Assessing the microstructure of motion correspondences with non-retinotopic feature attribution

Laboratory of Psychophysics, Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Thomas U. Otto

Center for Neuro-Engineering & Cognitive Science, University of Houston, USA, & Department of Electrical & Computer Engineering, University of Houston, USA

Haluk Öğmen

Michael H. Herzog

The motion correspondence problem, one of the classical examples of perceptual organization, addresses the question of how elements are grouped across space and time. Here, we investigate motion correspondences using a new feature attribution technique. We present, for example, a grating of four lines followed by a spatially shifted grating of three lines. Observers perceive a contracting grating. To study individual line-to-line correspondences, (1) we add, as a "perceptual marker," a small Vernier offset to one line of the first grating and (2) determine to which line of the second grating this offset is attributed. This procedure allows us inferring motion correspondences because this kind of feature attribution follows perceptual grouping in dynamic displays (H. Öğmen, T. U. Otto, & M. H. Herzog, 2006). Our results show that feature attribution between outer lines of the grating is more consistent than between inner lines. We interpret our results according to the principle of the "primacy of bounding contours," which states that bounding contours of an object provide a framework for element correspondences that is more important than the internal structure of that object.

Keywords: apparent motion, correspondence problem, ambiguity, Ternus–Pikler display, perceptual grouping, feature attribution, indirect measure


Introduction

The human visual system organizes discrete elements of a display into “perceptual wholes.” While much research dealt with the rules governing grouping operations across space only (e.g., Koffka, 1935), Pikler (1917) and Ternus (1926) addressed the question of grouping across space and time. When two elements are presented successively at different spatial locations, observers may perceive either a single moving object or two objects separated in space. In the first case, the two elements are grouped and, according to Ternus’ terminology, exhibit “phenomenal identity.” How the visual system establishes phenomenal identities in displays with multiple elements received considerable attention under the term motion correspondence problem (Attneave, 1974; Dawson, 1991; Marr, 1982; Ullman, 1979).

As an example, consider the two frame display depicted in Figure 1 adapted from Pikler (1917) and Ternus (1926). Three lines in the first frame (at positions a, b, and c) are followed by another three lines in the second frame (at positions b', c', and d'). In principle, each of the lines in the first frame can correspond to each of the lines in the second frame generating 3^3 potential matches (3! matches if one assumes one-to-one correspondences). However, only two of the possible correspondences are perceived depending on the blank interstimulus interval (ISI) between frames (e.g., Pantle & Picciano, 1976). For an ISI of 100 ms, observers perceive a grating of three lines moving from left to right (group motion, Figure 1A). For an ISI of 0 ms, observers perceive a single line moving from the outer left position to the outer right position thereby crossing two static lines (element motion, Figure 1B). Hence, with group motion, lines b and c'
Figure 1. Ternus–Pikler display. Three lines appear at positions a, b, and c in the first frame and, successively, at positions b', c', and d' in the second frame. (A) Group motion. For an ISI of 100 ms, the three lines a, b, and c are perceived to move as a group to b', c', and d', respectively (for an animation, see Supplementary Movie 1A). (B) Element motion. For an ISI of 0 ms, the line a is perceived to move to d' whereas the lines b and c are perceived as stationary (see Supplementary Movie 1B). (C) Non-retinotopic feature attribution. We inserted a small Vernier offset to one of the lines in the first frame (here at line b). Observers were instructed to attend to one of the lines in the second frame and to report the perceived offset direction. Note that lines in the second frame were not offset. In the case of group motion (for an ISI of 100 ms), the offset presented at line b is primarily perceived at line c' in the second frame (Ögmen et al., 2006). Hence, feature attribution is in accordance with the established motion correspondence.

exhibit “phenomenal identity,” whereas with element motion, this holds for lines b and b'.

In most studies of the correspondence problem, direct reports are employed; that is, observers’ task consists of reporting directly the perceived motion correspondence in the stimulus. For example, with the Ternus–Pikler display, observers are asked whether they perceive group or element motion (e.g., Alais & Lorenceau, 2002; Breitmeyer & Ritter, 1986; Dawson, Nevin-Meadows, & Wright, 1994; He & Ooi, 1999; Kramer & Rudd, 1999; Kramer & Yantis, 1997; Pantle & Petersik, 1980; Pantle & Picciano, 1976; Petersik, 1984; Scott-Samuel & Hess, 2001; for a tactile version, see Harrar & Harris, 2007; for a review, see Petersik & Rice, 2006). Such direct reports are well suited to determine the established motion correspondences for the ensemble of elements as a whole (or for individual elements in displays wherein a sparse set of elements undergoes simple motion). In contrast, for displays containing a relatively dense set of elements with ambiguous and complex motion (e.g., translation combined with expansion), observers may have difficulty reporting directly the microstructure of element correspondences; that is, the exact correspondences between individual elements in the display. In fact, individual line-to-line correspondences were only rarely investigated using direct reports (i.e., although the individual element-to-element correspondences are of interest, the observers’ task usually concerns the entire ensemble of elements; e.g., Burt & Sperling, 1981; Gepshtein & Kubovy, 2000, 2007). Like the stimuli in the aforementioned complex displays, many natural objects possess intricate texture patterns which create a formidable matching ambiguity during motion. Therefore, understanding of how the visual system analyzes element correspondences at the micro level is fundamental to an understanding of how texture, shape, and motion information are combined to establish object identities in natural scenes.

Using a version of the Ternus–Pikler display, we showed recently that feature attribution is in accordance with the perceived motion correspondence (Ögmen, Otto, & Herzog, 2006). We included, as a feature, a small Vernier offset to line b in the first frame and asked observers to attend to one of the lines in the second frame (which were all straight) and to indicate the perceived offset direction (Figure 1C). In accordance with the percept of group motion (for an ISI of 100 ms), the offset of line b in the first frame was primarily perceived at line c’ in the second frame and not at line b’ as it would be expected by retinotopic feature attribution. Hence, feature attribution is in accordance with the established motion correspondence.

Here, we propose that determining how features are attributed between individual elements can be used as a tool complementary to direct reports to investigate the microstructure of motion correspondences. In analogy to biological marker techniques, the Vernier offset in the first frame serves as a “perceptual marker” and, by determining the element of “reappearance” in the second (or later) frame, we can indirectly infer the perceived motion correspondences. Moreover, by employing an offset discrimination task, we do not ask observers to judge their motion percept, and consequently, no cognitive strategies or directly motion-related criteria are imposed by our task demands.

We applied this method to Ternus–Pikler displays that are composed of line gratings with different numbers of elements in the two frames. For example, we presented a grating of four lines in the first and a grating of three lines in the second frame. In this case, no unique one-to-one correspondences between individual lines can be established because of the mismatch in the number of lines presented in the two frames. Such displays elicit the percept
of transformational group motion; that is, observers report a moving and transforming line grating (e.g., a contracting grating in the case of 4 + 3 lines). Using the feature attribution as a measure for perceptual organization, we show, for example, that outer lines of the gratings—as the lines $a$ and $b'$ in Figure 1—have a privileged role in solving the motion correspondence problem.

### Methods

#### Observers

Data were obtained from one of the authors (T.O.) and naive paid students. The general purpose of the experiment and the possible consequences of the studies were explained to each observer. Moreover, observers were told that they could quit the experiment at any time they wished. After observers signed informed consent, we determined visual acuity by means of the Freiburg visual acuity test (Bach, 1996). To participate in the experiments, observers had to reach a value of 1.0 at least for one eye (corresponding to a Snellen ratio of 20/20). The experiments were undertaken with the permission of the local ethics committee.

#### Apparatus

Stimuli appeared on an X-Y-display (HP-1332A, Tektronix 608) controlled by a PC via fast 16 bit D/A converters. Line stimuli were composed of dots drawn with a dot pitch of 250–350 $\mu$m at a dot rate of 1 MHz. The dot pitch was selected so that dots slightly overlapped; that is, the dot size (or line width) was of the same magnitude as the dot pitch. Stimuli were refreshed at 200 Hz. Luminance of the stimuli was 80 cd/m$^2$ as measured with a Minolta LS-100 luminance meter by means of a dot grid (with the same dot pitch and refresh rate as above). Background luminance on the screen was below 1 cd/m$^2$. The room was dimly illuminated (approximately 0.5 lx). Viewing distance was 2 m.

#### Stimuli

We presented variations of the Ternus–Pikler display as used by Ögmen et al. (2006). We presented three or four lines in the first frame followed by other three or four lines that were shifted to the right in the second frame. The length of lines was 1260$''$ (arcsec) including a central gap of 60$''$. The spacing between lines was 800$''$. Frames were presented for 70 ms each and separated by a blank screen of 100 ms. Hence, the total stimulus duration was 240 ms.

### Procedures

Each trial was initiated with four markers at the corners of the screen presented for 500 ms followed by a blank screen for 200 ms. Then, the actual stimulus was presented. After stimulus presentation, a blank screen appeared until observers responded. A new trial was initiated 500 ms after the observer gave a response.

Stimuli were presented in blocks of 80 trials. In each block, we inserted a spatial (Vernier) offset to one of the lines in the first frame; that is, the lower line segment was slightly offset randomly to the left or right with respect to the upper segment. At the beginning of a block, we instructed observers to attend to one of the lines in the second frame (see Figure 1C). We presented no fixation dot to ease this deployment of attention. In a binary forced choice task, observers were asked to report the perceived offset direction of this attended line by pressing one of two buttons. Observers pushed the left (right) button when the lower line segment was perceived offset leftward (rightward) with respect to the upper segment. Note that the lines in the second frame were not offset. Naive observers had no knowledge about where the Vernier offset was presented. No feedback was given.

In order to achieve comparable performance levels across observers, we determined individual offset discrimination thresholds using an adaptive staircase method. Test levels were chosen following the PEST procedure (Taylor & Creelman, 1967; start value: 70$''$; initial step size: 3.0 dB; Wald constant: 1.5; 80 trials). We estimated the threshold (and the slope) of the psychometric function (cumulative Gaussian; chance level: 50%; rate of response lapses: 2.5%) by means of a maximum likelihood analysis, taking all trials into account (left and right offset directions were pooled). We determined thresholds using the standard display with an offset at line $b$ in the first frame and asked observers to attend to line $c'$ in the second frame (see Figure 1C). For each observer, we used an offset size according to the individual threshold level throughout the experiments. Individual offset sizes are specified in the particular result section.

In each experiment, block by block, we measured all combinations of offset line in the first frame and attended line in the second frame. The order of conditions was randomized across observers to reduce the influence of hysteresis, learning, or fatigue effects in the averaged data. For each observer, each condition was measured twice (i.e., 160 trials per observer). After each condition had been measured once, the order of conditions was reversed for the second set of measurements.

### Data analysis

Block by block, observers attended to one straight line in the second frame while one line in the first frame was
offset. Assume that the first frame contains \( m \) lines, each denoted by \( x_i \) with \( 1 \leq i \leq m \). Similarly, assume that the second frame contains \( n \) lines, each denoted by \( x_j' \) with \( 1 \leq j \leq n \). When line \( x_i' \) was attended and line \( x_i \) was offset, we determined as a quantitative measure \( P[x_i, x_i'] \), which is the percentage of responses in accordance with the offset direction of line \( x_i \) in the first frame.

We created correspondence diagrams to visualize how features are attributed between lines in the first and the second frame. In a correspondence diagram, the width of a connecting line between line \( x_i \) in the first frame and line \( x_j' \) in the second frame corresponds to the difference between \( P[x_i, x_j'] \) and 50%. If this difference was smaller than 5%, no connecting line between the corresponding positions was drawn for graphical clarity.

Next, we computed for each observer the mean accordance level \( \overline{P}[x_i] \) across all conditions in which line \( x_i \) was offset. \( \overline{P}[x_i] \) is given by

\[
\overline{P}[x_i] = \frac{1}{n} \sum_{j=1}^{n} P[x_i, x_j'].
\]  

Similarly, we determined the mean accordance level \( \overline{P}[x_j'] \) across all conditions in which line \( x_j' \) was attended. \( \overline{P}[x_j'] \) is given by

\[
\overline{P}[x_j'] = \frac{1}{m} \sum_{i=1}^{m} P[x_i, x_j'].
\]  

Next, we analyzed the distribution of feature attribution. The offset of a line in the first frame can be attributed specifically to one line in the second frame or rather loosely to more lines. To quantify this aspect, we computed for each observer the feature distribution index in terms of the fan-out concentration, \( D_{fo} \), for each line \( x_i \) by

\[
D_{fo}[x_i] = \frac{\max_{j} \left( P[x_i, x_j'] \right) - \overline{P}[x_i]}{\overline{P}[x_i]}.  
\]

If the offset of line \( x_i \) is attributed broadly across lines in the second frame, \( D_{fo}[x_i] \) is low. If the offset of line \( x_i \) is primarily attributed to one line, \( D_{fo}[x_i] \) is high.

Similarly, offsets attributed to one attended line can originate either primarily from one line in the first frame or from two or more lines. To quantify this aspect, we computed the feature distribution index in terms of the fan-in concentration for each line \( x_j' \) by

\[
D_{fi}[x_j'] = \frac{\max_{i} \left( P[x_i, x_j'] \right) - \overline{P}[x_j']}{\overline{P}[x_j']}. \]

If the offset perceived at line \( x_j' \) originated from two or more lines in the first frame, \( D_{fi}[x_j'] \) is low. If the offset perceived at line \( x_j' \) originated primarily from one line, \( D_{fi}[x_j'] \) is high.

To compare differences among offset lines in the first frame and attended lines in the second frame, respectively, we computed one-way analyses of variance with repeated measures (ANOVA), with Greenhouse–Geisser correction. Significant ANOVA outcomes were followed by pairwise least significant difference (LSD) comparisons.

Results

Experiment I: Four plus four lines

We presented a Ternus–Pikler display with four lines in both frames (Figure 2A). This experiment will be the basis of the ensuing experiments.

Methods

In one block, only one of the four lines in the first frame (labeled as \( a, b, c, \) and \( d \)) was offset. Individually determined offset sizes ranged from 30" to 40", (mean: 37", see Procedures). Block by block, five observers were asked to attend to one of the four lines in the second frame (labeled as \( b', c', d', \) and \( e' \)) and to discriminate the perceived offset direction. We tested all sixteen combinations of offset line in the first frame and attended line in the second frame.

Results and discussion

For a Ternus–Pikler display with three lines in each frame, a Vernier offset of the first frame is perceived retinotopically mislocalized in the second frame (Ögmen et al., 2006). For a display with four lines, we find a very similar result. A Vernier offset is primarily perceived at that line in the second frame that was shifted by one position to the right (Figure 2B). For example, if line \( b \) was offset, this offset is primarily reported when line \( c' \) was attended (see Supplementary Movie 2B). The pattern of feature attribution as a whole is visualized in the correspondence diagram (Figure 2C). Feature attribution according to the percept of group motion is evident from the strong connections between lines of the first frame and lines of the second frame which are shifted by one position to the right.

To analyze feature attribution in more detail, first, we computed the mean accordance level for each offset line of the first (Equation 1) and each attended line of the second frame (Equation 2), respectively. For offset lines of the first frame, we find no significant differences in mean accordance levels (Figure 2D). Note that the outer line \( d \) was followed by an overlapping line in the second frame while outer line \( a \) was not (see Figure 2A). Still, mean accordance levels for lines \( a \) and \( d \) are comparable.
Hence, differences in backward masking, which might have occurred because of the overlapping lines in the second frame, seem not to be at work here. For the second frame, the mean accordance level for line $e'$ seems to be slightly reduced compared to lines $b'$, $c'$, and $d'$ (Figure 2E). However, this difference fails to be significant. This trend might be caused by the fact that some feature attribution has occurred for lines presented at the same retinotopic position (e.g., $d \rightarrow d'$) in addition to the main effect of feature attribution according to group motion. For the attended line $e'$, this is not possible because there was no preceding line at position $e$ in the first frame.

Second, we analyzed whether an offset presented at a particular line in the first frame is attributed specifically to one line in the second frame or rather loosely to more lines. For example, considering the correspondence diagram (Figure 2C), does feature attribution differ for line $a$ compared to line $b$? To quantify this aspect, we computed the feature distribution index $D_{fo}$ for each line in the first frame (Equation 3). A high value of $D_{fo}$ indicates a specific distribution whereas a small value indicates a loose distribution. As shown in Figure 2F, $D_{fo}$ seems to be slightly elevated for line $a$ compared to lines $b$, $c$, and $d$. However, these differences are not significant. This trend might reflect that some retinotopic feature attribution has occurred. Such retinotopic feature attribution could not occur for line $a$ because it was not followed by an overlapping line in the second frame.

Analogously, we determined the distribution index $D_{fi}$ for each attended line (Equation 4). An ANOVA showed a
significant effect of attended line \((F(2,1,8,6) = 5.58, p = 0.029, \text{Greenhouse–Geisser corrected; Figure 2G})\). According to post hoc LSD comparisons, \(D_{\alpha}\) differs for attended lines \(b'\) and \(d'\) (mean difference: 0.17; \(p = 0.026\)) and for \(c'\) and \(e'\) (mean difference: \(-0.19; p = 0.020\)). Moreover, although not significant, the difference in \(D_{\alpha}\) for lines \(b'\) and \(c'\) reveals a trend (mean difference: 0.20; \(p = 0.058\)). Hence, while feature attribution is primarily determined by the percept of group motion, attribution is more specific when the outer lines of the second frame were attended.

This finding might be influenced by two extraneous factors. First, in some trials, observers may have erroneously attended to another line than instructed. This is more likely to occur for central lines than for outer lines and hence causing lower values of \(D_{\alpha}\) for central lines. To rule out this possibility, we conducted a control experiment in which we measured discrimination thresholds for real Vernier offsets presented in the second frame. For these thresholds, we find no significant effects of attended line (see Supplementary data). Hence, the less specific feature attribution for central lines is not explained by attentional lapses.

Second, when observers attended the various lines of the second frame, the eccentricity of the stimulus changed slightly. For example, when the inner line \(c'\) was in the center of gaze compared to the outer line \(b'\), the stimulus was shifted by about 800'' (i.e., less than 0.25 deg). This eccentricity difference might change the probability of group motion and consequently might change feature attribution. However, this is unlikely for two reasons. First, we tested feature attribution for displays with 3 + 3, 4 + 4, and 5 + 5 lines. Although the number and eccentricity of unattended lines was changed, we found the same result for all displays: feature attribution is less specific for inner lines than for outer lines (data not shown for the 3 + 3 and the 5 + 5 display). Second, in a previous study with the 3 + 3 display, the pattern of feature attribution was the same when the spacing between lines was either 800'' or 1600'' (Ögmen et al., 2006). Hence, although the eccentricity was changed, feature attribution was not.

To summarize, in our previous study, we find that feature attribution follows primarily the percept of group motion. We find that feature attribution is more specific when outer lines in the second frame were attended compared to central ones. In the following experiments, we induce correspondence ambiguities by varying the number of the lines in the first and/or second frame to study how these changes influence feature attribution and the microstructure of element correspondences.

**Experiment II: Four plus three lines**

In Experiment I, both frames contained 4 lines. Hence, one-to-one correspondences could be established. In this and the next experiment, we presented ambiguous displays with different numbers of lines across frames. Thus, a one-to-one mapping of lines is impossible. These displays lead to the perception of group motion with figural transformations in form of an expanding or contracting line grating.

**Methods**

We presented a display with four lines in the first and three lines in the second frame (Figure 3A). This display corresponds to the display in Figure 2A by removing line \(b'\) in the second frame. In the first frame, only one line was offset. Individually determined offset sizes ranged from 30'' to 40'' (mean: 36''). We asked five observers to attend to one line of the second frame and to discriminate the perceived offset direction. We tested all twelve combinations of offset line and attended line.

**Results and discussion**

Although the number of lines in the first frame did not match the number of lines in the second frame, a percept of group motion is elicited (see Supplementary Movie 3). Accordingly, as in the previous experiment, we find that the offset, which was presented in the first frame, is primarily perceived in the second frame to be shifted in the direction of the group motion (Figures 3B and 3C). Remarkably, when line \(c'\) was attended, performance is higher when line \(a\) than when line \(b\) was offset, although the distance \(a \rightarrow c'\) was larger than \(b \rightarrow c'\).

We were particularly interested how feature attribution was influenced by the ambiguity in the display. First, if one line in the first frame did not attribute its offset to any of the lines in the second frame, the corresponding mean accordance level should drop to 50%. However, for offset lines in the first frame, we find no significant differences in mean accordance levels (Figure 3D). Hence, as in the previous experiment, all lines of the first frame attribute “their” offset to the second frame. Likewise, for the attended lines of the second frame, we find a pattern with no significant differences in mean accordance levels comparable to the previous experiment (Figure 3E).

Next, we computed the feature distribution indices. For \(D_{10}\), an ANOVA showed a significant effect of offset line \((F(1,6,6.3) = 11.94, p = 0.009, \text{Greenhouse–Geisser corrected; Figure 3F})\). Post hoc LSD comparisons revealed that \(D_{10}\) differed for lines \(a\) and \(b\) (mean difference: 0.19; \(p = 0.008\)), for \(a\) and \(c\) (mean difference: 0.12; \(p = 0.001\)), and for \(b\) and \(d\) (mean difference: \(-0.16; p = 0.004\)). This finding indicates that feature attribution for outer lines of the first frame is more precise than for central lines and especially for line \(b\). This finding is also visualized by the correspondence diagram showing weak but equal correspondences from line \(b\) to lines \(c'\) and \(d'\) (Figure 3C). For lines of the second frame, \(D_{\alpha}\) seems to be elevated for line \(e'\) compared to lines \(c'\) and \(d'\) (Figure 3G). However, these differences are not significant. Note that, compared to the previous experiment, line \(b'\) was not displayed here.
To summarize, we induced a correspondence ambiguity by presenting four lines in the first but only three lines in the second frame. Despite this mismatch, each line of the first frame attributes to the second frame as expressed by roughly comparable mean accordance levels. Moreover, our results indicate a strong preference to establish correspondence matches between the outer lines and bounding contours. Accordingly, the correspondence ambiguity induced by the unequal number of lines seems to concern mostly the central line, which attributes its offset partially to lines c' and d'.

Experiment III: Three plus four lines

In this experiment, we studied the inverse case of Experiment II; that is, we presented three lines in the first and four lines in the second frame. The display corresponds to the display in Figure 2A by removing line d in the first frame (Figure 4A). One line in the first frame was offset (here at line c). Observers attended to one line of the second frame. Offsets presented at the outer lines a and d are primarily perceived at the outer lines c' and d' in the second frame. The offset of line c is primarily perceived at line d', whereas the attribution of the offset of line b does not follow a clear trend.

Results and discussion

As in Experiments I and II, according to the ISI of 100 ms, a percept of group motion is elicited. In addition, the grating seems to expand in the direction of the group motion (see Supplementary Movie 4).
Regarding the first frame, mean accordance levels seem to increase from line \(a\) to line \(c\) (Figure 4D). However, these differences fail to be significant. For the second frame, an ANOVA shows a significant effect of attended line \((F(2.2,11.0) = 7.10, p = 0.009, \text{Greenhouse–Geisser corrected}; \text{Figure 4E})\). Post hoc LSD comparisons revealed that mean accordance levels differed for lines \(b'\) and \(d'\) (mean difference: 6.7%; \(p = 0.050\)), for \(b'\) and \(e'\) (mean difference: 8.6%; \(p = 0.009\)), and for \(c'\) and \(e'\) (mean difference: 7.5%; \(p = 0.019\)). Thus, mean accordance levels decreased from line \(b'\) to line \(e'\). Hence, the study of mean accordance levels indicates some net transfer of feature attribution from the right side of the grating in the first to the left side in the second frame.

As in the previous experiments, we determined the feature distribution indices. For the first frame, an ANOVA shows a significant effect of offset line \((F(1.7,8.3) = 24.34, p < 0.001, \text{Greenhouse–Geisser corrected}; \text{Figure 4F})\). Post hoc LSD comparisons revealed that \(D_{fo}\) differs for offset lines \(a\) and \(b\) (mean difference: 0.18; \(p = 0.005\)) and for \(a\) and \(c\) (mean difference: 0.20; \(p < 0.001\)). Hence, feature attribution is less specific for lines \(b\) and \(c\) compared to line \(a\). This is also evident in the correspondence diagram showing one strong connection originating from line \(a\) whereas two or more connections originate from lines \(b\) and \(c\) (Figure 4C).

For the second frame, \(D_{fo}\) seems to increase slightly from line \(c'\) to \(e'\) and to be highest for line \(b'\) (Figure 4G). However, these differences are not significant. Moreover, values of \(D_{fo}\) are in general lower than in the previous experiments. One explanation might be that the lower indices reflect the slightly lower accordance levels in this compared to the previous experiments (e.g., \(P[a, b'] = 80.9\%\); \textbf{Experiment I}: \(P[a, b'] = 89.3\%\)).

To summarize, we find that the outer line \(a\) attributes “its” offset primarily to the corresponding outer line \(b'\) whereas the outer line \(c\) attributes to multiple lines of the second frame. Hence, the induced correspondence ambiguity seems to concern mainly the right side of the line grating. This finding is in good accordance with the
subjectively observed expansion of the grating (see Supplementary Movie 4).

Note that the stimuli shown in Experiments II and III were symmetric; that is, the spatial distances between individual lines were the same in both displays, only the temporal order of frames was reversed. Despite the spatial symmetry, we find that the established correspondences differ. When the grating was contracting, we find strong feature attribution between the outer lines. When the grating was expanding, a strong feature attribution was found for the left outer lines \((a \text{ and } b^\prime)\) in Figure 4 but not for the right ones \((c \text{ and } e^\prime)\) in Figure 4. This difference can be understood by considering the fact that contraction occurs mainly for the left outer line from its initial position \(a\) to its final position \(c^\prime\) (Figure 3). As contraction takes place gradually from \(a\) to \(c^\prime\), there are no other elements in the second frame on the path of contraction to act as an “intermediate outer line” during this gradual contraction. On the other hand, when expansion takes place gradually from \(c\) to \(e^\prime\), elements \(e^\prime, d^\prime,\) and \(e^\prime\) are on the expansion path and can act as intermediate outer lines during this gradual expansion. As a result, one may expect the outer line \(c\) to attribute its offset to \(c^\prime, d^\prime,\) and \(e^\prime\), which act as progressive positions of the expanding outer line. Hence, motion correspondences depend not only on the distances between single lines but also on the transformation of the group of elements.

**Experiment IV: Three plus three lines—gap at \(d^\prime\)**

In the previous experiments, we induced correspondence ambiguities by presenting unequal numbers of lines across frames. In this and the next experiment, we presented displays with three lines in each frame. Hence, as in Experiment I, a one-to-one mapping of lines is possible. However, we induced a “gap” in the grating of lines in the second frame to investigate whether inhomogeneous distances between lines change feature attribution and motion correspondence.

**Methods**

The display corresponded to Figure 2A by removing line \(d\) in the first and line \(d^\prime\) in the second frame (Figure 5A). Individually determined offset sizes ranged from 30° to 40° (mean: 33°). We asked five observers to attend to one line in the second frame and to discriminate the perceived offset direction. We tested all nine combinations of offset line and attended line.

**Results and discussion**

As in the previous experiments, a group motion percept is elicited. However, the grating of lines seems to split into two parts during motion (see Supplementary Movie 5). Lines \(a\) and \(b\) appear to move to positions \(b^\prime\) and \(c^\prime\), respectively. The line \(c\) seems to move to position \(e^\prime\) but appears to be faster and to separate from the others. Despite this inhomogeneous motion percept, the pattern is rather comparable to Experiment I. Offsets presented in the first frame are primarily attributed to the right (Figure 5B).

The mean accordance level for line \(b\) seems to be reduced compared to lines \(a\) and \(c\) (Figure 5D). This trend is comparable to the trend of slightly lower mean accordance levels for central lines found with the 4 + 4 display. For the second frame, an ANOVA shows a significant effect of attended line \((F(1,3,5,2) = 9.07, p = 0.025,\) Greenhouse–Geisser corrected; Figure 5E). Post hoc LSD comparisons reveal that mean accordance levels differ for lines \(c^\prime\) and \(e^\prime\) (mean difference: 6.9%; \(p = 0.031\)). Moreover, although not significant, the differences in mean accordance levels for lines \(b^\prime\) and \(c^\prime\) (mean difference: \(-4.1%; p = 0.057\)) and for lines \(b^\prime\) and \(e^\prime\) (mean difference: 2.8%; \(p = 0.056\)) reveal a trend. This finding bears analogy to the trend of a reduced mean accordance level for line \(e^\prime\) with the 4 + 4 display.

According to feature distribution indices, we find no significant differences across offset lines of the first frame (Figure 5F). For the second frame, an ANOVA shows a significant effect of offset line \((F(1,8,7,3) = 13.99, p = 0.004,\) Greenhouse–Geisser corrected; Figure 5G). Post hoc LSD comparisons reveal that \(D_b\) differs for lines \(b^\prime\) and \(c^\prime\) (mean difference: 0.15; \(p = 0.005\)) and for \(c^\prime\) and \(e^\prime\) (mean difference: \(-0.12; p = 0.023\)). Hence, feature attribution is more specific when the outer lines of the second frame were attended. This is also evident in the correspondence diagram showing only one strong connection terminating at lines \(b^\prime\) and \(e^\prime\), respectively, but multiple connections terminating at line \(c^\prime\) (Figure 5C).

To summarize, we presented three lines in both frames. Hence, a one-to-one mapping of lines is possible as in Experiment I. Here, we introduced a gap in the grating of the second frame. Despite this gap, the pattern of feature attribution is comparable to Experiment I. Hence, the inhomogeneous distances between lines seem not to change feature attribution and motion correspondences in this case.

**Experiment V: Three plus three lines—gap at \(c^\prime\)**

We repeated the previous experiment with a gap at position \(c^\prime\) in the grating of lines of the second frame.

**Methods**

The display corresponds to Figure 2A by removing line \(d\) in the first frame and line \(c^\prime\) in the second frame.
The same five observers as in the previous experiment participated, and we used the same offset sizes as above. We tested all nine combinations of offset line and attended line.

**Results and discussion**

While the percept of group motion is elicited, the grating of lines seems to split in two parts similar to the previous experiment (see Supplementary Movie 6).

For mean accordance levels, for the first frame, an ANOVA shows a significant effect of offset line ($F(1.0,4.1) = 7.85$, $p = 0.047$, Greenhouse–Geisser corrected; Figure 6D). Post hoc LSD comparisons revealed that mean accordance levels differed for lines $b$ and $c$ (mean difference: $-6.4\%$; $p = 0.001$). For the second frame, the mean accordance level for line $e$ seems to be slightly lower compared to lines $b'$ and $d'$ (Figure 6E). However, this difference failed to be significant. Overall, these findings are rather comparable to the previous experiment.

According to feature distribution indices, for the first frame, an ANOVA shows a significant effect of offset line ($F(1.2,4.6) = 25.34$, $p = 0.005$, Greenhouse–Geisser corrected; Figure 6F). Post hoc LSD comparisons reveal that $D_{fo}$ differs for offset lines $a$ and $b$ (mean difference: $-0.17$; $p < 0.001$) and for lines $a$ and $c$ (mean difference: $-0.13$; $p = 0.012$). Hence, feature attribution is less specific for lines $b$ and $c$ compared to line $a$. Such a difference is not observed in the previous experiment. Moreover, the feature attribution index for line $b$ seems to be reduced compared to the previous experiment (post hoc analysis; two-tailed, paired $t$ test: $p = 0.014$). Hence, with the gap at position $d'$, feature attribution is less specific for line $b$ compared to the display with a gap at position $d'$.

For the second frame, an ANOVA shows a significant effect of attended line ($F(1.7,6.7) = 17.56$, $p = 0.001$, two-tailed, paired $t$ test: $p = 0.001$). The cluster $D_{a}$ differs for attended lines $a$ and $b$ (mean difference: $-0.17$; $p < 0.001$) and for lines $a$ and $c$ (mean difference: $-0.14$; $p = 0.001$). Hence, feature attribution is less specific for line $b$ compared to the display with a gap at position $d'$.

(Figure 6A). The same five observers as in the previous experiment participated, and we used the same offset sizes as above. We tested all nine combinations of offset line and attended line.
Post hoc LSD comparisons reveal that $D_{fi}$ differed for lines $bV$ and $dV$ (mean difference: 0.21; $p = 0.008$) and for line $dV$ and $eV$ (mean difference: 0.18; $p = 0.003$). Hence, attribution is more specific when the lines $bV$ and $eV$ were attended compared to line $dV$. This is also evident in the correspondence diagram showing only one strong connection terminating at lines $bV$ and $eV$, respectively, and two weaker connections terminating at line $dV$ (Figure 6C). This finding indicates that the correspondence for the central line $dV$ is less specific compared to the central line $cV$ in the previous experiment.

To summarize, despite the similarity of the displays here and in Experiment IV, we observe different patterns of feature attribution. With a gap at position $d'$, feature attribution is in favor of a unique one-to-one matching. With a gap at position $c'$, our results reveal a correspondence ambiguity. Hence, although a one-to-one matching of lines is possible in principle, different gap positions can yield different element correspondences.

**Discussion**

**Non-retinotopic feature attribution—A measure for motion correspondences**

When static elements are presented in successive frames, the vivid illusion of moving objects is often elicited. To generate this percept, the visual system has to establish correspondences between subsequent elements. However, establishing such motion correspondences is an ill-posed problem because each element in one frame can correspond, in principle, to any of the elements in the next frame (e.g., Attneave, 1974; Dawson, 1991; Marr, 1982; Ternus, 1926; Ullman, 1979).

To investigate how the visual system solves the motion correspondence problem direct reports are usually used as, for example, the rating of the motion “smoothness” or choosing between two (or more) possible motion percepts. In studies on the Ternus–Pikler display, typically, the

Figure 6. Gap at $c'$. (A) We presented a similar display as shown in Figure 5. Here, the gap in the line grating was at position $c'$ (see Supplementary Movie 6). In this example, line $b$ is offset. (B) Accordance levels for all combinations of offset line and attended line. When the outer lines $b'$ and $e'$ were attended, offsets of lines $a$ and $c$ are primarily perceived, respectively. When line $d'$ was attended, offsets of lines $b$ and $c$ are perceived. (C) Correspondence diagram. Compared to Figure 5C, the central line $d'$ seems to correspond to the outer line $c$ and to the central line $b$. (D and E) Mean accordance levels. (F and G) Feature distribution indices. Means and SEM for five observers.
latter method is employed by asking observers whether they perceive element or group motion (e.g., Alais & Lorenceau, 2002; Breitmeyer & Ritter, 1986; Dawson et al., 1994; He & Ooi, 1999; Kramer & Rudd, 1999; Kramer & Yantis, 1997; Pantle & Petersik, 1980; Pantle & Picciano, 1976; Petersik, 1984; Scott-Samuel & Hess, 2001). It is also possible to design displays for which the motion direction for element motion differs from the direction for group motion allowing observers to choose between element and group motion by indicating one or the other motion direction (Gepshtein & Kubovy, 2000). Most studies using direct reports have in common that the observer judges the global motion percept rather than motion correspondences based on individual elements. Moreover, a quantification of the element-to-element correspondences is usually rather difficult (see Supplementary movies).

Here, we introduce a complementary approach to determine and to quantify the microstructure of motion correspondences by taking advantage of the recently discovered non-retinotopic feature attributions in motion displays (see also, Nishida, Watanabe, Kuriki, & Tokimoto, 2007; Otto, Ögmen, & Herzog, 2006; Shimozaki, Eckstein, & Thomas, 1999). With the Ternus–Pikler display, for example, we showed previously that feature attribution for single elements is in accordance with the global motion percept of either element or group motion (Ögmen et al., 2006). Here, we determine and quantify the line-to-line correspondences by measuring feature attribution systematically for all possible combinations of lines in a Ternus–Pikler display. In Experiment I, our findings indicate primarily one-to-one “correspondences” between lines according to the global group motion percept (Figure 2B). However, there was also a small amount of feature attribution not in accordance with the group motion, which might have occurred for two reasons. First, these cases reflect feature attribution violating the established motion correspondence or, second, point to a correspondence ambiguity causing an unspecific feature attribution. Interestingly, most of these “erroneous” feature attributions occur retinotopically; that is, according to element motion. This might indicate that, with an ISI of 100 ms, the visual system interprets element motion as a second but much less likely solution to solve the motion correspondence problem. Moreover, we never found feature attribution in the opposite direction to the global motion percept (e.g., $c \rightarrow b'$), which, from a pure spatial perspective, is as likely as attribution in the global motion direction (e.g., $b \rightarrow c'$). Hence, we think it is justified to assume that cases with ambiguous feature attribution point to ambiguous motion correspondences.

While we found previously non-retinotopic feature attribution to occur in a range of up to about 0.5 deg between single elements (Ögmen et al., 2006), the exact spatiotemporal limits wherein our feature attribution approach can be used remains to be determined. Moreover, as a limitation for our approach, we note that feature attribution is strongly reduced for the “jumping line” in the case of element motion (e.g., $a \rightarrow d'$ in Figure 1, unpublished data). It seems that static elements along the motion trajectory can block feature attribution.

**Element correspondences and transformational group motion**

In the Experiments II–V, we applied the feature attribution approach to Ternus–Pikler displays, in which we varied the configuration of the line gratings across frames. We were primarily interested to investigate how these figural transformations of the line grating affect feature attribution and the motion correspondences.

When an object moves in a scene, its shape may undergo strong spatial transformations. For example, when a square is followed by a circle, observers perceive a square that changes its shape into a circle (e.g., Kolers, 1972; Kolers & Pomerantz, 1971). Despite this large change in appearance, the square and the circle exhibit “phenomenal identity” because they are perceived as one object. Hence, shape differences do not prevent the establishment of motion correspondences for single elements. Accordingly, many researchers assumed that “figural identity” plays only a minor role in the assignment of motion correspondences (e.g., Dawson, 1991; Kolers, 1972; Navon, 1976; Ullman, 1979) and investigated primarily the effects of the spatiotemporal distances between single elements in motion displays (e.g., Burt & Sperling, 1981; Gepshtein & Kubovy, 2007; Korte, 1915). In our experiments, we presented identical lines (except for the small Vernier offset of one line in the first frame), and hence we did not investigate whether the “figural identity” of single elements affects motion correspondences (e.g., Green, 1986; Kramer & Yantis, 1997; Nishida, Ohtani, & Ejima, 1992; Nishida & Takeuchi, 1990; Scott-Samuel & Hess, 2001; Shechter, Hochstein, & Hillman, 1988; Werkhoven, Snippe, & Koenderink, 1990; Werkhoven, Sperling, & Chubb, 1993, 1994). Instead, we changed the configuration of the line gratings across frames and investigated how these differences affect the microstructure of element correspondences.

In Experiment II, for example, we presented a grating of four lines followed by a shifted grating of three lines. In analogy to the transformation of a square to a circle, observers perceived a grating that appears to contract in appearance, the square and the circle exhibit “phenomenal identity” because they are perceived as one object. Hence, shape differences do not prevent the establishment of motion correspondences for single elements. Accordingly, many researchers assumed that “figural identity” plays only a minor role in the assignment of motion correspondences (e.g., Dawson, 1991; Kolers, 1972; Navon, 1976; Ullman, 1979) and investigated primarily the effects of the spatiotemporal distances between single elements in motion displays (e.g., Burt & Sperling, 1981; Gepshtein & Kubovy, 2007; Korte, 1915). In our experiments, we presented identical lines (except for the small Vernier offset of one line in the first frame), and hence we did not investigate whether the “figural identity” of single elements affects motion correspondences (e.g., Green, 1986; Kramer & Yantis, 1997; Nishida, Ohtani, & Ejima, 1992; Nishida & Takeuchi, 1990; Scott-Samuel & Hess, 2001; Shechter, Hochstein, & Hillman, 1988; Werkhoven, Snippe, & Koenderink, 1990; Werkhoven, Sperling, & Chubb, 1993, 1994). Instead, we changed the configuration of the line gratings across frames and investigated how these differences affect the microstructure of element correspondences.
In the case of three lines in the first frame and four lines in the second frame (Figure 4A), the percept of an expanding grating is elicited (see Supplementary Movie 4). This percept of an expanding grating bears some similarity to the line motion illusion—when a dot is followed by a line, the percept of one line expanding from the position of the dot is elicited (e.g., Hikosaka, Miyachi, & Shimojo, 1993; for an early report of the illusion, see Pikler, 1917). In agreement with this subjective impression, we found correspondences between the outer line c in the first frame, which is in the leading position during motion, and various lines in the second frame (Figures 4B and 4C; expanding outer line: c → c', d', e'). Hence, in the case of expansion, the attribution from the outer line c to multiple lines in the second frame seems to reflect the gradual expansion of the right bounding contour from its original position c to its final position e'.

The stimuli presented in Experiments II and III are symmetric in the sense that the spatial distances between single lines were the same in both displays, only the temporal order of frames was reversed. For example, the correspondence a → c' in Figure 3 matches the inverted correspondence e' → c in Figure 4. Hence, when the visual system solves the motion correspondence problem primarily by spatiotemporal cues, the same line-to-line correspondences should be established in both displays. However, we find that the established correspondences differ in the corresponding experiments depending on the figural transformation of the motion group (i.e., contraction or expansion). Still, in both cases, the outer lines of the grating seem to have a privileged role in solving the correspondence problem.

In Experiments IV and V, we presented displays with three lines in each frame and included a “gap” between lines in the second frame. Subjectively, the regular line grating, presented in the first frame, splits into two parts in the second frame (see Supplementary Movies 5 and 6). The gap in the line grating should not affect the line correspondences because a one-to-one mapping of lines is possible in both cases (e.g., a → b', b → c', and c' → e' in Figure 5 and a → b', b → d', and c' → e' in Figure 6). However, our results indicate that the gap in the grating changes the element correspondences. Hence, figural transformations of the moving grating can affect the microstructure of element correspondences.

In Experiment V, lines a and c were equidistant to line b'. When line b' was attended, primarily the offset of line a, and not the offset of line c, is reported (Figure 6B). In Experiment II, when line c' was attended, correspondence is higher when line a was offset than when line b was offset, although the distance between a and c' was larger than the distance between b and c' (Figure 3B). These findings are in violation of the “nearest neighbor principle” (e.g., Ullman, 1979) and suggest the importance of the bounding contours for motion correspondences. In Experiment V, the left bounding contour in the first frame (a) is mapped to the left bounding contour of the second frame, while the right bounding contour of the first frame (c) is mapped either to the global right bounding contour of the second frame (e') or to the left bounding contour of the group d'–e'. In Experiment II, the correspondence for line c' was neither with the line c in the same position nor its nearest neighbors b or d but, instead, with the left bounding contour a (see Figure 3).

Previous research showed that element correspondences in Ternus–Pikler displays depend on both temporal (e.g., Pante & Picciano, 1976) and form factors (e.g., Kramer & Yantis, 1997). These findings lead to the suggestion that perceptual grouping takes places interactively in space and time (e.g., Wallace & Scott-Samuels, 2007). We have shown that object features are perceived according to perceptual grouping relations, even when these relations imply the violation of pure retinotopic rules for feature processing. We suggest that these spatiotemporal feature attributions result from interactions between form and motion computations that take place in the synthesis of dynamic form.

From this perspective, motion correspondences and form computations are correlated, and the primacy of bounding contours emerges from the central role that the bounding contours play in the computations of both motion and form. In the standard Ternus–Pikler display (Figure 1), no motion is perceived if the leftmost line in the first frame and the rightmost line in the second frame are removed. Accordingly, we found previously that feature attribution in this case is determined by retinotopic coordinates (Ögmen et al., 2006). Moreover, motion cues are basically given by the outer lines (see also, Watanabe & Cole, 1995), and hence the correspondences between the outer lines may be less ambiguous than those between the inner lines.

From the perspective of form computation, the importance of boundary contours has long been recognized as a key factor in figural organization (e.g., Grossberg & Mingolla, 1985; Koffka, 1935; Mitchison & McKee, 1985; Tse & Caplovitz, 2006; Tse & Logothetis, 2002). Hence, the primacy of bounding contours finds its roots both in motion and form processing: the motion induced by the outer lines is that of group motion; that is, the elements bounded by the outer lines are organized into a single object, whose form is that of a line grating. Inspection of correspondence diagrams obtained in all of our experiments shows that correspondences between bounding contours are always strong. On the other hand, the strengths of other correspondences are more variable and depend on the spatiotemporal configurations of the stimuli.

**Summary**

In this contribution, we introduced non-retinotopic feature attribution as a tool to unravel the perceptual organization of individual elements in motion displays.
We applied this new measure to transformational group motion for which the exact element-to-element correspondences are subjectively less obvious. Our results show that the figural transformations of moving line gratings affect the microstructure of line-to-line correspondences. We interpret this finding according to the principle of the “primacy of bounding contours”; that is, bounding contours of an object provide a framework for motion correspondences, which is more important than the internal structure of that object.

Acknowledgments

T.O. was supported by the Swiss National Science Foundation. We thank Marc Repnow for excellent technical support, Kristoffer Åberg for discussions, and Daniela Högl and Maria Casado-Hernandez for help with data collection.

Commercial relationships: none.
Corresponding author: Thomas U. Otto.
Email: tom.otto@epfl.ch.
Address: EPFL SV INS LPSY, Station 15, 1015 Lausanne, Switzerland.

References


