Research Report

ONE-SHOT VIEW INVARIANCE IN A MOVING WORLD

Zoe Kourtzi and Maggie Shiffrar

Rutgers University

Abstract—*How do people recognize an object in a novel orientation? Psychophysical and neurophysiological studies have suggested that extensive practice is required before observers can recognize an object that has been rotated to a new orientation. Because object orientation frequently varies with object movement, we examined whether observers might more readily recognize a moving object in a new orientation. Results from a priming study indicate that motion significantly and readily enhances the recognition of new object orientations when those orientations fall within the path of the motion. That is, motion promotes view-invariant object recognition without practice.*

Successful interaction with objects in the environment requires the accurate recognition of objects across changes in their orientation. Two general classes of theories attempt to explain how observers recognize objects in novel orientations. *View-dependent* theories propose that objects are represented in specific orientations relative to the observer. View-dependent representations are suggested by a slowing of object recognition with increases in the angular difference between novel and previously seen object orientations (Jolicoeur 1985, 1988, 1990, Jolicoeur & Landau, 1984, Rock, DiVita, & Barbeito, 1981, R. N. Shepard & Cooper, 1982, R. N. Shepard & Metzler, 1971, S. Shepard & Metzler, 1988, Tarr & Pinker, 1989). *View-invariant* theories suggest that objects are represented as structural descriptions that are independent of orientation (Marr & Nishihara, 1978).

Hummel and Biederman (1992) specified the conditions under which view invariance can be achieved. Specifically, the object must be decomposable into view-invariant parts (geons), it must have a distinct geon structural description, and its geon structure should not change with any object transformation. When these conditions are satisfied, view invariance is achieved in object-naming and -matching tasks (Biederman & Gerhardstein, 1993).

One violation of the conditions required for view invariance occurs when the observer is discriminating between objects that are from the same category and share the same geon structure (Hummel & Biederman, 1992). Numerous studies suggest that view-invariant discrimination of such objects emerges only after extensive exposure to multiple object orientations (Bulthoff & Edelman, 1992, Logothetis, Pauls, Bulthoff, & Poggio, 1994, Logothetis, Pauls, & Poggio, 1995, Tarr & Pinker, 1989). For example, discrimination between different faces, cars, or birds presented in unfamiliar orientations is a difficult task that requires extensive practice if not expertise.

New orientations of objects in the same category may be recognized by an alignment between the new and old orientations (Koriat & Norman, 1988, 1989, Palmer, Rosch, & Chase, 1981, R. N. Shepard & Cooper, 1982, R. N. Shepard & Metzler, 1971, Tarr & Pinker, 1989) or by interpolation between previously stored two-dimensional views (Bulthoff & Edelman, 1992, Edelman & Weinshall, 1991, Poggio & Edelman, 1990). When either of these two processes occurs, recognition of new object orientations may be limited to a generalization field extending only up to approximately 45° from known orientations (Bulthoff & Edelman 1992). For example, monkeys trained with a set of objects can generalize only to slightly rotated versions of those objects (Logothetis et al., 1994, 1995). Moreover, many inferotemporal neurons show bell-shaped firing patterns that are centered on the training orientation of an object and fall off with changes in the object's orientation. Thus, view-invariant recognition of objects from the same category can be achieved after extensive exposure to a small number of object views.

In the laboratory, recognition performance across variations in object orientation is traditionally tested with static objects. However, outside the laboratory, changes in object orientation usually result from movement of the object or the observer. The visual system might take advantage of this association by readily linking different views of a moving object so that the object's movement appears as a single, continuous event. Such "on the fly" linkage of different orientations of moving objects should yield one-shot view-invariant recognition of dynamic objects even when those objects belong to the same category.

The goal of this research was to investigate the contribution of motion to object recognition (Ungerleider & Mishkin 1982). Specifically, we tested whether observers might achieve view-invariant object recognition more readily with moving than with static objects. To measure conservatively whether motion facilitates view-invariant object recognition, we chose conditions favoring view-dependent object representations. Novel asymmetrical objects from the same category, as shown in Figure 1, were used because they are thought to be represented in a view-dependent manner (Tarr & Pinker, 1990). Rotations in the frontal plane were employed because they violate the conditions necessary for view invariance (Hummel & Biederman, 1992). Specifically, frontal-plane rotations change an object's structural description relative to the observer by altering the *top-of, bottom-of,* and *side-of* relations between the object parts as well as the *vertical, horizontal,* and *oblique* values of the individual parts.

We used an immediate priming paradigm (Sekuler & Palmer, 1992) in which a briefly presented prime object is followed by a pair of targets. The prime consisted of a two-frame sequence of different object orientations in which the second orientation was a rotated version of the first. The rotation angle varied between 30° and 150° in 30° steps. The two prime orientations were either linked by motion (apparent motion condition) or presented as static snapshots (control condition). Subjects reported whether two subsequent targets matched each other. Priming was indicated by faster reaction times when the two targets were the same as the prime object. We were interested in whether priming of new object orientations would differ when those orientations fall within or outside the path of apparent motion. Because object representations generalize across orientation changes less than 45° (Bulthoff & Edelman, 1992, Logothetis et al., 1994, 1995), small changes in object orientation should have been primed in both conditions. Any differences in the priming across orientation between

---

Address correspondence to Zoe Kourtzi, Psychology Department, 101 Warren St., Rutgers University, Newark, NJ 07102, e-mail kourtzi@pegasus.rutgers.edu

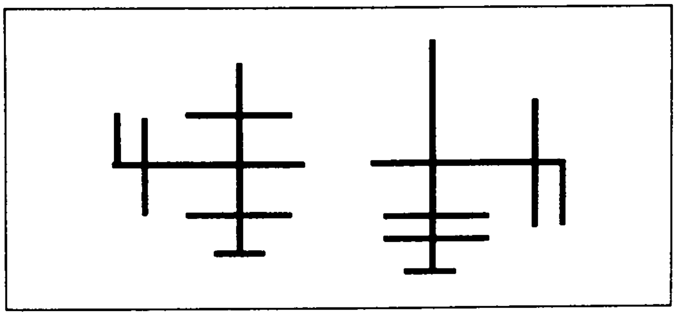

Fig 1. Example of the asymmetrical objects (Tarr & Pinker, 1990) used as stimuli The objects were similar configurations of line segments, with a line width of 3 6 min of visual angle in different spatial arrangements

moving and static objects would be identifiable for larger orientation changes

METHOD

Subjects

One hundred undergraduate students, recruited from the Rutgers subject pool, consented to participate in this experiment All subjects had normal or corrected-to-normal vision and were naive to the hypothesis under investigation

Stimuli

Stimuli were presented on a 21-in color monitor with a 1,024 × 768 pixel resolution and 60-Hz refresh rate controlled by a PowerMac 7100 The monitor was positioned 95 cm from a chin rest, and the stimuli were drawn within a square area subtending 4 82° × 4 82° of visual angle on the screen Subjects viewed the stimuli through a circular aperture to minimize framing effects from the monitor

The stimuli consisted of 40 objects adapted from the asymmetrical characters of Tarr and Pinker (1990) The set of prime objects consisted of 10 asymmetrical objects Two prime views of each object differed only by a rigid rotation of the object about its base The primes were sequentially presented in the center of the screen, the targets were simultaneously presented 0 6° of visual angle to the left and right of the center of the screen

Procedure

Figure 2 illustrates the experimental procedure Each trial began with a fixation point presented for 1,500 ms, followed by the first prime frame shown for a variable duration, as described in the following paragraph Then the second prime frame followed for the same duration as the first A blank screen was then displayed for 500 ms, followed by a pair of targets presented until the subject responded (with a 3-s maximum) Subjects carefully observed the prime objects and then pressed a key if the two subsequent targets matched each other This "go–no go" task was used to reduce the variability often observed in priming studies that require subjects to select one of two different motor responses (Biederman & Gerhardstein, 1993) Subjects were instructed that both reaction time and accuracy were important Overall feedback (mean reaction time and percentage of correct responses) was provided at the end of each block of trials

The first prime had one of five possible orientations relative to the observer 0° (upright), 90°, 180°, −45°, −135° The second prime was a rotated version of the first The two primes were separated by one of five possible rotation angles 30°, 60°, 90°, 120°, or 150° The second prime was rotated clockwise from the first prime in half the trials and counterclockwise in the rest The duration of the two prime frames varied with the rotation angle between them such that the optimal apparent motion was achieved Pilot studies showed that good apparent motion was observed when each prime frame was presented for the duration used by R N Shepard and Judd (1976) for the corresponding angle plus a constant of 100 ms This yielded durations of 232, 265, 298, 331, and 364 ms for the five rotation angles, respectively The interstimulus interval (ISI) between the two prime frames was 0 ms in the apparent motion condition and 450 ms in the control condition In the apparent motion condition, the first and second prime frames were presented so that the prime object appeared to rotate smoothly about its base In the control condition, the second prime frame was displaced 2 41° of visual angle to the right of the first This spatiotemporal separation between the two prime frames eliminated the perception of apparent motion in the control condition

Before beginning the experimental trials, each subject completed a block of 20 practice trials with objects that differed from those of the experimental trials Most subjects obtained reaction times less than 1,000 ms by the end of the practice block Subjects having longer reaction times completed a second practice block

The experimental session consisted of five blocks each containing 40 trials The target objects in each block were presented in one of five

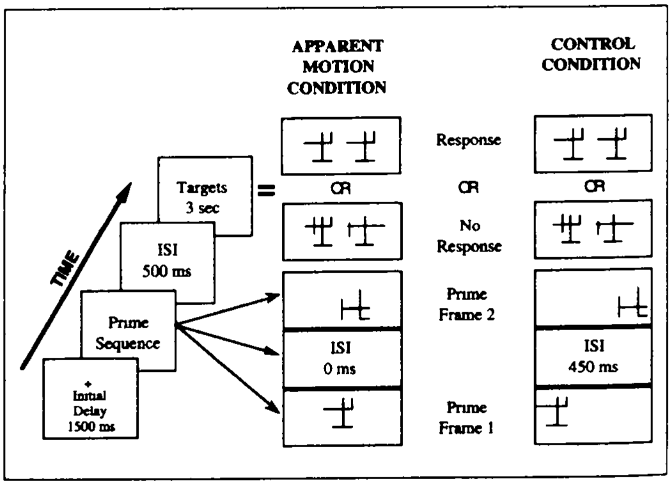

Fig 2 Experimental design for the apparent motion and control conditions A sequence of two prime views was followed by a blank screen (500 ms), and then a pair of targets appeared The subjects had to respond when the targets matched each other No response was required when the targets differed from each other The targets might be the same object as the prime or different from the prime The examples illustrated here show targets matching the prime, as well as each other, and targets matching neither the prime nor each other ISI = interstimulus interval

Table 1 *Orientation of the two targets as a function of the rotation angle of the prime*

Rotation angle of prime	Target orientations				
	Same as prime		Different from prime		
	Frame 1	Frame 2	Intermediate	Extra 1	Extra 2
30°	0	30	15	−15	45
	90	120	105	75	135
	180	−150	−165	165	−135
	−45	−15	−30	−60	0
	−135	−105	−120	−150	−90
60°	0	60	30	−30	90
	90	150	120	60	180
	180	−120	−150	150	−90
	−45	15	−15	−75	45
	−135	−75	−105	−165	−45
90°	0	90	45	−45	135
	90	180	135	45	−135
	180	−90	−135	135	−45
	−45	45	0	−90	90
	−135	−45	−90	−180	0
120°	0	120	60	−60	180
	90	−150	150	30	−90
	180	−60	−120	120	0
	−45	75	15	−105	135
	−135	−15	−75	165	45
150°	0	150	75	−75	−135
	90	−120	165	15	−45
	180	−30	−105	105	45
	−45	105	30	−120	180
	−135	15	−60	150	90

orientations: the first orientation of the prime (Frame 1), the second orientation of the prime (Frame 2), the orientation halfway between the two prime orientations (Intermediate), an orientation before the first orientation of the prime (Extra 1), or an orientation beyond the second orientation of the prime (Extra 2). The orientation of the Intermediate target equaled the first orientation of the prime plus half the rotation angle. The Extra 1 orientation equaled the first orientation of the prime minus half of the rotation angle. The Extra 2 orientation equaled the second orientation of the prime plus half of the rotation angle. Thus, the orientation of the Extra 1 target deviated from the orientation of the first prime by the same amount that the orientation of the Extra 2 target deviated from the orientation of the second prime and by the same amount that the orientation of the Intermediate target deviated from the orientations of both primes. Table 1 shows all of the target orientations, with 0° referring to the upright orientation.

Each target orientation was run in a separate block. Block order was counterbalanced across subjects. Stimulus order was randomized within each block. Each block contained 10 trials in which the targets matched each other as well as the prime, 10 trials in which the targets matched each other but differed from the prime, and 20 trials in which the targets differed from each other and the prime.

In a between-subjects design, five groups of 10 subjects completed the apparent motion condition, and five groups of 10 subjects completed the control condition. Each group of subjects observed stimuli at only one rotation angle so that every subject saw only novel prime objects, that is, the subjects had not previously seen the prime objects in other orientations.

RESULTS

Only reaction times for correct responses are reported because all subjects exhibited ceiling levels of performance. Priming is reported as a repeated measurement, or as the difference in reaction time between trials in which the prime and targets were identical and trials in which the prime and targets differed. The results are reported on the basis of subjects and collapsed over items, orientation of the first prime, and rotation direction (clockwise or counterclockwise) because no systematic pattern of differences was observed for these variables.

Did Priming Occur?

Repeated analyses of variance (ANOVAs) with priming as the within-measure variable indicated significant priming for Frame 1 (F[1, 90] = 115.3, $p < .001$), Frame 2 (F[1, 90] = 112.8, $p < .001$), Intermediate (F[1, 90] = 59.5, $p < .001$), and Extra 2 (F[1, 90] = 12.6, $p < .001$) but not for Extra 1 (F[1, 90] = 0.005, $p > .94$). Table 2 reports priming in the apparent motion and control conditions across rotation angle. The reported p values were derived from one-tail

Table 2 *Amount of priming (expressed in milliseconds) across rotation angles for all possible target orientations in the apparent motion condition and the control condition*

Rotation angle	Frame 1		Frame 2		Intermediate		Extra 1		Extra 2	
	Priming	p	Priming	p	Priming	p	Priming	p	Priming	p
					Apparent motion condition					
30°	91 7	003	126 5	000	87 1	022	−61 4	019	42 2	049
60°	46 7	001	114 2	000	85 5	017	−24 1	290	19 4	127
90°	113 0	001	101 8	022	56 7	025	−12 2	383	14 0	182
120°	44 2	013	74 5	028	68 0	010	−80 9	038	5 2	446
150°	56 6	004	102 5	015	62 9	006	12 7	302	−76 4	013
					Control condition					
30°	163 1	001	126 3	003	160 2	000	54 8	003	84 6	000
60°	83 0	001	83 2	000	43 4	043	83 6	034	71 1	015
90°	114 1	000	101 7	000	63 5	000	84 0	023	100 4	004
120°	89 4	005	49 0	015	0 8	486	−38 8	048	17 6	258
150°	106 0	018	138 6	000	4 6	433	−24 5	068	7 0	382

paired t tests used to test the unidirectional hypothesis that faster reaction times are observed when the targets match the prime than when the targets differ from the prime

Amount of Priming

A repeated ANOVA with priming as the within-measure variable and condition (apparent motion or control), rotation angle (difference between orientations of the first and second primes), and test frame (Frame 1, Frame 2, Intermediate, Extra 1, or Extra 2) as the independent variables indicated significant main effects of priming (F[1, 450] = 203 2, $p <$ 001), condition (F[1, 450] = 12 2, $p <$ 001), rotation angle (F[4, 450] = 8 6, $p <$ 001), and test frame (F[4, 450] = 23 1, $p <$ 001) A significant interaction was shown between condition and test frame, F(4, 450) = 4 2, $p =$ 002

Priming differences between the apparent motion and control conditions are summarized in Figure 3 A repeated ANOVA with priming as the within-measure variable and rotation angle and test frame as the independent variables showed significant main effects of priming, F(1, 225) = 52 8, $p <$ 001, and test frame, F(4, 225) = 18 6, $p <$ 001, in the apparent motion condition No significant effect of rotation angle was observed, F(4, 225) = 1 3, $p =$ 256 Fisher's post hoc comparisons showed that Frame 1, Frame 2, and Intermediate were significantly more primed than Extra 1 and Extra 2 ($p <$ 001) The same analysis in the control condition showed significant main effects of priming (F[1, 225] = 174 3, $p <$ 001), rotation angle (F[4, 225] = 9 6, $p <$ 001), and test frame (F[4, 225] = 7 8, $p <$ 001) Fisher's post hoc comparisons showed that Frame 1 and Frame 2 were significantly more primed than Intermediate, Extra 1, and Extra 2 ($p <$ 01)

For small rotation angles (30°–90°), a one-way ANOVA with priming as the dependent variable and test frame as the independent variable showed a main effect of test frame (F[4, 145] = 11 9, $p <$ 001) in the apparent motion condition but not in the control condition (F[4, 145] = 1 2, $p =$ 269) Fisher's post hoc comparisons showed that Frame 1, Frame 2, and Intermediate were significantly more primed than Extra 1 ($p <$ 001) and Extra 2 ($p <$ 01) in the apparent motion condition Frame 1 was significantly more primed than Extra 1 ($p <$ 05) in the control condition

For large rotation angles (120° and 150°), a main effect of test frame was found in both the apparent motion condition (F[4, 95] = 7 3, $p <$ 001) and the control condition (F[4, 95] = 9 4, $p <$ 001) Fisher's post hoc comparisons showed that Frame 1, Frame 2, and Intermediate were significantly more primed than Extra 1 ($p <$ 01) and Extra 2 ($p <$ 01) in the apparent motion condition In the control condition, Frame 1 and Frame 2 were significantly more primed than Intermediate ($p <$ 01), Extra 1 ($p <$ 01), and Extra 2 ($p <$ 01)

DISCUSSION

Motion appears to play a critical role in object representation That is, observers more readily recognize moving, as compared with static, objects in novel orientations More precisely, the results indicate that targets having the same orientation as the primes were readily primed in both the apparent motion and the control conditions across all rotation angles However, when the target orientation differed from the prime orientations, response times differed significantly across the two conditions Novel target orientations falling in between the two prime orientations were primed across all rotation angles in the apparent motion condition In the control condition, these Intermediate orientations were primed only for small angles Target orientations falling outside either end of the rotation path were primed at small rotation angles in the control condition However, in the apparent motion condition, novel orientations falling outside the rotation path were primed only at the smallest rotation angle and only when the orientation was in the direction of rotation (i e , Extra 2)

It is important to note that the apparent motion and control conditions differed in two ways perception of apparent motion and ISI duration In a subsequent control study, we examined whether the perception of apparent motion, rather than ISI, caused these priming differences To that end, path-guided apparent motion (R N Shepard & Zare, 1983) was employed by presenting a dim, gray band connecting the two prime views during the long ISI of the control con-

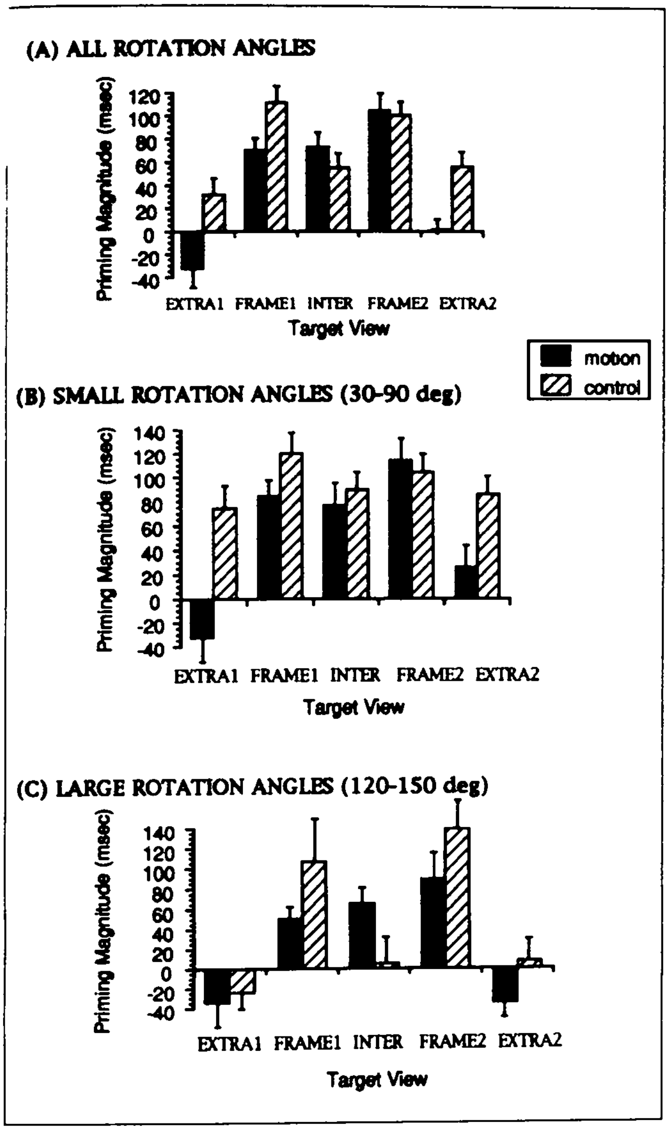

Fig 3 Magnitude of priming for target orientations in the apparent motion and control conditions collapsed across (a) rotation angles, (b) small rotation angles only (30°–90°), and (c) large rotation angles only (120° and 150°) INTER = Intermediate target

dition With this manipulation, subjects reported the perception of apparent rotation The pattern of priming found in this new control condition did not differ significantly from that of the apparent motion condition reported here This result further supports the hypothesis that enhanced priming within paths of motion results from the perception of motion and not from the use of short ISIs

Taken together, these results support the hypothesis that object recognition is based on restricted generalization fields around "known" object views (Bulthoff & Edelman, 1992) Yet the current results are significant because they demonstrate that these generalization fields can be modified without practice When a static object is presented in a novel orientation, recognition of that object is primed only for small orientation differences That is, observers can generalize only up to 45° rotations in the frontal plane from known orientations However, when two views of an object in different orientations are linked by apparent motion, generalization fields appear to be tuned or sharpened such that priming is facilitated within but inhibited outside the motion path Thus, an object can be recognized more rapidly in novel orientations as long as those orientations fall within the path of apparent motion Interestingly, an object can also be readily recognized in novel orientations just outside the motion path as long as the orientations are consistent with momentum (Freyd, 1987)

In summary, observers can readily recognize a moving object in a novel orientation Thus, although view invariance for within-category object recognition requires extensive practice one-shot view invariance can be achieved with moving objects The behavioral tuning of generalization fields for dynamic objects appears to be precisely restricted to the path and direction of motion It would be particularly interesting to investigate whether neural generalization fields are similarly tuned for moving objects

Acknowledgments—We would like to thank Shimon Edelman Peter Gerhardstein and Ron Rensink for their helpful comments and suggestions on previous versions of this manuscript This research was supported by a grant from the Sigma Xi Foundation to the first author and a grant from the National Eye Institute (EY099310) to the second author Some of the results discussed in this article were first presented at the May 1996 meeting of the Association for Research in Vision and Ophthalmology in Ft Lauderdale Florida

REFERENCES

Biederman I & Gerhardstein P C (1993) Recognizing depth rotated objects Evidence and conditions for three-dimensional viewpoint invariance *Journal of Experimental Psychology Human Perception and Performance 19* 1162–1182

Bulthoff H H & Edelman S (1992) Psychophysical support for a 2 D view interpolation theory of object recognition *Proceedings of the National Academy of Sciences USA 89* 60–64

Edelman S & Weinshall D (1991) A self organizing multiple view representation of 3D objects *Biological Cybernetics 64* 209–219

Freyd J J (1987) Dynamic mental representations *Psychological Review 94* 427–438

Hummel J E & Biederman I (1992) Dynamic binding in a neural network for shape recognition *Psychological Review 99* 480–517

Jolicoeur P (1985) The time to name disoriented natural objects *Memory & Cognition, 13* 289–303

Jolicoeur P (1988) Mental rotation and the identification of disoriented objects *Canadian Journal of Psychology 42* 461–478

Jolicoeur P (1990) Identification of disoriented objects A dual systems theory *Mind & Language 5* 387–410

Jolicoeur P & Landau M J (1984) Effects of rotation in the identification of similar visual patterns *Canadian Journal of Psychology 38* 80–93

Koriat, A & Norman J (1988) Frames and images Sequential effects in mental rotation *Journal of Experimental Psychology Learning Memory and Cognition 14* 93–111

Koriat, A & Norman J (1989) Establishing global and local correspondence between successive stimuli *Journal of Experimental Psychology Learning Memory and Cognition 15* 480–495

Logothetis N Pauls J Bülthoff H H & Poggio T (1994) View-dependent object recognition by monkeys *Current Biology 4* 401–414

Logothetis N Pauls J & Poggio T (1995) Shape representation in the inferior temporal cortex of monkeys *Current Biology 5* 552–563

Marr D & Nishihara, H K (1978) Representation and recognition of three-dimensional shapes *Proceedings of the Royal Society of London, Series B 200* 269–294

Palmer S Rosch E & Chase P (1981) Canonical perspective and the perception of objects In J Long & A Baddeley (Eds) *Attention and performance IX* (pp 135–151) Hillsdale NJ Erlbaum

Poggio T & Edelman S (1990) A network that learns to recognize three-dimensional objects *Nature 343* 263–266

Rock, I DiVita, J & Barbeito R (1981) The effect on form perception of change of orientation in the third dimension *Journal of Experimental Psychology Human Perception and Performance 7* 719–732

Sekuler A B & Palmer S E (1992) Perception of partly occluded objects A microgenetic analysis *Journal of Experimental Psychology General 121* 95–111

Shepard R N & Cooper L A (1982) *Mental images and their transformations* Cambridge MA MIT Press

Shepard, R N & Judd S A (1976) Perceptual illusions of rotation of three-dimensional objects *Science 191* 952–954

Shepard, R N & Metzler J (1971) Mental rotation of three-dimensional objects *Science 171* 701–703

Shepard R N & Zare S (1983) Path-guided apparent motion *Science 220* 632–634

Shepard, S & Metzler D (1988) Mental rotation Effects of dimensionality of objects and type of task *Journal of Experimental Psychology Human Perception a*
Performance 14 3–11

Tarr M J & Pinker S (1989) Mental rotation and orientation-dependence in sha
recognition *Cognitive Psychology 21* 233–282

Tarr M J & Pinker S (1990) When does human object recognition use a viewe
centered reference frame? *Psychological Science 1* 253–256

Ungerleider L G & Mishkin M (1982) Two cortical visual systems In D G Ingl
M A Goodale & R J W Mansfield (Eds) *Analysis of visual behaviour* (pp 549
586) Cambridge MA MIT Press

(RECEIVED 9/6/96 ACCEPTED 2/19/97)