulus presentation type from one of three possibilities: a single Vernier, corresponding to complete fusion (“Single”); motion from the first to the second Vernier, corresponding to partial fusion (“Motion”); and two independent Vernier stimuli, corresponding to segregation/failure of stimulus fusion (“Double”). In both response phases, subjects were also allowed to indicate if they did not see the stimulus clearly (e. g. due to blinking at the wrong moment or lapses of attention) by clicking the right mouse button, discarding the trial.

3. Analysis and results

3.1. Analysis

In total, 26 subjects completed the experiment; all subjects reported normal or corrected-to-normal vision. All of the subjects provided informed written consent, and all experiments were performed according to the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the guidelines of the University of Trento ethics committee. To verify that subjects were indeed performing the task as expected, a post hoc test was performed: in 50% of the trials, both the first and the second Vernier featured identical spatial offsets. In theory, subjects should never have a percept of motion in this condition, for the spatial locations of all stimulus components are fixed in screen coordinates. Still, most subjects reported a percept of the “motion” type for a certain percentage of these trials. To allow for a certain lapse rate, only those subjects reporting motion percepts in more than 20% of the relevant trials were excluded from the study ($N = 5$). The remaining subjects ($N = 21$, 17F 4M, age 19–34) were included in all further analysis. Of these, 8 subjects completed 3 spatial offsets ($1'$, $2'$ and $3'$), 13 subjects completed all 4 spatial offsets. 4 subjects completed 12 blocks, 3 subjects completed 8 blocks and 14 subjects completed 4 blocks of trials. The total number of trials collected per subject varied between 212 and 788. Leftward (negative) and rightward (positive) offsets were rectified and pooled. All of the included subjects used the right mouse button to discard trials less than 1% of the time, with the exception of two individuals (1.4% and 2.0%). Discarded trials were not repeated and thus excluded from the analysis.

3.2. Results

Pooled across all conditions, the different spatial offsets resulted in clearly different subject estimates of spatial offset. This difference was confirmed in a one-way ANOVA showing a main effect of spatial offset ($F(3) = 289.12$, $p < 0.01$. In all statistical analysis, empty cells due to uneven numbers of trials between subjects were treated as missing values, while retaining all cells with existing values). Interestingly, subjects systematically underestimated the spatial offsets of the Vernier stimuli for both same-offset and opposite-offset stimulus pairings (see Fig. 2, left panel). Underestimation was stronger for opposite-offset stimulus conditions, as expected, and differed significantly from the same-offset conditions (anova, $p < 0.01$, $F(1) = 29.13$), a first indication that feature fusion of the Vernier stimuli was indeed taking place. Only in the case of $0.5'$ spatial offset was there a lack of a significant difference between same-offset and opposite-offset conditions.

We proceeded by separating the results by ISI (see Fig. 2, right panel). Subject estimates in the opposite-offset condition depended strongly on ISI (anova, $F(8) = 9.32$, $p < 0.01$), while no significant dependency was found with the same-offset condition (anova, $F(8) = 0.84$, $p = 0.57$). In other words, the influence of ISI on perceived offset was limited to the opposite offset condition. In these trials, the weakest spatial offset underestimation (and

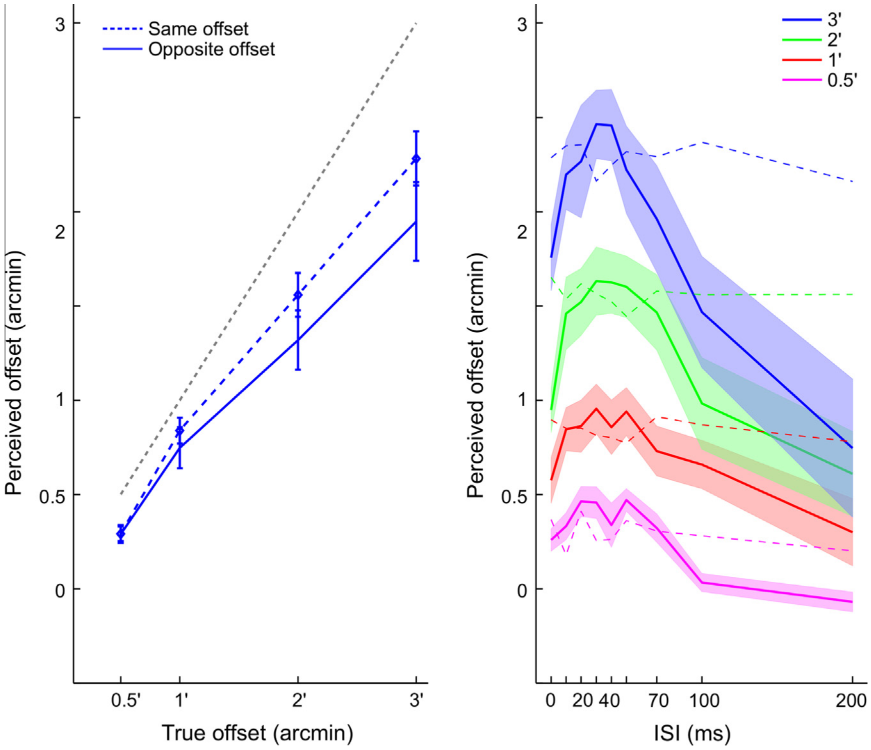

Fig. 2. Offset estimation. Left panel: true spatial offset against perceived offsets, averaged across all ISIs. Solid lines represents opposite-offset condition, dashed lines represent same-offset condition. Gray dashed line represents veridical perception. Right panel: perceived spatial offset, per-ISI. Different colors represent different true offsets (see legend). Solid lines represent opposite-offset condition, dashed lines represent same-offset condition. All data given as mean and s.e.m. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

most veridical percept) was found at medium intervals, with the least underestimation at ca. 30–40 ms ISI. The strongest underestimation occurred at the longest interval (200 ms).

Subsequently, the data was separated by the reported type of event percept. While for the opposite-offset stimuli, the frequency of the 3 different percepts was again strongly dependent on both ISI (anova, Single: $F(8) = 6.16$, $p < 0.01$, Motion: $F(8) = 11.38$, $p < 0.01$, Double: $F(8) = 26.21$, $p < 0.01$) and spatial offset (anova, Single: $F(3) = 125.72$, $p < 0.01$, Motion: $F(3) = 57.25$, $p < 0.01$, Double: $F(3) = 4.89$, $p < 0.01$), the percept frequency for same-offset stimuli was dependent only on ISI (Single: $F(3) = 16.16$, $p < 0.01$, Double: $F(3) = 13.53$, $p < 0.01$) and not on spatial offset (Single: $F(3) = 0.22$, $p = 0.89$, Double: $F(3) = 0.16$, $p = 0.92$); see Figs. 3 and 4.

Across spatial offsets, shorter ISIs resulted in a higher frequency of the “Single” percept, reaching almost 90% with the smallest spatial offset and 0 ms ISI. As expected, the prevalence of the “Single” percept decreased for longer ISIs, and the frequency of the “Motion” percept increased, until the “Motion” percept reached a peak at approximately 40 ms. After this, with increasing ISIs, both “Single” and “Motion” percepts were further reduced in frequency with the “Double” percept becoming dominant (see Fig. 3).

Across ISIs, smaller spatial offsets resulted in increased frequency of the “Single” percept, which was reported more than 65% of the trials at an offset of 0.5′, but less than 14% of the trials at an offset of 3′ (see Fig. 4). The “Motion” percept followed a mirrored trend, with a reporting frequency of approximately 16% at the smallest offset and more than 55% at the largest offset. The frequency of the “Double” percept increased only slightly with the spatial offset, starting at just less than 20% at the smallest offset and reaching approximately 31% at the largest offset.

To analyze the extent to which the perceived presentation (event) type and the estimated spatial offset were correlated, we continued data analysis after separating trials by reported percept. For the same-offset conditions, no significant difference was found between “Single” and “Double” (anova, $F(1) = 1.744$, $p = 0.187$); the “Motion” trials were ignored, as they were deemed lapses (see above). Data for the same-offset condition was hence pooled across percepts for visual comparison in Fig. 5. For the opposite-offset trials, both a general anova ($F(2) = 29.038$, $p < 0.01$) and pairwise t-tests between the percepts were significant (Single vs. Motion: $p < 0.001$, Single vs. Double: $p = 0.015$; Motion vs. Double: $p < 0.001$). In those opposite-offset trials in which subjects reported percepts of the “Single” type (35% of trials), across ISIs, underestimation was stronger than when averaging across all perceptual types (compare Figs. 5 and 2).

Underestimation was reduced to near-identity with the same-offset trials in those cases where subjects reported to see “Motion” (38% of trials). In the particular case of 0.5′ spatial offset, the result did not differ significantly from veridical. In the cases where subjects reported to see two separate stimulus presentations (“Double”, 27% of trials), the underestimation was again stronger than when averaging across all perceptual types, and even stronger than in the “Single” case. The trend towards stronger underestimation with larger offsets remained across all percepts.

When analyzing the same data per ISI, the timing-dependent trend in the opposite-offset data remained visible (see Fig. 5, bottom row) and significant (anovas, Single: $F(8) = 3.39$, $p < 0.01$; Motion: $F(8) = 5.41$, $p < 0.01$; Double: $F(8) = 2.75$, $p = 0.01$) across all three percepts.

3.3. Summary

Our results revealed that the perceived offset size for any Vernier/Anti-Vernier pairing depended in a non-trivial manner not only on the actual spatial offset, but also on the time that was allowed to pass between the stimulus presentations (ISI). This

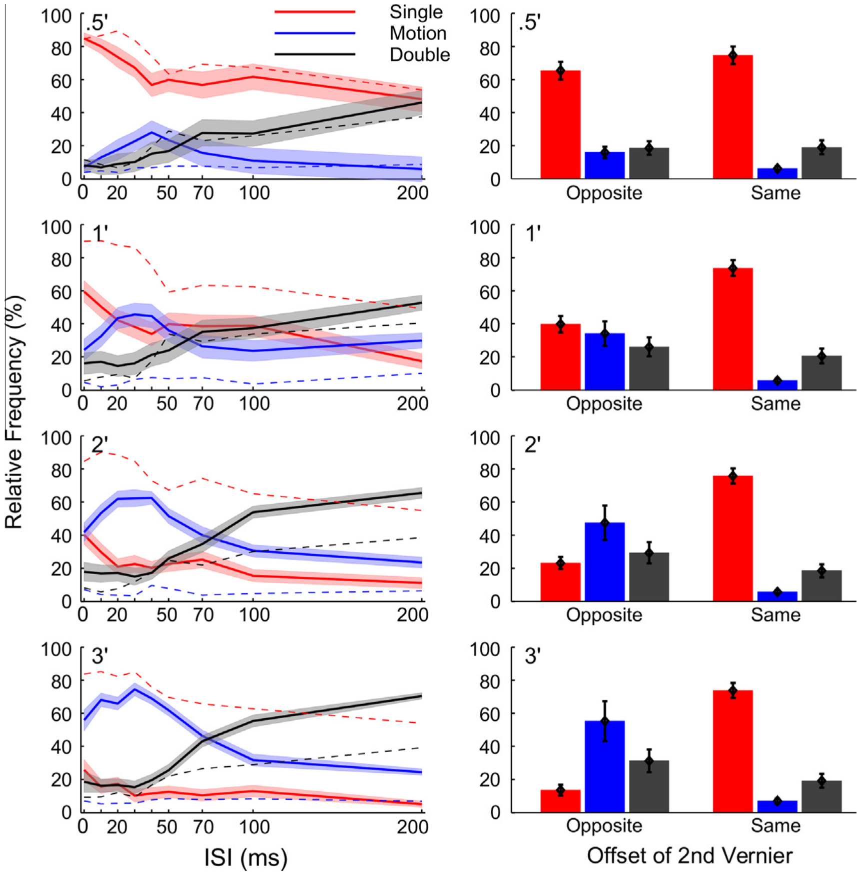

Fig. 3. Relative frequency of percept types, averaged across subjects, per spatial offset. Solid lines represent opposite-offset conditions, dashed lines represent same-offset conditions. Red represents trials perceived as "single" stimulus, blue represents "motion", black represents "double". All data given as mean and s.e.m. The number in the top-left corner of each plot represents the spatial offset. Left column: perceived presentation type per ISI, averaged across subjects. Right column: perceived presentation types, averaged across ISIs and subjects. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

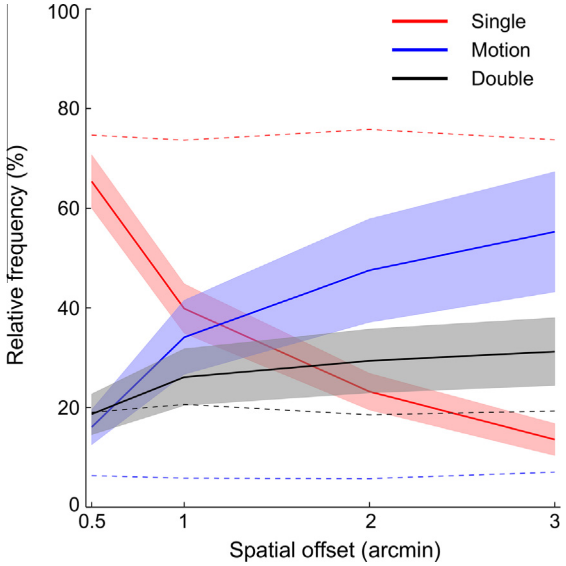

Fig. 4. Frequency of perceived presentation type vs. spatial offset, averaged across ISIs and subjects. Mean and s.e.m. for opposite-offset conditions (solid lines), mean for same-side conditions (dashed lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

timing-dependency did not exist in the same-offset condition. Unlike previous reports, under the given conditions subjects did not fuse the stimuli into a single percept in about 65% of the trials. Still, in the opposite-offset condition spatial fusion occurred, even in those cases where subjects reported temporal segregation of the stimuli – as evidenced by increased underestimation of the spatial offset relative to the same-offset condition.

The type of percept (Single, Motion, Double) also depended both on spatial offset and the time passing between stimulus presentations; here too, the dependency on the timing existed only in the opposite-offset condition.

4. Discussion

In this study, we investigated the nature of the spatial and temporal integration occurring on single trials in a Vernier fusion design. We systematically varied both spatial offsets and ISIs to probe the dynamics of Vernier feature fusion in both time and space.

4.1. Summary of findings

1. Underestimation of spatial offset occurred in both same-offset and opposite-offset conditions.
2. Underestimation was significantly stronger in the opposite-offset condition, indicating feature fusion.
3. The magnitude of the perceived (fused) spatial offset depended on the ISI and the type of percept (single, motion, double).
4. Spatial fusion still occurred even when stimuli were perceived as temporally segregate.

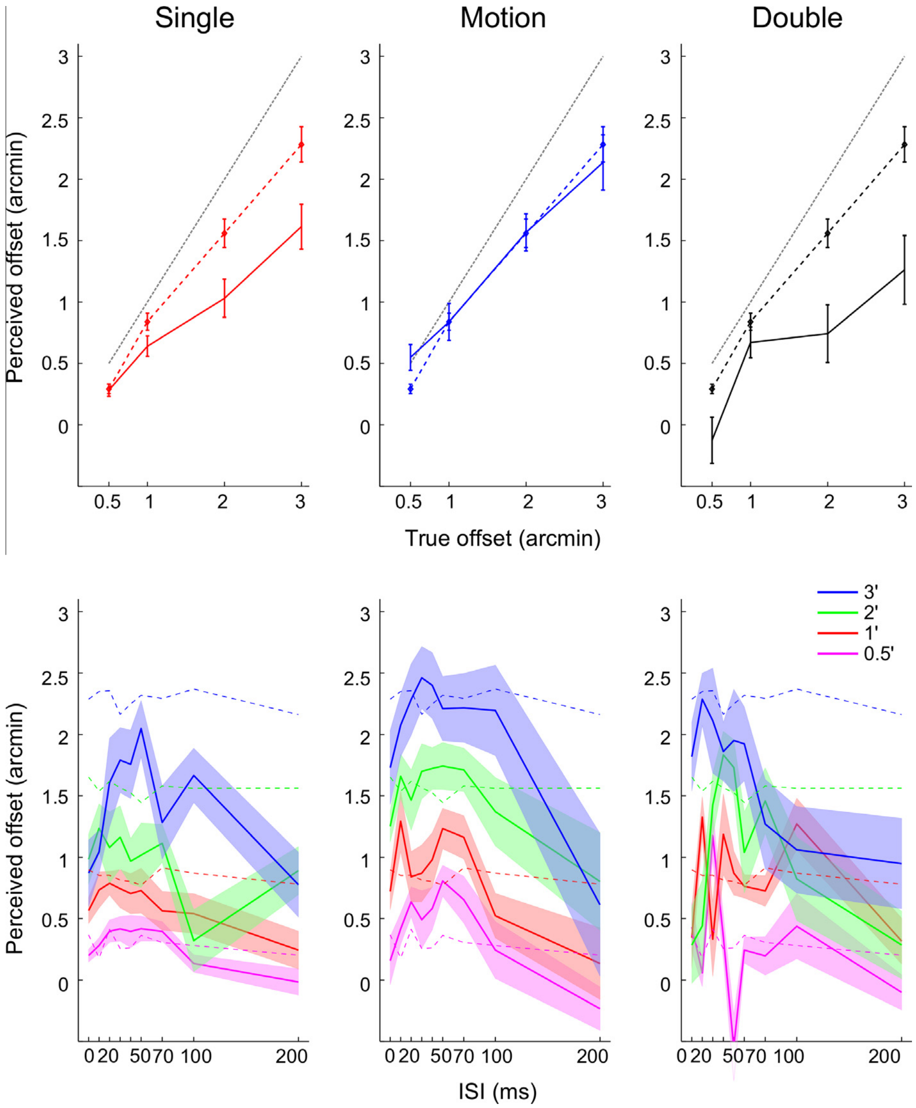

Fig. 5. Offset estimation separated by reported type of percept, averaged across subjects. Solid lines represent opposite-offset condition (mean and s.e.m.). Same-offset conditions were not significantly different between percepts (see text), and data was pooled across percepts, represented by dashed lines. Perceived presentation types are sorted in columns (Single, Motion, Double). Top row: averaged across ISI. Red represents "single" percept, blue represents "motion" and black represents "double". Gray dashed line represents veridical. Bottom row: data separated by ISI. Solid colors represent different true offsets (see legend). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In no condition did our subjects fuse more than 90% of the time; therefore, at least 10% of the time, subjects noticed at least to a certain extent that more than a single stimulus presentation had occurred. In other conditions, with bigger spatial offsets, but also with longer temporal intervals, subjects would report the "single" percept type less often, in some conditions less than 50% of the time. In the relevant previous studies, this phenomenon was mentioned only by Herzog, Scharnowski, and Hermens (2007), who stated that "for long ISIs, apparent motion or flickering could not be avoided for most observers"; however, their paradigm differed in several aspects from the present study. Rüter, Kammer, and Herzog (2010) report in very similar circumstances, that subjects perceived the stimulation either as motion or as 2 independent stimuli, yet they do not mention the occurrence of "single" type percepts.

All previous studies only requested a judgment of offset direction. While this may be sufficient to demonstrate Vernier dominance and related phenomena, the actual magnitude of the resulting percept is lost. In this way, the results of the first experiment in Rüter, Kammer, and Herzog (2010), and all previously mentioned studies, only correspond to a *sign* transformation of the results we present in our work. This sign-transformed perceived offset saturates against the ceiling when the actual perceived offset magnitude shows a more complex and non-monotonous time course. Our data might even seem to converge against zero in the time-infinite; however, this kind of long-term integration would not agree with most theories of temporal–spatial perception. It seems more likely that at even longer ISIs, subjects would perceive both Verniers as completely independent entities. With longer ISIs, offset judgments should therefore converge with those achieved with the same-offset condition.

In our experiment, very similar stimulus presentations led to very different percepts, depending systematically on both the spatial offset and the temporal distance of our Vernier stimuli. A single integration mechanism with a limited temporal and spatial range may seem to be an appropriate explanation at first glance – whenever both stimulus presentations fall into the spatio-temporal integration range, feature fusion occurs and subjects perceive only one single stimulus. However, the temporal and spatial limits of this mechanism would be expected to depend on each other. One might expect either a space–time trade-off– the wider the spatial offset, the narrower the temporal integration range and vice versa (see Fig. 3), or conversely a hierarchical system in which higher levels have receptive fields with larger spatial and temporal integration windows. Such a single mechanism would not, however, explain the cases when subjects reported temporal segregation of the stimuli ("double"), but still very significantly fused the spatial offset of the Vernier stimuli (Fig. 5). A hypothesis based on a single mechanism would demand that stimuli be either integrated both temporally and spatially, or neither. The observed dissociation between temporal and spatial integration would be better explained with at least a partial independence between the spatial and temporal integration mechanisms, such that spatial fusion and temporal segregation would not be mutually exclusive. This kind of mechanism may also explain the percept of motion occurring on those ISIs that are of intermediate length. In this range of ISIs, temporal integration may begin to fail, leading to less-than-complete fusion of the stimuli, yet still partially connecting both stimulus presentations in the perception of the subject. As far as the timing of the spatial and temporal integration is concerned, the frequencies of the different percepts may be seen as a twofold indicator. Firstly, the ratio of "Double" percepts seems to increase mostly with increased ISIs, and only to a much lesser extent with increased spatial offset. The ratio of "Single" vs. "Motion" however seems to be influenced by the timing, but also very strongly by the spatial offset. We may postulate that the transition from other percepts to the "Double" percept signifies the closing of a temporal integration window, while the transition from "Single" to "Motion" may be attributed additionally to a spatial integration process, such as motion integration. An alternative interpretation might be three at least partially independent mechanisms in competition with each other, each with a preferred range of ISIs. In either case, when averaging across spatial offsets, the point of turnover, after which the majority of trials is perceived as "Double", is located between 70 and 100 ms ISI. This may be the approximate timing where the temporal integration starts to fail, as has been reported in masking studies (Breitmeyer & Ögmen, 2006; Enns & Di Lollo, 2000).

The type of percept clearly influences the perceived offset magnitude: in those trials where subjects reported to see motion in the opposite-offset condition, offset judgments across ISIs are nearly the same as the same-side judgments; still, when separated by ISI, a clear temporal trend emerges. The peak in offset estimation is found at around 30–40 ms ISI. An interval consistent with the shortest integration windows found in studies of perceived simultaneity (Mach, 1865). Offset judgments in this interval even exceed those of the same-side condition. It might be speculated that motion extrapolation is the underlying reason for this; however, our results offer no way to reach certainty.

Classically, Korte's law would predict that "good motion" would be perceived when the time between two stimuli is reduced as the distance is increased (Korte, 1915). Alternatively, it has been postulated that to experience "good motion", temporal and spatial distances should be increased or decreased together (Gepshtein & Kubovy, 2007). Interestingly, in our data the peak in the frequency of the motion percept is at a constant temporal location of 30–40 ms ISI, independent of the spatial offset; the laws of "good motion" therefore do not seem to apply in this context. One possible confound in the judgments of our subject may stem from a phenomenon first described by Bowen (1989)), in which two subsequent pulses may be seen as three flashes. In our paradigm, such a phenomenon may have shifted some judgments towards the "Double" percept. This phenomenon appears to be most frequent at timings around 100 ms, while the incidence is near zero at the shortest (0 ms) and longest (200 ms) timings (see also Purushothaman, Öğmen, & Bedell, 2003). These results apply to pulses displayed in the same spatial location. While in our baseline (same-offset) condition, the Vernier line segments are indeed stationary, in the more interesting opposite-offset condition, Vernier line segments are never shown in the same spatial location; it is therefore not clear if the 3-flashes illusion can at all occur in our experimental conditions. In any case, however, the 0 ms and 200 ms conditions can be assumed to be free or nearly free of such an artifact, further supporting the conclusion that Vernier fusion is occurring.

4.2. Limits of the current study

In order to measure the magnitude of the perceived (fused) offset of the Vernier stimuli, we presented subjects with a symmetric, pseudo-continuous scale of Vernier offsets (see Fig. 1). When choosing the perceived offset, subjects may have been repelled from the ends of the scale, and/or attracted by the center of the scale (0-offset). This may have caused or contributed to the general underestimation of spatial offsets, even in the same-offset condition; however, same-offset and different-offset stimuli would certainly have been affected in the same way, indicating that there was spatial compression during Vernier fusion, as expected.

In addition, the actual tick marks on the scale may have attracted individual judgments. In particular, the near-veridical judgments in the case of the smallest spatial offset (0.5′) would coincide with the very first tick mark of the scale. The small overall magnitude of the judgment combined with the lack of near-enough alternative choices (tick marks) may have led to a "snap-in" effect. With larger offsets, the relative distance between tick marks was reduced compared to the magnitude of the actual judgment, leading to less severe "snap-in".

4.3. Conclusions

We investigated the perceived spatial offsets and perceived types of presentation in a Vernier fusion paradigm on a trial-by-trial basis. We found the perceived magnitude of the (fused) Vernier offset to depend on the timing between the two stimulus presentations in a non-trivial way that was not reported in previous studies. The same kind of presentation may lead to substantially different perceived offset magnitudes, depending on the perceived type of presentation. Additionally, we found a dissociation between the involved mechanisms for temporal and spatial integration, allowing one to succeed even when the other fails.

Acknowledgments

We would like to thank Maddalena Costanzo for her help with data collection. We also thank Michael Herzog and one anonymous reviewer for their helpful suggestions.

This research was supported by a European Research Council (ERC) grant (Grant Agreement No. 313658). WZ was supported by a National Natural Science Foundation of China (62263042, 61005087), China Scholarship Council Grant.

References

Beard, B. L., Levi, D. M., & Klein, S. A. (1997). Vernier acuity with non-simultaneous targets: The cortical magnification factor estimated by psychophysics. *Vision Research, 37*(3), 325–346.

Bowen, R. W. (1989). Two pulses seen as three flashes: A superposition analysis. *Vision Research, 29*(4), 409–417. http://dx.doi.org/10.1016/0042-6989(89)90005-9.

Brand, A., Kopmann, S., & Herzog, M. (2004). Intact feature fusion in schizophrenic patients. *European Archives of Psychiatry and Clinical Neuroscience, 254*(5), 281–288. http://dx.doi.org/10.1007/s00406-004-0493-2.

Brand, A., Kopmann, S., Marbach, S., Heinze, M., & Herzog, M. H. (2005). Intact and deficient feature fusion in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience, 255*(6), 413–418. http://dx.doi.org/10.1007/s00406-005-0590-x.

Breitmeyer, B. G., & Öğmen, H. (2006). *Visual masking: Time slices through conscious and unconscious vision* (2nd ed.). Oxford University Press.

Efron, R. (1967). The duration of the present. *Annals of the New York Academy of Sciences, 138*(2), 713–729. http://dx.doi.org/10.1111/j.1749-6632.1967.tb55017.x.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4*(9), 345–352. http://dx.doi.org/10.1016/S1364-6613(00)01520-5.

Fahle, M. (1991). A new elementary feature of vision. *Investigative Ophthalmology & Visual Science, 32*(7), 2151–2155.

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision, 7*(8), 9. http://dx.doi.org/10.1167/7.8.9.

Harter, M. R. (1967). Excitability cycles and cortical scanning: A review of two hypotheses of central intermittency in perception. *Psychological Bulletin, 68*(1), 47–58.

Herzog, M. H., & Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proceedings of the National Academy of Sciences, 98*(7), 4271–4275. http://dx.doi.org/10.1073/pnas.071047498.

Herzog, M. H., Lesemann, E., & Eurich, C. W. (2006). Spatial interactions determine temporal feature integration as revealed by unmasking. *Advances in Cognitive Psychology, 2*(1), 77–85. http://dx.doi.org/10.2478/v10053-008-0046-8.

Herzog, M. H., Parish, L., Koch, C., & Fahle, M. (2003). Fusion of competing features is not serial. *Vision Research, 43*(18), 1951–1960. http://dx.doi.org/10.1016/S0042-6989(03)00278-5.

Herzog, M. H., Scharnowski, F., & Hermens, F. (2007). Long lasting effects of unmasking in a feature fusion paradigm. *Psychological Research, 71*(6), 653–658. http://dx.doi.org/10.1007/s00426-006-0062-6.

Korte, A. (1915). Kinematoskopische Untersuchungen. *Zeitschrift Für Psychologie, 72*, 193–296.

Mach, E. (1865). *Über die Wirkung der räumlichen Vertheilung des Lichtreizes auf die Netzhaut: vorgelegt in der Sitzung am 3. October 1865*. K. k. Hof- und Staatsdruckerei.

Purushothaman, G., Öğmen, H., & Bedell, H. E. (2003). Suprathreshold intrinsic dynamics of the human visual system. *Neural Computation, 15*(12), 2883–2908. http://dx.doi.org/10.1162/089976603322518786.

Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proceedings of the National Academy of Sciences of the United States of America, 93*(8), 3693–3697.

Rüter, J., Kammer, T., & Herzog, M. H. (2010). When transcranial magnetic stimulation (TMS) modulates feature integration. *The European Journal of Neuroscience, 32*(11), 1951–1958. http://dx.doi.org/10.1111/j.1460-9568.2010.07456.x.

Rüter, J., Sprekeler, H., Gerstner, W., & Herzog, M. H. (2013). The silent period of evidence integration in fast decision making. *PLoS ONE, 8*(1), e46525. http://dx.doi.org/10.1371/journal.pone.0046525.

Scharnowski, F., Hermens, F., & Herzog, M. H. (2007a). Bloch's law and the dynamics of feature fusion. *Vision Research, 47*(18), 2444–2452. http://dx.doi.org/10.1016/j.visres.2007.05.004.

Scharnowski, F., Hermens, F., Kammer, T., Öğmen, H., & Herzog, M. H. (2007b). Feature fusion reveals slow and fast visual memories. *Journal of Cognitive Neuroscience, 19*(4), 632–641. http://dx.doi.org/10.1162/jocn.2007.19.4.632.

Stroud, J. M. (1967). The fine structure of psychological time. *Annals of the New York Academy of Sciences, 138*(2 Interdiscipli.), 623–631. http://dx.doi.org/10.1111/j.1749-6632.1967.tb55012.x.

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences, 7*(5), 207–213.

Varela, F. J., Toro, A., John, E. R., & Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia, 19*(5), 675–686.

Watt, R. J. (1984). Towards a general theory of the visual acuities for shape and spatial arrangement. *Vision Research, 24*(10), 1377–1386. http://dx.doi.org/10.1016/0042-6989(84)90193-7.

Westheimer, G., & McKee, S. P. (1977). Spatial configurations for visual hyperacuity. *Vision Research, 17*(8), 941–947. http://dx.doi.org/10.1016/0042-6989(77)90069-4.

Wülfing, E. A. (1893). Über den kleinsten Gesichtswinkel. *Zeitschrift Für Biologie, XXIX*(2), 199–203.

Wutz, A., & Melcher, D. (2013). Temporal buffering and visual capacity: The time course of object formation underlies capacity limits in visual cognition. *Attention, Perception & Psychophysics, 75*(5), 921–933. http://dx.doi.org/10.3758/s13414-013-0454-9.

Wutz, A., Weisz, N., Braun, C., & Melcher, D. (2014). Temporal windows in visual processing: "Prestimulus brain state" and "poststimulus phase reset" segregate visual transients on different temporal scales. *The Journal of Neuroscience, 34*(4), 1554–1565. http://dx.doi.org/10.1523/JNEUROSCI.3187-13.2014.