

Cognition and Neurosciences

Common mechanisms in apparent motion perception and visual pattern matching

AXEL LARSEN and CLAUS BUNDESEN

University of Copenhagen, Denmark

Larsen, A. & Bundesen, C. (2009). Common mechanisms in apparent motion perception and visual pattern matching. *Scandinavian Journal of Psychology*, 50, 526–534.

There are close functional similarities between apparent motion perception and visual pattern matching. In particular, striking functional similarities have been demonstrated between perception of rigid objects in apparent motion and purely mental transformations of visual size and orientation used in comparisons of objects with respect to shape but regardless of size and orientation. In both cases, psychophysical data suggest that differences in visual size are resolved as differences in depth, such that transformation of size is done by translation in depth. Furthermore, the process of perceived or imagined translation appears to be stepwise additive such that a translation over a long distance consists of a sequence of smaller translations, the durations of these steps being additive. Both perceived and imagined rotation also appear to be stepwise additive, and combined transformations of size and orientation appear to be done by alternation of small steps of pure translation and small steps of pure rotation. The functional similarities suggest that common mechanisms underlie perception of apparent motion and purely mental transformations. In line with this suggestion, functional brain imaging has isolated neural structures in motion area MT used in mental transformation of size.

Key words: Apparent motion, mental rotation, mental size transformation, visual matching, pattern matching.

Axel Larsen, Center for Visual Cognition, Department of Psychology, University of Copenhagen, Øster Farimagsgade 2A, DK-1361 Copenhagen K, Denmark. E-mail: Larsen.Axel@psy.ku.dk

When a stationary object is presented first in one position and, after an appropriate time interval, again in a second position, the human observer has the illusory impression of a single object moving smoothly across the field from the first position to the second one. The phenomenon is called “apparent movement” to emphasize that there is no physical movement, but merely an impression of movement. It was discovered by Exner (1875) and later studied extensively in Gestalt psychology by Wertheimer (1912), Korte (1915), Neuhaus (1930), and others. Later studies (e.g., Bundesen, Larsen & Farrell, 1983; Shepard & Judd, 1976) have shown that when the two objects differ in orientation, observers get a visual impression of one rigid object in smooth rotational motion, and when the objects differ in size, one rigid object is seen to recede or approach the observer, while keeping distal size constant.

The impression of motion in these paradigms depends critically on temporal as well as spatial parameters (distance, angular difference, and size ratio). In the limiting case where stimulus onsets and offsets are identical (i.e., with simultaneous presentation) and only spatial parameters are varied, no motion is perceived. However, if the task is now to determine whether the two objects are identical except for differences in size or orientation, then subjects seem to solve the task by use of visual mechanisms that are partly the same as the mechanisms underlying perception of apparent motion. In this article we first review motion studies on translated, size transformed, and rotated patterns, and next document a striking functional relationship between motion perception and elementary visual recognition of patterns that differ in position, size, or orientation.

In addition, we point to converging evidence from brain imaging research, which suggests that brain systems allocated for visual motion processing also provide part of the neural underpinnings for elementary pattern recognition.

VISUAL APPARENT MOVEMENT

Apparent translatory motion and Korte's third law

The simplest type of illusory movement is seen when two laterally separated spots of light are presented in succession. In the first systematic exploration of the spatiotemporal conditions for the illusion of apparent continuous translation of a spot of light Korte (1915) concluded that, when stimulus intensity and duration are held constant, the inter-stimulus interval (ISI) required for apparent continuous movement (“optimal apparent movement” or “beta movement”) is directly related to the spatial separation between the stimuli (Korte's third law). Korte used the time from stimulus offset to stimulus onset – the interstimulus interval (ISI) – as the main temporal parameter, but later investigators usually followed Kahneman (1967) and used the time between stimulus onsets, the stimulus-onset asynchrony (SOA).

Korte's third law has not stood the test of time. Further studies (e.g., Neuhaus, 1930; Caelli & Finlay, 1981) have shown that if a given ISI provides good movement (beta movement) over a given spatial separation, it provides beta movement over any smaller separation too. Nevertheless, the law does appear to

embody an important principle. The cited studies suggest that when stimulus intensity and duration are kept constant, the minimum SOA required for apparent movement (the simultaneity threshold for apparent continuous translation) increases with the spatial separation.

To test the generality of this relationship, Larsen, Farrell, and Bundesen (1983) presented subjects with two point sources in sequential alternation (with zero ISI and zero intercycle interval) and measured the minimum SOA required for apparent continuous translation as a function of the (lateral) separation of the sources with separations ranging from about one-tenth of a degree up to 5 degrees of visual angle. Viewing distance was varied as a parameter.

With viewing distance kept constant, the minimum SOA required for apparent continuous translation increased with the angular separation of the sources over the entire range of visual angles investigated. However, as illustrated in Fig. 1 (upper panel), the function showed a strong linear increase over angles

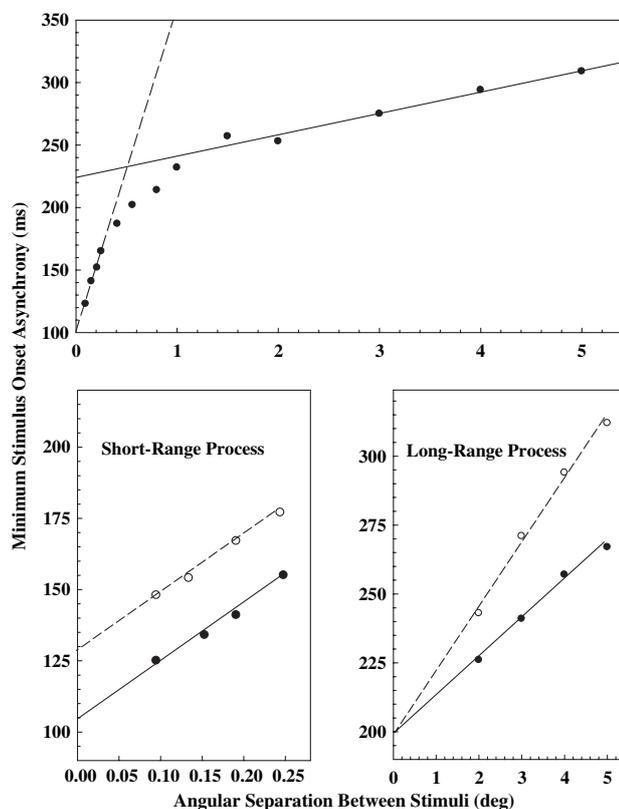


Fig. 1. Upper panel: Minimum stimulus-onset asynchrony required for perception of apparent movement as a function of visual angle separating the stimuli (redrawn after Larsen, Farrell & Bundesen, 1983). Separate least-squares straight lines have been fitted to the data for small ($\leq 0.25^\circ$) and large ($\geq 1.25^\circ$) visual angles. Lower panels: Minimum stimulus-onset asynchrony required for perception of apparent movement as a function of visual angle separating the stimuli. Viewing distance was 1.5 m (closed circles) and 6 m (open circles). Left panel: Short-range results fitted by straight parallel lines. Right panel: Long-range results fitted by diverging straight lines with a common zero intercept.

less than about a quarter of a degree (short-range component of the function) and a much weaker linear increase over angles greater than about one degree (long-range component). The short- and long-range components of the function were differentially affected by a fourfold increase in viewing distance (see Fig. 1, lower panels): The long-range component increased in slope, but the extrapolated zero-intercept was constant; the short-range component changed in intercept, but not in slope.

Our results supported a two-process theory for apparent movement proposed by Braddick (1974). In this theory, visual impressions of movement may result from either of two processes, a central process that underlies apparent movement over large visual angles ("long-range apparent movement") and a more peripheral process that produces impressions of movement over small visual angles ("short-range apparent movement" or "real movement"). Long-range movement is thus supposed to be indirectly inferred by a high-level process. For motion stimuli in the short or in the long range, then, a modern version of Korte's third law may be formulated as follows: the minimum SOA required for apparent movement (the simultaneity threshold for apparent continuous translation) increases as a linear function of the spatial separation between the stimuli. A simple interpretation of Korte's third law is that there is an upper limit to the speed of apparent translatory movement in visual space.

Further studies by Sperling and co-workers (e.g., Lu & Sperling, 2001), among others, have led to an extension of the two-process theory. In the extended theory, there are three systems of motion perception. The first-order system that responds to moving luminance patterns corresponds to the short-range process. A new second-order system responds to moving modulations of feature types. The first- and second-order systems are predominantly monocular, sensitive, and fast. The third order system embodies the long-range mechanism. This system is binocular, rather slow, and to some extent influenced by attention, and it generates impressions of movement from a broad range of stimuli.

Representation of illusions of movement in primary visual cortex

Following up on research based on single cell recordings in primates, which had revealed a brain area (called MT or V5) particularly sensitive to moving stimuli, several research groups (e.g., Tootell, Reppas, Kwong *et al.*, 1995; Watson, Myers, Frackowiak *et al.*, 1993) used the new functional brain imaging techniques (PET, fMRI) to map a corresponding motion area in humans. Area MT is located approximately at the junction of occipital, parietal and temporal cortex. It is highly sensitive to moving stimuli and it also responds when objects do not move, but just appear to do so. However, MT is also highly active when stimuli follow so slowly after one another that there is no impression of movement (see Larsen, Kyllingsbæk, Law & Bundesen, 2005). Activation of area MT may thus be necessary, but is clearly not sufficient for our experience of objects moving across regions of visual space that are not directly stimulated.

To cast light on this issue we conducted an fMRI experiment (Larsen, Madsen, Lund & Bundesen, 2006) in which we monitored activity in the primary visual cortex (V1). We presented observers with two widely separated flashing dots. In one condition onsets and offsets were simultaneous, and in the (long-range) apparent motion condition one dot went on as the other went off. In a third condition, we presented a flashing bar that connected the dots, and the fourth real motion condition just displayed one moving dot that traversed the space between the locations of the flashing dots.

Primary visual cortex is retinotopically organized, and the flashing dots mapped to fairly widely separated regions of activation in V1 (see Fig. 2a), while the flashing bar (Fig. 2b) and the dot in real motion (Fig. 2c) yielded a trace of activation connecting the regions representing the flashing dots. Moreover, when the flashing dots appeared in sequential alternation and generated illusions of translatory movement this was accompanied by a corresponding trace of activation in the primary visual cortex (Fig. 2d).

Several types of evidence support the conjecture that long-range apparent motion is generated by interaction between MT and V1. Transcranial magnetic stimulation (TMS) of MT can

induce visual sensations of moving phosphenes (spots of light), but the sensations of motion fail to appear if a “blocking” TMS pulse is delivered over V1 some 30 ms later (Pascual-Leone & Walsh, 2001). This latency closely corresponds to the time needed for communicating a signal from MT back to V1.

Extensions of Korte’s third law: translations in depth

In an unpublished study with 18 subjects tested in a fully lit classroom with one shielded computer available for each subject (Larsen, 2003), we presented two sequentially alternating stimulus patterns that were the same in shape, but differed in size in the frontoparallel picture plane (see Fig. 3, left panel). The experiment was self-paced and each trial displayed a new pair of random same-shaped stimuli. When the rate of alternation was sufficiently slow, the presentation produced a visual impression of a rigid object moving in three-dimensional space in which the object appeared to be moving to and fro in depth rather than changing in distal size. ISI was zero, and the subject first pressed the left-hand button on the mouse a number of times, each time decreasing SOA by 5%, until the impression of rigid motion broke down (SOA1). The subject next pressed

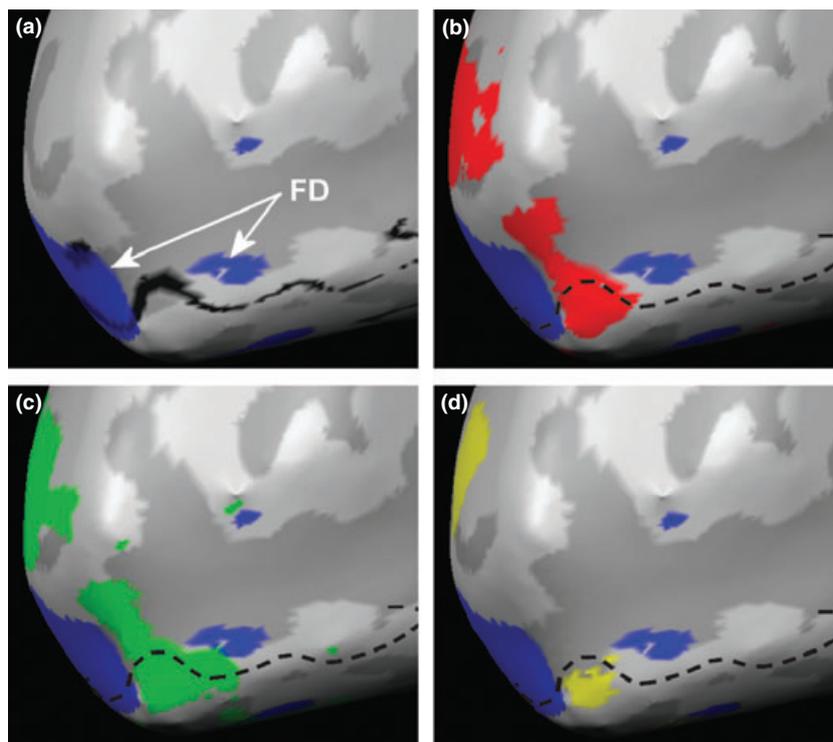


Fig. 2. Cortical activations in 31-year-old male observer. The activations appear as colored patches on the inflated left occipital hemisphere. Gyri are indicated by light gray and sulci by darker gray. (a) BOLD response (blue) due to flashing dots (FD) relative to a baseline formed by the BOLD response in the fixation-point condition. The computed location of the representation of the horizontal meridian in the calcarine sulcus is shown in black. The computation was based on the mean activation caused by a rotating wedge passing through the horizontal orientation (mean of clockwise and counterclockwise rotations). (b) Contrast (red) between BOLD response to flashing bar and BOLD response to flashing dots along with the (blue) BOLD response to the flashing dots. The approximate location of the representation of the horizontal meridian in the calcarine sulcus is indicated by the dashed black line. (c) Contrast (green) between BOLD response to real motion and BOLD response to flashing dots along with the (blue) BOLD response to the flashing dots. (d) Contrast (yellow) between BOLD response to apparent motion and BOLD response to flashing dots along with the (blue) BOLD response to the flashing dots. For all activations, $p < 0.05$, uncorrected. (Reprinted from Larsen, Madsen, Lund & Bundesen, 2006).

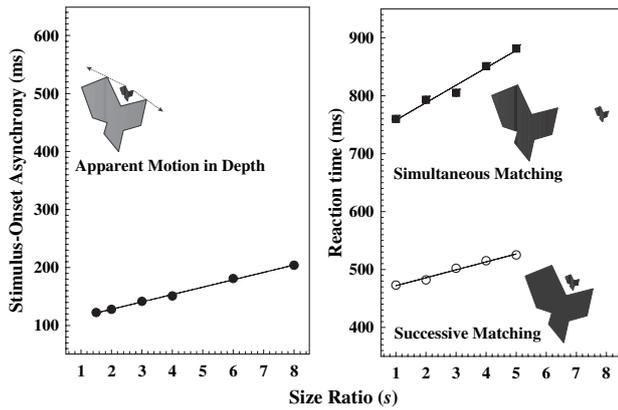


Fig. 3. Left panel: Minimum SOA as a function of size ratio (s). Right panel, upper curve: Mean reaction time across positive and negative stimulus pairs as a function of size ratio in simultaneous matching (redrawn from Bundesen & Larsen, 1975, Experiment 2). Right panel lower curve: Mean reaction time across positive and negative stimulus pairs in successive matching (redrawn from Larsen & Bundesen, 1978, Experiment 1). Examples of the type of stimulus material used in the experiments are shown in the figures.

the right-hand button a number of times, each time increasing SOA by 5%, until good rigid motion reappeared (SOA2), which terminated the trial. The left panel of Fig. 3 shows mean SOA (i.e., the mean of SOA1 and SOA2) across subjects as a function of size ratio. The linear fit accounts for 99.99% of the variance.

The subjective reports suggested that changes in stimulus size were visually resolved as changes in apparent distance such that, within trials, the apparent distal size of the moving object was approximately constant. To analyze the situation geometrically, let x and y be alternating stimulus patterns presented in a frontoparallel plane at a distance D from the subject (Fig. 4, lower panel). Let L_x and L_y , where $L_x < L_y$, be the linear sizes (say, the greatest linear extents) of x and y , and let s be the size ratio L_y/L_x . Complete visual resolution of the size disparity between x and y as a difference in depth should correspond to central projection of the patterns onto frontoparallel planes at distances d_x and d_y , respectively, such that apparent distal sizes l_x and l_y are equalized. It is easy to show geometrically that $l_x = l_y$ if, and only if,

$$d_x = s d_y \tag{1}$$

so that the length of the motion trajectory, d , is given by

$$d = d_x - d_y = s d_y - d_y = d_y(s - 1) \tag{2}$$

which is proportional to $s - 1$ on the assumption that in the fully lit room, the apparent visual distance d_y to the largest pattern on the screen was approximately constant. The nearly perfect linear fit to mean SOA as a function of the size ratio s (Fig. 3, left panel) strongly supports the idea that differences in size are first resolved as differences in depth, followed by a computation of the motion path, which is then implemented as a time consuming movement in depth. The illusion of movement is constrained by a maximum upper velocity and the average

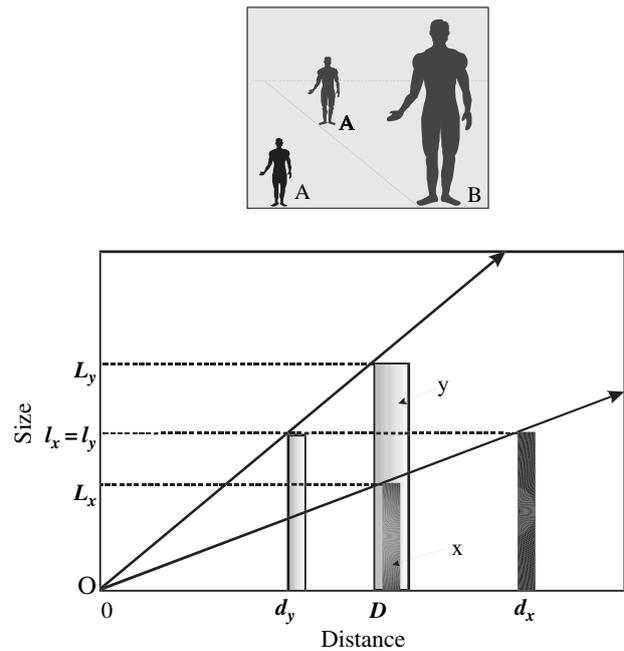


Fig. 4. Upper panel: Graphical illustration of how discrepancies of size in the left and rightmost silhouettes may be interpreted as depicting same-sized silhouettes at difference distances. Lower panel: Geometric resolution of disparity of size as a difference in depth. Objects x (smaller dark rectangle) and y (larger light rectangle) with linear sizes L_x and L_y are shown at a distance D from a subject. By geometric multiplications with respect to the subject's point of view (the origin of the coordinate system), the disparity of size between x and y is resolved as a difference in depth between the images of x (larger dark rectangle) and y (smaller light rectangle). The sizes and distances of the images are designated by l_x , l_y and d_x , d_y .

time (102 ms) needed for resolving size discrepancies and computing motion trajectories (which connect corresponding parts of the alternating patterns, see Ullman, 1979) corresponds to the intercept for $s = 1$.

Illusion of rotation of three-dimensional objects

In their landmark investigation of mental rotation Shepard and Metzler (1971) used perspective views of the same three-dimensional object displayed in two different orientations. Axis of rotation was either the line of sight (rotation in the picture plane) or a vertical line (rotation in depth). In either case, Shepard and Metzler found that the time taken to decide whether the perspective views portrayed the same three-dimensional object was a linear function of the angular difference in orientation between the two views. The function had essentially the same slope but slightly different intercepts for rotations in the picture plane and rotations in depth. Based on phenomenological reports in which subjects claimed to mentally rotate an image of one of the stimuli into congruence with the other one by a gradual process much like a real physical rotation, Shepard and Metzler hypothesized that this was indeed how the task was solved. In a subsequent study Shepard and Judd (1976)

displayed the very same three-dimensional objects in sequential alternation (zero ISI). With sufficiently long SOA participants reported vivid impressions of a rigid three-dimensional object rotating back and forth in the picture plane or in depth. In this study, Shepard and Judd (1976) found that the SOA at which the impression of rigid rotational motion broke down increased approximately linearly as a function of angular difference in orientation with nearly the same slopes for rotations in the picture plane and in depth. The linear increase suggests that there is an upper limit to the velocity of the rotational movement. Farrell, Larsen, and Bundesen (1982) showed that the limit relates to angular velocity, and not the linear velocity of the fastest moving subpattern.

The similarity between experiencing illusory rotations of an object and mentally rotating an image of an object is suggestive. Moreover, even though slopes and intercepts differed greatly in the motion and recognition tasks, the functional dependencies of reaction time and stimulus-onset asynchrony on angular difference in orientation show strong similarities between mental rotation and apparent rotational movement.

Translation and rotation combined

Bundesen, Larsen and Farrell (1983) presented subjects with two sequentially alternating stimulus patterns that were the same in shape, but differed in size and angular orientation in the frontoparallel picture plane. When the rate of alternation was sufficiently slow, the presentation produced a visual impression of a rigid object moving in three-dimensional space. If stimulus size was kept constant, the object appeared to be rotating back and forth in a frontoparallel plane. When stimulus size was changing, the object appeared to be moving to and fro in depth rather than changing in distal size. Apparent changes in distance and orientation seemed to be gradual and concurrent in the sense that, when the object appeared at intermediate distances, its apparent angular orientation was intermediate between those of the two stimuli. The experience could be described as the impression of a "helical, screwlike motion in depth".

The experiment was run in a completely darkened room and here a plausible constraint on the visual resolution of the disparity of size as a difference in depth was that the average apparent distance of that object, $(d_x + d_y)/2$, was a constant a regardless of L_x and L_y . The constant average apparent distance a should be determined by the constant indications of distance from accommodation, convergence, and binocular disparity; the deviation from a represented by $d_x - d_y$ should be a function of the sole cue to such a difference in depth, the disparity of size between x and y .

By use of Equation 1, it is easy to verify that

$$d_x - d_y = (d_x + d_y)(s - 1)/(s + 1) \quad (3)$$

Equation 3 implies that, if the average apparent distance, $(d_x + d_y)/2$, is a constant, then the difference in apparent depth

between the stimulus patterns x and y is directly proportional to the ratio $(s - 1)/(s + 1)$.

Our results are illustrated in Fig. 5 which shows the SOA threshold for the appearance of a rigidly moving object as a function of the ratio $(s - 1)/(s + 1)$ with the angular extent of the transformation ν as the parameter. The SOA threshold was approximately a linearly increasing function of ν for simple rotation,

$$SOA_{rot} = \alpha + \beta\nu \quad (4)$$

and a linearly increasing function of the ratio $(s - 1)/(s + 1)$ for simple translation in depth,

$$SOA_{trans} = \gamma + \delta(s - 1)/(s + 1) \quad (5)$$

and the extrapolated zero-intercept α of the threshold function for apparent rotation was higher than the extrapolated zero-intercept γ of the threshold function for apparent translation in depth. For composite transformations, the SOA threshold was fitted by

$$SOA_{comp} = \begin{cases} SOA_{trans} + \beta\nu & \text{if } SOA_{trans} \geq \alpha \\ \alpha + \beta\nu & \text{if } SOA_{trans} < \alpha \end{cases} \quad (6)$$

Equation 6 fits in with the hypothesis that rigid movements are decomposable into translations and rotations, and a translation

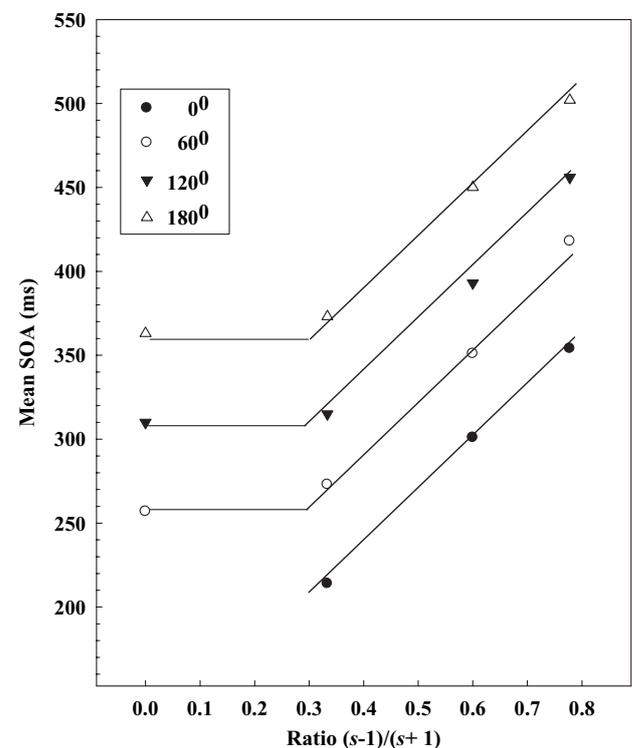


Fig. 5. Mean of minimum SOAs as a function of $(s - 1)/(s + 1)$, where s is the linear size ratio between stimuli, with angular difference in orientation as parameter. Solid straight lines indicate the fits by Equation 6.

over a long distance is decomposable into smaller translations, whereas a rotation over a wide angle is decomposable into smaller rotations. A simple, general procedure for generating a sequence of visual representations of an object in rigid movement might consist in iterating one routine that transforms a visual representation of an object into a visual representation of the same object after an incremental translation (a translation over a certain small distance) and another routine that transforms a visual representation of an object into a visual representation of the same object after an incremental rotation (a rotation over a certain small angle). Assume that long-range apparent movement of rigid objects is *stepwise additive* in the sense that it is implected in successive small steps such that each step is an incremental translation or an incremental rotation. The main features of our parametric data on SOA thresholds may then be explained by assuming that the minimum time required to perform an incremental transformation (translation or rotation) is a constant regardless of the sequence of transformations in which it is embedded.

In the suggested interpretation, the minimum implection time for apparent translation is directly proportional to the linear extent of the transformation; the minimum implection time for apparent rotation is directly proportional to the angular extent of the transformation; and the implection time for a composite transformation of apparent translation and rotation equals the sum of the implection times for the components. Thus, in Equations 4 and 5, slope constants β and δ are measures for the maximum speed of implecting apparent rotation and apparent translation, respectively. The zero-intercepts α and γ of the threshold functions measure the time taken to establish correspondence and specify the trajectory to be implected. Intercept α for simple rotation was higher than intercept γ for simple translation, and Equation 6 means that, in composite transformations too, implection of apparent translation was released with the shorter latency γ whereas implection of apparent rotation was released with the longer latency α .

VISUAL PATTERN MATCHING

The way we determine that two objects are identical with respect to shape regardless of their size or orientation was first discussed by Mach (1902) and described by many later researchers (e.g. Woodworth & Schlossberg, 1954) as a basic case of stimulus equivalence. Modern treatments of shape invariance began with Pitts and McCulloch's (1947) neural network model for extraction of size invariant features in the visual cortex and gained momentum in the late 1960s and early 1970s (Bundesen & Larsen, 1975; Rock, 1973; Shepard & Metzler, 1971).

Spatial separation in visual pattern matching

Whereas size and orientation invariance have been extensively investigated, the simplest case of all, the case of translational displacement across the retina, has attracted rather little atten-

tion. Cave *et al.* (1994) and Foster and Kahn (1985; see also, Kahn & Foster, 1981) reported effects of spatial separation on the detection of translational equivalence of successively presented patterns, which was explained by assuming that one pattern was encoded into visual memory and then mentally translated to fit the other pattern.

When we controlled for eye movements and eccentricity by showing the two patterns to be compared (postmasked filled polygons) simultaneously on the periphery of an imaginary circle for short exposure durations (below 194 ms), we found no evidence of mental translation (Larsen & Bundesen, 1998). Detection of equivalence was equally good regardless of the spatial separation, suggesting that the initial visual computations released by the presentation of the simultaneous stimuli sufficed to detect stimulus equivalence without any mental translations aligning the two patterns. We then changed the task so that members of a "same" pair could differ by both a translation and a rotation. Detection of equivalence (measured by d') now decreased as a function of the angular difference in orientation suggesting that the task now required a mental rotation. On the critical trials, however, the rotational component was zero. On these trials detection decreased as a function of spatial separation, suggesting that the subjects had adopted a general strategy of mental alignment of the patterns to be compared, which was maintained on the critical trials as a gradual mental translation of one pattern to the position of the other one.

Mental transformation of size

Size invariance in visual pattern recognition has been thoroughly investigated by reaction time methods (Bundesen & Larsen, 1975; Larsen & Bundesen, 1978, Experiment 2; Sekuler & Nash, 1972). In a typical experiment, the participant is instructed to determine "as quickly as possible without making too many errors" whether two objects that differ with respect to size have the same shape (see Fig. 3, right panels). The pattern of results is similar whether the objects are presented simultaneously or one at a time, that is, as one-back match-to-sample. In either case RT is found to increase approximately linearly as a function of size ratio (see Fig. 3, right panels), but usually with a larger slope and a higher intercept with simultaneous presentation of the stimuli. We still subscribe to our original hypothesis, namely, that the task is solved by (a) encoding the visual impression of one of the patterns to be compared as a mental image, (b) gradually transforming the image so that the represented size changes to the size of the other pattern, and then (c) matching the transformed visual image against the visual impression of the other pattern. We also hypothesize that mental transformation of size is realized as a mental translation in imagined three-dimensional space.

The difference between RT functions for simultaneous and successive matching is a difference in degree, not in kind. There is good evidence that the difference is due to the fact that subjects in simultaneous matching trials repeat a mental size transformation that usually can be performed only once with one-

back match-to-sample presentations (see Larsen, McIlhagga & Bundesen, 1999). The repetitions inflate both intercept and slope and these effects can be predicted by a random walk model in which elementary mental size transformations are repeated until a fixed degree of evidence for a match or a mismatch is achieved (see Larsen *et al.*, 1999).

To investigate the neural basis of mental size transformation we performed a one-back match-to-sample PET study in which one-back matches of the same size were contrasted with one-back matches in which stimulus size increased/decreased by a factor of 6 (Larsen, Bundesen, Kyllingsbæk, Paulson & Law, 2000). As shown in Fig. 6, mostly structures in the dorsal visual pathways, including areas in the occipital, parietal and temporal transition zone (for the most part in the left hemisphere), were implicated. Relating to our functional analyses, a crucial finding was that the motion area, MT, was activated during mental transformation of size.

Mental rotation

The original paper by Shepard and Metzler (1971) on mental rotation is one of the most cited papers in cognitive psychology. We have already described their basic findings and the hypothesis that the task is performed by a “holistic” mental rotation corresponding to an actual physical rotation, with a fixed maximum angular velocity, by which objects are brought into align-

ment. The hypothesis has been much debated from a purely theoretical viewpoint (see recent discussions on “imagery” in Pylyshyn, 2003; Kosslyn, Thompson & Ganis, 2006), and the hypothesis may seem difficult to reconcile with eye movement recordings (Just & Carpenter, 1976) and with the wide variation in slope constants (about 60 deg/s reported by Shepard & Metzler, 1971; about 600 deg/s reported by Bundesen *et al.*, 1981) resulting from minor modifications of the task. However, a fairly straightforward account for these discrepancies is that subjects mentally rotate subpatterns one by one, in particular when patterns are complex and exceed visual short-term memory capacity or when positive and negative pairs are highly similar. We thus conjecture that slope differences in RT functions for mental rotation (see, e.g., Folk & Luce, 1987) may be explained like slope differences in mental size transformation studies, that is, by variations in the number of times the subjects repeat the mental transformation to increase the probability of a correct response.

The neurobiological basis of mental rotation has been extensively investigated. Overall, the activation resembles the activation due to mental transformation of size, with significant activation in the visual dorsal pathways (posterior parietal cortex, occipital cortex). Also, several investigators have found activation in area MT (e.g., Alivisatos & Petrides, 1997; Cohen, Kosslyn, Breiter *et al.*, 1996; Vanrie, Béatse, Wagemans, Sunaert & van Hecke, 2002).

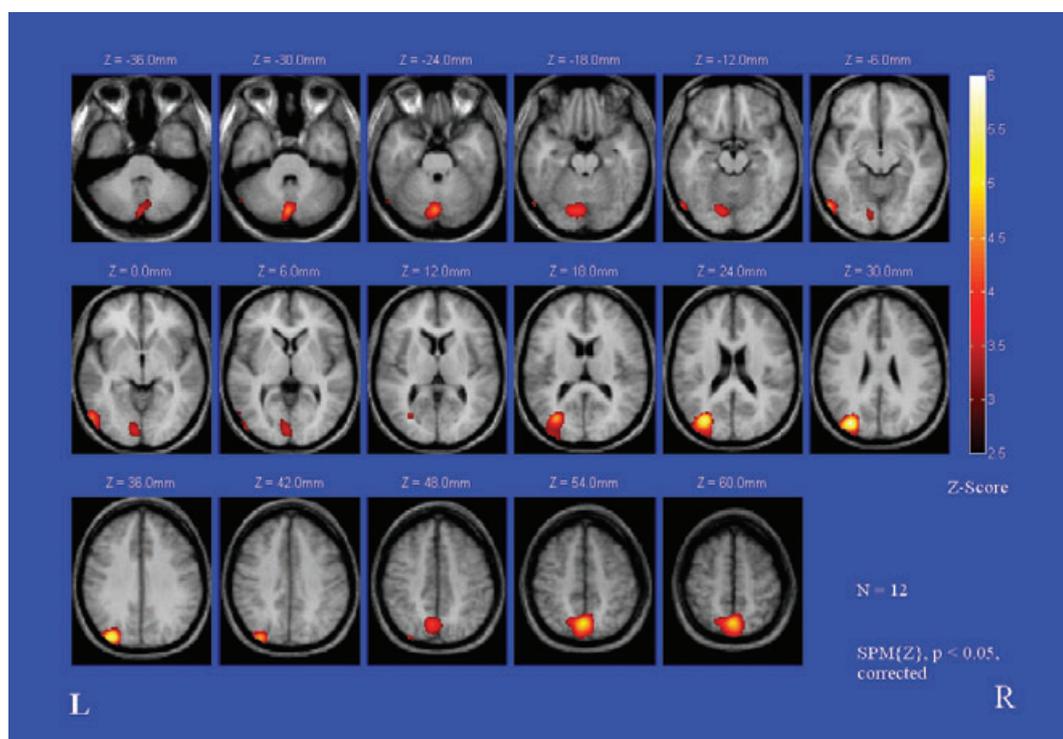


Fig. 6. Transverse sections of the statistical parametric map of the contrast between blocks of one-back match-to-sample stimuli changing size format in the sequence 1:6:1:6:1 ... and blocks of trials in which size format was always 6 or always 1. The contrast is rendered on the average image of the normalized MR brain images of the 12 subjects. Distances above/below the AC-PC plane are given in millimeters. The Z-scores are mapped to a hot metal color scale. The activation in the vicinity of MT is in the left hemisphere and is strongest at a Z-value close to zero.

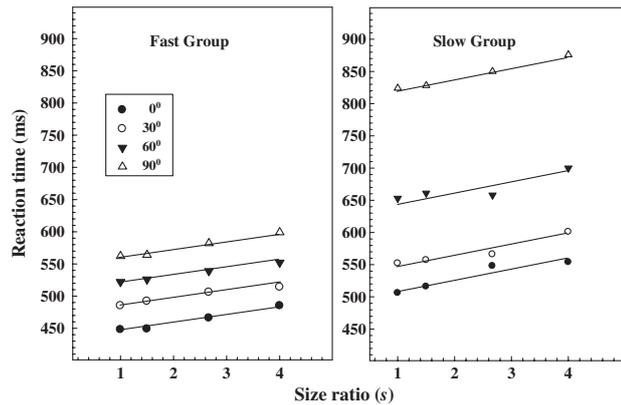


Fig. 7. Mean RT as functions of linear size ratio (s) with angular difference between stimulus pairs as the parameter (redrawn after Bundesen, Larsen & Farrell, 1981). Left panel shows group data for fast subjects fitted with equidistant lines. Right panel shows group data for slow subjects fitted by parallel lines.

Size transformation and rotation combined

To gain further insight into mental transformations of size and orientation we constructed a successive matching task in which stimulus patterns differed with respect to size, orientation, or both (Bundesen *et al.*, 1981). The effects of size ratio and angular difference in orientation were additive and highly similar to the results for visual motion (see Fig. 7). In a later study, Larsen (1985) corroborated and extended these findings. Thus, there are strong functional similarities between the mechanisms underlying perception of apparent motion and the mechanisms underlying transformation of mental images. In exactly the same way as movements of rigid objects appearing to revolve while moving to and fro are decomposable into translations and rotations, we suggest that mental transformations of size and orientation are decomposable into imagined translations and rotations. And, again, that an extended mental translation is decomposable into smaller translations whereas an extended mental rotation is decomposable into smaller rotations.

CONCLUDING DISCUSSION

There are close functional similarities between perception of apparent motion and visual pattern matching. In particular, striking functional similarities have been found between perception of rigid objects in long-range (“high-level”) apparent motion and purely mental transformations of visual size and orientation used in comparisons of objects with respect to shape but regardless of size and orientation. In both cases, psychophysical data suggest that differences in visual size are resolved as differences in depth, such that transformation of size is done by translation in depth. Furthermore, the process of perceived or imagined translation appears to be stepwise additive such that a translation over a long distance consists of a sequence of smaller translations, the durations of these steps being additive. Both perceived and imagined rotation also appear to be stepwise

additive, and combined transformations of size and orientation appear to be done by alternation of small steps of pure translation and small steps of pure rotation. The functional similarities suggest that common mechanisms underlie perception of apparent motion and purely mental transformations. In line with this suggestion, functional brain imaging has isolated neural structures in motion area MT used in mental transformation of size. Mental imagery seems to be a kind of simulated perception, not only psychologically but also neurophysiologically.

This research was supported by grants from University of Copenhagen’s Center of Excellence Program to the Center for Visual Cognition and by grants from the Nordic Research Councils to the Nordic Center of Excellence in Cognitive Control.

REFERENCES

- Alivisatos, B. & Petrides, M. (1997). Functional activation of the human brain during mental rotation. *Neuropsychologia*, *35*, 111–118.
- Braddick, O. (1974). A short-range process in apparent motion. *Vision Research*, *14*, 519–527.
- Bundesden, C. & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 214–220.
- Bundesden, C., Larsen, A. & Farrell, J. E. (1981). Mental transformations of size and orientation. In J. Long & A. Baddeley (Eds.), *Attention and performance IX* (pp. 279–294). Hillsdale, NJ: Erlbaum.
- Bundesden, C., Larsen, A. & Farrell, J. E. (1983). Visual apparent movement: Transformations of size and orientation. *Perception*, *12*, 549–558.
- Caelli, T. & Finlay, D. (1981). Intensity, spatial frequency, and temporal frequency determinants of apparent motion: Korte revisited. *Perception*, *10*, 183–189.
- Cave, K. R., Pinker, S., Giorgi, L., Thomas, C. E., Heller, L. M., Wolfe, J. M. & Lin, H. (1994). The representation of location in visual images. *Cognitive Psychology*, *26*, 1–32.
- Cohen, M. S., Kosslyn, S. M., Breiter, H. C., DiGirolamo, G. J., Thompson, W. L., Anderson, A. K., Bookheimer, S. Y., Rosen, B. J. & Belliveau, J. (1996). Changes in cortical activity during mental rotation: A mapping study using functional MRI. *Brain*, *119*, 89–100.
- Exner, S. (1875). Experimentelle Untersuchung der einfachsten psychischen Prozesse: IV. Abhandlung, Die Empfindungszonen des Sehnervenapparats. *Archiv für die gesamte Physiologie des Menschen und der Tiere*, *11*, 581–602.
- Farrell, J. E., Larsen, A. & Bundesden, C. (1982). Velocity constraints on apparent rotational movement. *Perception*, *11*, 541–546.
- Folk, M. D. & Luce, R. D. (1987). Effects of stimulus complexity on mental rotation rate of polygons. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 395–404.
- Foster, D. H. & Kahn, J. I. (1985). Internal representations and operations in the visual comparison of transformed patterns: Effects of pattern point-inversion, positional symmetry, and separation. *Biological Cybernetics*, *51*, 305–312.
- Just, M. A. & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441–480.
- Kahn, J. I. & Foster, D. H. (1981). Visual comparison of rotated and reflected random-dot patterns as a function of their positional symmetry and separation in the field. *Quarterly Journal of Experimental Psychology*, *33A*, 155–166.
- Kahneman, D. (1967). An onset-onset law for one case of apparent motion and metacontrast. *Perception and Psychophysics*, *2*, 577–584.

- Korte, A. (1915). Kinematoskopische Untersuchungen. *Zeitschrift für Psychologie*, 72, 193–296.
- Kosslyn, S. M., Thompson, W. L. & Ganis, G. (2006). *The case for mental imagery*. New York: Oxford University Press.
- Larsen, A. (1985). Pattern matching: Effects of size ratio, angular difference in orientation, and familiarity. *Perception & Psychophysics*, 38, 63–68.
- Larsen, A. (2003). Resolving differences in size as apparent motion in depth. Unpublished manuscript.
- Larsen, A. & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 1–20.
- Larsen, A. & Bundesen, C. (1998). Effects of spatial separation in visual pattern matching: Evidence on the role of mental translation. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 719–731.
- Larsen, A., Bundesen, C., Kyllingsbæk, S., Paulson, O. B. & Law, I. (2000). Brain activation during mental transformation of size. *Journal of Cognitive Neuroscience*, 12, 763–774.
- Larsen, A., Farrell, J. E. & Bundesen, C. (1983). Short- and long-range processes in visual apparent movement. *Psychological Research*, 45, 11–18.
- Larsen, A., Kyllingsbæk, S., Law, I. & Bundesen, C. (2005). Activation in the MT-complex during visual perception of apparent motion and temporal succession. *Neuropsychologia*, 43, 1060–1071.
- Larsen, A., Madsen, K. H., Lund, T. E. & Bundesen, C. (2006). Images of illusory motion in primary visual cortex. *Journal of Cognitive Neuroscience*, 18, 1174–1180.
- Larsen, A., McIlhagga, W. & Bundesen, C. (1999). Visual pattern matching: Effects of size ratio, complexity, and similarity in simultaneous and successive matching. *Psychological Research*, 62, 280–288.
- Lu, Z.-L. & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America, A*, 18, 2331–2370.
- Mach, E. (1902). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen* (3rd edn). Jena, Germany: Gustav Fischer.
- Neuhaus, W. (1930). Experimentelle Untersuchung der Scheinbewegung. *Archiv für die Gesamte Psychologie*, 75, 315–455.
- Pascual-Leone, A. & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, 292, 510–512.
- Pitts, W. & McCulloch, W. S. (1947). How we know universals: The perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 9, 127–147.
- Pylyshyn, Z. (2003). Return of the mental image: “Are there really pictures in the brain?”. *Trends in Cognitive Sciences*, 7, 13–118.
- Rock, I. (1973). *Orientation and form*. New York: Academic Press.
- Sekuler, R. & Nash, D. (1972). Speed of size scaling in human vision. *Psychonomic Science*, 27, 93–94.
- Shepard, R. N. & Cooper, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- Shepard, R. N. & Judd, S. (1976). Perceptual illusion of rotation of three-dimensional objects. *Science*, 191, 952–952.
- Shepard, R. N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701–703.
- Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R. & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, 15, 3215–3230.
- Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- Vanrie, J., Béatse, E., Wagemans, J., Sinaert, S. & van Hecke, P. (2002). Mental rotation versus invariant features in object perception from different viewpoints: An fMRI study. *Neuropsychologia*, 40, 917–930.
- Watson, J. D. G., Myers, R., Frackowiak, R. S. J., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., Shipp, S. & Zeki, S. (1993). Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, 3, 79–94.
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, 61, 161–265.
- Woodworth, R. S. & Schlossberg, H. (1954). *Experimental psychology*. London: Methuen.

Received 4 August 2009, accepted 7 August 2009