

Contrast- and illumination-invariant object recognition from active sensation

INGO RENTSCHLER^{1,*}, EROL OSMAN¹ and MARTIN JÜTTNER²

¹ *Institute of Medical Psychology, University of Munich, Munich, Germany*

² *Psychology, School of Life and Health Sciences, Aston University, Birmingham, UK*

Received 27 August 2008; accepted 29 December 2008

Abstract—It has been suggested that the deleterious effect of contrast reversal on visual recognition is unique to faces, not objects. Here we show from priming, supervised category learning, and generalization that there is no such thing as general invariance of recognition of non-face objects against contrast reversal and, likewise, changes in direction of illumination. However, when recognition varies with rendering conditions, invariance may be restored and effects of continuous learning may be reduced by providing prior object knowledge from active sensation. Our findings suggest that the degree of contrast invariance achieved reflects functional characteristics of object representations learned in a task-dependent fashion.

Keywords: Object recognition; invariant recognition; contrast reversal; direction of illumination; categorization.

1. INTRODUCTION

Contrast reversal disrupts the visual recognition of human faces from grey-level (Galper, 1970; Hayes *et al.*, 1986; Liu *et al.*, 1999) and two-tone images (Phillips, 1972), whereas little or no effect of contrast reversal on the recognition of chairs (Subramaniam and Biederman, 1997) and contraptions of geometric components ('Greebles'; Vuong *et al.*, 2005) has been reported. A possible explanation of this discrepancy is the existence in test objects of differences in contrast-invariant shape features resulting from discontinuities of orientation and depth (Nederhouser *et al.*, 2007). To exclude this possibility, these last authors compared the effects of contrast reversal on the recognition of faces and non-face stimuli composed of spherical harmonics. The results obtained were suggestive to these authors of contrast reversal being deleterious to the recognition of faces but not objects.

*To whom correspondence should be addressed. E-mail: Ingo.Rentschler@med.uni-muenchen.de

An alternative account of the dependence of visual recognition on contrast reversal has been offered by Jüttner *et al.* (2004). These authors studied supervised category learning (Caelli *et al.*, 1987) and generalization to contrast reversal for several sets of compound Gabor signals (grey-level patterns composed of Gabor waveforms; Caelli *et al.*, 1986). The sets differed either locally in their grouping in feature space or globally in the grey-level distribution of learning patterns. Jüttner and co-workers found an inverse relationship between the set-specific duration of category learning to criterion and generalization to contrast reversal whereby fast learning coincided with invariance to contrast reversal. They analysed their behavioural data by means of computer simulations employing an evidence-based, structural pattern classifier (*cf.* Caelli and Dreier, 1994; Jain and Hoffman, 1988; Jüttner *et al.*, 1997). For a given categorization task, Jüttner *et al.* obtained sets of part and part-relational attributes, or feature states, as solutions within this evidence-based system (EBS). They found the psychophysically observed speed of category learning to be directly proportional to the number of EBS solutions in search space. Importantly, fast learning was associated with high relative frequency of contrast-invariant feature states used for generating class descriptions.

On the one hand, the findings by Jüttner *et al.* (2004) indicate that invariance to contrast in pattern recognition depends on task- or context-dependent characteristics of learned representations (*cf.* Albright, 1995). On the other hand, the visual recognition of three-dimensional (3-D) objects could involve more abstract types of representation in the brain (Biederman, 1987; Biederman and Kalocsai, 1997) that entail a generic, learning-independent invariance to contrast reversal. To resolve this issue, we examine the effects of changed rendering within a category learning paradigm involving unfamiliar 3-D objects (*cf.* Osman *et al.*, 2000; Rentschler *et al.*, 2008). More specifically, we study supervised category learning as well as generalization to novel viewpoints and changed rendering using 3-D objects composed of a fixed number of identical spheres. For probing the nature of object representations, we control the amount of prior knowledge subjects have about test objects. This is achieved through priming, a technique from memory research using the beneficial influence of pre-exposure to a stimulus in the absence of explicit instructions to remember the stimulus (see Biederman and Cooper, 1991; Cooper *et al.*, 1992; Rentschler *et al.*, 2008).

The conjecture of invariance characteristics of object recognition being task-dependent entails the possibility that internalised representations are susceptible to continuous learning (Keith-Magee, 2001; Tagscherer, 2001; Tagscherer and Protzel, 1998). That is, visual object recognition may not only be enabled by category learning but continuously modified by interfering input signals (*cf.* Edelman, 1987). We investigate this possibility in an additional experiment, where participants classify object views in standard rendering after having performed with test views in changed rendering.

2. METHODS

2.1. Subjects

Forty-five subjects participated for pay. They ranged in age from 19 to 45 years and were free of ophthalmological and neurological disorders.

2.2. Stimuli

2.2.1. Learning objects. A set of three test objects was constructed using four identical sphere parts in varying spatial arrangements (Fig. 1(a)). Objects were designed according to the concept of chirality. Chiral objects cannot be carried into their mirror images by rotations and translations. Handed chiral objects allow categorization into ‘left’ and ‘right’ objects (King, 2003). Bilaterally symmetric objects are achiral by definition. Thus, one non-handed achiral object (Object 1, *cf.* Fig. 1(a)) and one pair of handed chiral objects (Objects 2 and 3) were used. Physical models were constructed from directly connected polystyrene balls each measuring 6 cm in diameter. Three spheres formed a rectangular isosceles triangle. The fourth sphere was placed perpendicularly above the centre of one of the base spheres. Corresponding virtual models were generated and displayed for 250 ms each as perspective 2-D projections by means of the Open Inventor™ (Silicon Graphics, Inc.) 3-D developer’s toolkit. A lighting model of mixed directed and

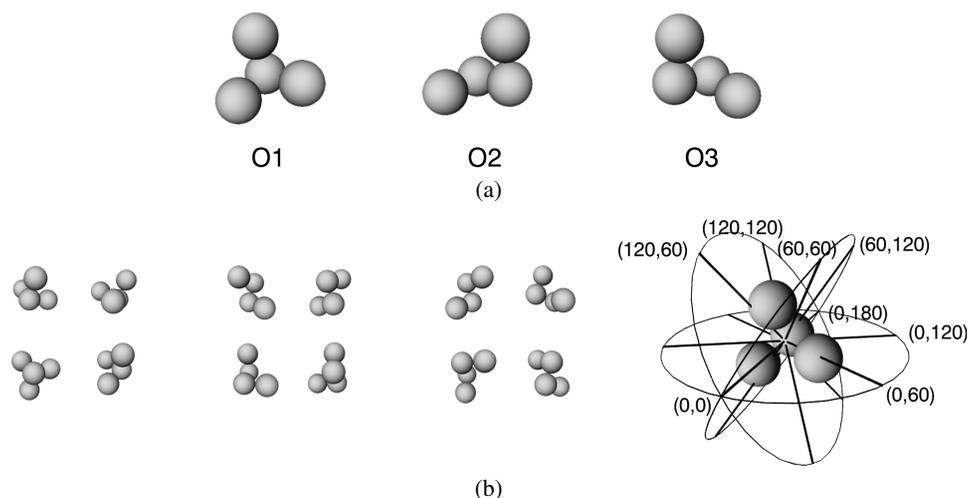


Figure 1. 2-D views of the set of objects for category learning and generalization. (a) Achiral (bilaterally symmetric; left) Object 1 (O1) and handed chiral (mirror image; centre and right) Objects 2 (O2) and 3 (O3). Three spheres formed a rectangular isosceles triangle. The fourth sphere was placed perpendicularly above the centre of one of the base spheres. All spheres were of same size and reflectance. (b) Sample of four 2-D views for each of the objects (left). Views obtained from sampling azimuth and polar angle of the viewing (half) sphere in 60° steps, thus, yielding 8 views per object (right). 2 views of Object 1 were redundant due to symmetry reasons. Discarding them, 22 views (6 for Object 1 and 8 for each of Objects 2 and 3) were used for category learning to criterion.

diffuse illumination lacking cast shadows was used. At the viewing distance of 1 m, virtual objects subtended in average 7.4° .

2.2.2. Learning views. Discarding views redundant due to rotational symmetry, 22 learning views (6 for Object 1, and 8 for each of Objects 2 and 3) were generated by sampling azimuth and polar angle of the viewing-sphere in 60° steps. Steps were taken on three great circles inclined against each other by 60° . The equatorial plane was horizontal and contained the symmetry axis of each object (three-fold rotational symmetry of Object 1 and two-fold rotational symmetries of Objects 2 and 3). The origin of the coordinate system used for sampling was at the centre of gravity of the component spheres and, thus, on the symmetry axis of each object. To reduce occlusion artefacts, reference views (Fig. 1(a)) were chosen such that the objects' symmetry axes were perpendicular to the picture plane and no centre of any sphere was directly in the equatorial plane. An angle of rotation in the picture plane, randomly selected from the values of 0° , 60° , 120° and 180° , was employed to prevent observers from using the frame of the computer screen as a reference. To examine the possibility of recognition depending on pixel-based similarities of input signals and stored 2-D views (*cf.* Bülthoff and Edelman, 1992), classification probabilities were predicted from a Minimum-Distance Classifier (see Ahmed and Rao, 1975) using the maximum of the 2-D cross-correlation function as a measure of similarity (*cf.* Caelli and Rentschler, 1986). Predicted classification probabilities were at chance level for all three objects, demonstrating that no differences of pixel-based similarities exist within the learning set. Thus, the 2-D views of the learning set qualify as 'structure-only' stimuli.

2.2.3. Test views. In order to test spatial generalization, 64 views from novel viewpoints were obtained from sampling the viewing sphere as above but in 30° steps. Views redundant due to rotational symmetry or identical to those used during supervised learning were discarded. Finally, the views were manipulated using three different rendering conditions. For views in the standard rendering condition the lighting model remained unaltered. In the contrast reversal condition (Fig. 2(a)), the grey level of each pixel was subtracted from the maximum value 255 to obtain reversed intensities relative to the mean (background) level of 128 (8 bit resolution with linearized grey-level-to-luminance function). For views in the condition of changed direction of illumination (Fig. 2(b)), the default position of the light source was changed from the upper left (infinite distance, 10° to the left and up from the viewer's position) to the lower right from the observer's position (infinite distance, 45° to the right and 90° down from the viewer's position).

2.3. Procedure

The main experiment consisted of three parts: priming, supervised category learning and generalization. Subjects who had carried out the generalization test under the

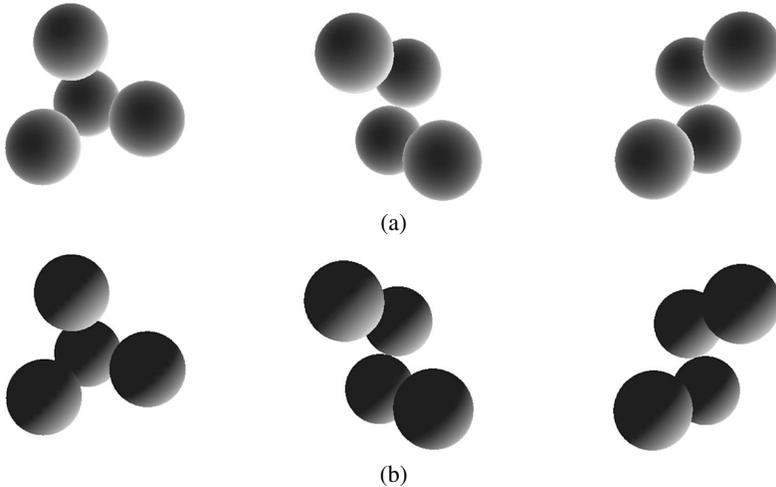


Figure 2. Changed object rendering. (a) For contrast reversal, original intensities were subtracted from the maximum value 255 to obtain reversed intensities. Background intensities remained fixed under these conditions. (b) For a change in the direction of illumination, the default position of the virtual light source was changed from the upper left corner to the lower right corner from the observer's position.

rendering conditions of contrast reversal or changed direction of illumination (see below) completed an additional generalization re-test, using the test views of the standard rendering condition.

2.3.1. Priming phase. During priming, subjects either explored with their hands the physical objects occluded from sight (*active touch* priming) or explored virtual models by grasping 2-D views on the computer display with the mouse and a cursor and rotating them within the viewing sphere (*active vision* priming). Priming lasted for 5 min in either condition. No instruction other than the invitation to familiarize themselves with the objects was given. As a baseline for the two priming conditions, a neutral (*control*) condition was included, where subjects entered the second part of the experiment, visual category learning, directly, i.e., without acquiring prior object knowledge. Participants were randomly assigned to one of the three experimental conditions, resulting in 15 participants per condition.

2.3.2. Supervised category learning. The second part of the experiment involved a procedure of supervised learning partitioned into learning units (for details, see Osman *et al.*, 2000; Rentschler *et al.*, 1994). Each learning unit consisted of a learning phase and a test phase. During the learning phase each of the 22 views of the learning set was presented once for 250 ms and in random order, followed by the corresponding object label ('1', '2' or '3') displayed for 1 s. During the test phase, each learning view was presented once, whereupon the observer assigned it to one of the three objects. Upon completion of each test phase, participants received feedback as to the percent correct value of their responses. No instructions as to

the possible duration of classification decisions were given. On average duration, a learning unit took 5 min to complete. The series of learning units terminated upon reaching a criterion of 90% correct classification.

2.3.3. Generalization test. The final part of the main experiment assessed generalization to novel viewpoints and to changes in rendering. Here, each of 64 novel test views of one rendering condition was shown three times for 250 ms and in random order. Upon each presentation, the observer had to assign the view to one of the object categories 1 to 3. The 15 participants within each priming condition were randomly assigned to one of the three rendering conditions, resulting in 5 participants per priming and rendering condition.

2.3.4. Generalization re-test. To test for effects of continuous learning, participants who had completed the generalization test under the conditions of changed rendering (*contrast reversal* or *changed direction of illumination*) were tested again for generalization to novel viewpoints, this time using the test views of the standard rendering condition. The same procedure as during the generalization test was employed. However, the test views were presented only once in random order.

3. RESULTS

Table 1 summarises durations of supervised category learning to criterion (90% correct classification) for the control condition, where no prior knowledge about test objects was provided, as well as the two priming conditions of active vision and active touch. Learning duration, as measured by the number of learning units, is distinctly reduced by either form of priming relative to the control group.

From the confusion matrix of each observer, scores of percentage correct for each of the three objects were computed to assess classification performance in the generalization test. Figure 3(a–c) shows mean generalization performance to novel viewpoints for the control condition and the priming conditions of active vision and active touch. Each graph shows three pairs of bars, with the left (bright) bar corresponding to generalization performance for non-handed Object 1 and the right (dark) bar to mean performance for handed Objects 2 and 3. In order to test

Table 1.

Duration of category learning to criterion for the three conditions of control, active vision priming and active touch priming

	Control	Active vision	Active touch
Mean	35.1	17.3	17.5
S.E. (<i>N</i>)	4.8 (15)	3.5 (15)	3.7 (15)

Mean values over groups of participants and standard errors in number of learning units (*N* = 15, standard rendering; *N* = 15, contrast reversal and *N* = 15, direction of illumination).

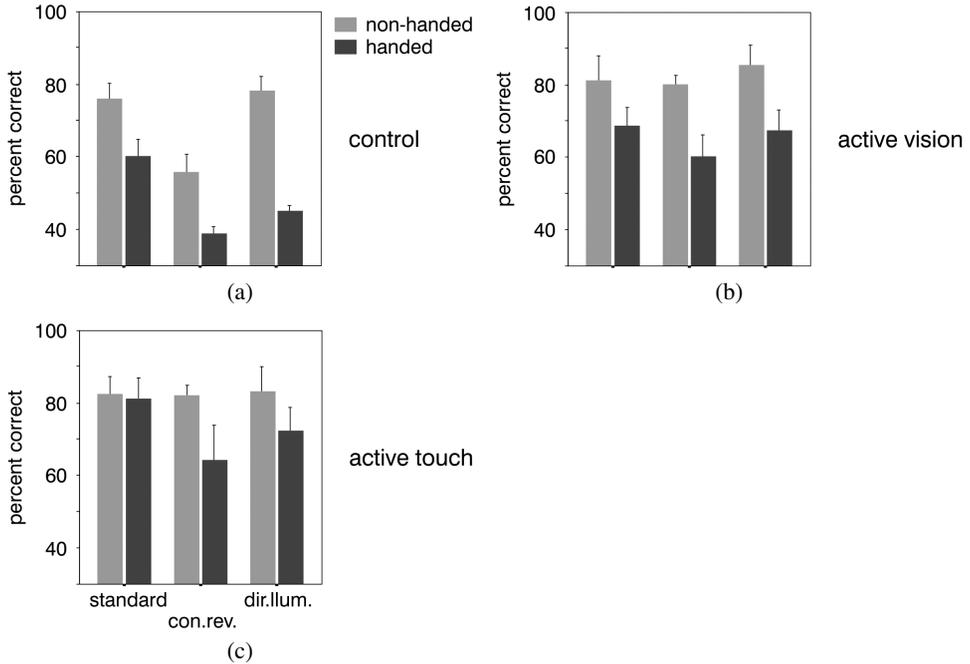


Figure 3. Spatial generalization under three conditions of rendering and three conditions of priming. Standard rendering (leftmost double bars), contrast reversal (centre) and changed direction of illumination (rightmost double bars). Priming conditions of control (top left), active vision (top right) and active touch (bottom). Generalization performance is given in percent correct classification. Left bar of double bars: Object 1; right bar of double bars: mean value of Objects 2 and 3. Error bars: 1 S.E. ($N = 5$; O1); mean values of S.E.s ($N = 5$ each; O2 and O3).

the impact of priming and rendering, we analysed the data in terms of separate 3 (priming condition: control vs. active vision vs. active touch) \times 3 (rendering condition: standard vs. contrast reversal vs. changed direction of illumination) ANOVAs for handed and non-handed objects. The analysis yielded a significant main effect of priming condition (non-handed object: $F(2, 36) = 6.72$, $p < 0.01$; handed objects: $F(2, 36) = 15.92$, $p < 0.001$). Rendering condition significantly affected generalization in case of handed Objects 2 and 3 ($F(2, 36) = 6.02$, $p < 0.01$) and marginally in case of non-handed Object 1 ($F(2, 36) = 3.13$, $p = 0.05$). There was no significant interaction between the two factors (non-handed object: $F(4, 36) = 1.80$, $p = 0.15$; handed objects: $F(4, 36) = 0.55$, $p = 0.70$).

To further assess the effect of rendering, additional one-way ANOVAs were conducted for each priming condition and object type. For the control condition, there was a significant effect of rendering on generalization for both object types (non-handed object: $F(2, 12) = 7.36$, $p < 0.01$; handed objects: $F(2, 12) = 16.69$, $p < 0.001$). As illustrated in Fig. 3(a) and confirmed by *post-hoc* Bonferroni comparisons, contrast reversal significantly impeded generalization for both the handed and non-handed object (all p 's < 0.05). A change in the direction of

illumination impeded generalization of the handed objects ($p < 0.01$) while leaving recognition of the non-handed object unaffected ($p = 0.99$).

No significant effect of rendering on generalization was observed in the case of active vision or active touch priming for both non-handed Object 1 (active vision: $F(2, 12) = 0.30$, $p = 0.75$; active touch: $F(2, 12) = 0.01$, $p = 0.99$) and handed Objects 2 and 3 (active vision: $F(2, 12) = 0.64$, $p = 0.54$; active touch: $F(2, 12) = 1.37$, $p = 0.29$). For both types of objects, categorization performance was invariant to contrast reversal and changes in the direction of illumination. Still, paired t -tests revealed that for priming from active vision performance for handed Objects 2 and 3 was consistently lower than for non-handed Object 1 in these two rendering conditions ($t(4) > 3.26$, $p < 0.05$ and $t(4) = 7.36$, $p < 0.01$, respectively), whereas there was no significant difference for the baseline condition of standard rendering ($t(4) = 1.98$, $p = 0.12$). Priming from active touch yielded a generalization performance for the two types of objects that differed significantly only in case of changed illumination ($t(4) = 4.55$, $p < 0.05$) but not for standard rendering ($t(4) = 0.27$, $p = 0.80$) or contrast reversal ($t(4) = 2.10$, $p = 0.11$).

Figure 4(a–c) shows results obtained in the re-test of spatial generalization to standard rendering for those subjects who had first been tested in the conditions of contrast reversal or changed direction of illumination (middle and right double bars). For comparison, generalization performance of the not re-tested group of participants (i.e., those who had directly been tested with novel views under conditions of standard rendering) is re-plotted from Fig. 3 (open bars).

Separate 3 (priming condition: control vs. active vision vs. active touch) \times 3 (rendering condition: standard direct vs. standard after contrast reversal vs. standard after changed direction of illumination) ANOVAs for both types of object revealed significant main effects of priming condition (non-handed object: $F(2, 36) = 4.28$, $p < 0.05$; handed objects: $F(2, 36) = 12.44$, $p < 0.001$) and rendering condition (non-handed object: $F(2, 36) = 6.56$, $p < 0.01$; handed objects: $F(2, 36) = 4.30$, $p < 0.05$). There was no significant interaction between the two factors (non-handed object: $F(4, 36) = 1.67$, $p = 0.18$; handed objects: $F(4, 36) = 0.34$, $p = 0.85$).

To further test for effects of continuous learning, additional one-way ANOVAs were conducted for each of the priming conditions and object types. For the control condition, there was a significant effect of the preceding test involving changed rendering on generalization performance in the re-test (non-handed object: $F(2, 12) = 12.21$, $p < 0.001$; handed objects: $F(2, 12) = 4.88$, $p < 0.05$). As illustrated in Fig. 4(a) and confirmed by *post-hoc* Bonferroni comparisons, this was mainly due to the detrimental effect of contrast reversal. Preceding testing with contrast-reversed images significantly impeded generalization in the re-test for both the handed and non-handed object (all p 's < 0.05), whereas a change in the direction of illumination left recognition of both types of objects during the re-test unaffected (all p 's > 0.24).

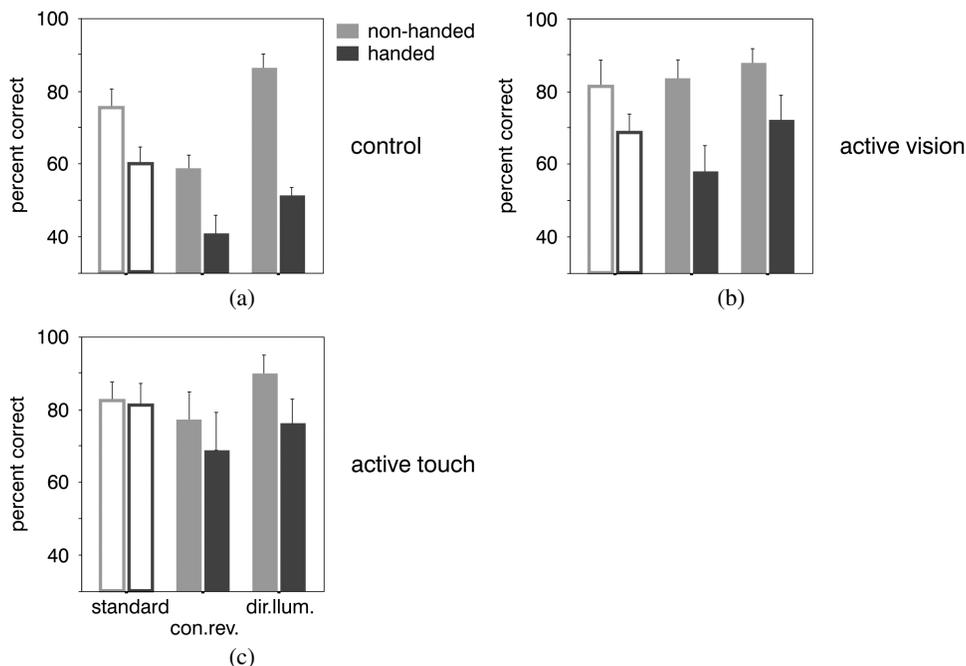


Figure 4. Re-test of spatial generalization under conditions of standard rendering following spatial generalization under conditions of contrast reversal (middle double bars), and changed direction of illumination (rightmost double bars). Data for participants, who were initially tested under conditions of standard rendering, reproduced from Fig. 3 (leftmost double bars). Priming conditions of control (top left), active vision (top right) and active touch (bottom). Generalization performance is given in percent correct classification. Left bars of double bars: Object 1; right bars of double bars: mean value of Objects 2 and 3. Error bars: 1 S.E. ($N = 5$; O1); mean values of S.E.s ($N = 5$ each; O2 and O3).

No significant evidence for continuous learning was found in case of active vision or active touch priming for non-handed Object 1 (active vision: $F(2, 12) = 0.37$, $p = 0.70$; active touch: $F(2, 12) = 1.29$, $p = 0.31$) and handed Objects 2 and 3 (active vision: $F(2, 12) = 1.42$, $p = 0.28$; active touch: $F(2, 12) = 0.72$, $p = 0.51$). For both types of objects, generalization performance in the re-test after a preceding test with changed rendering did not differ significantly from that without such test history. This indicates that prior knowledge from active visual or tactual sensation enabled the acquisition of more robust object representations.

4. DISCUSSION

We have shown that contrast reversal or a change in direction of illumination may be deleterious to the visual recognition of non-face objects. Yet it is possible to enable invariant recognition by providing observers with prior object knowledge from active sensation (active vision or active touch). Such type of information does not only diminish the effects of changed rendering but allows the construction of

object representations that are relatively robust against perturbation by intervening input signals, i.e., continuous learning.

Our results qualify previous claims of a generic invariance of non-face objects against contrast reversal (Nederhouser *et al.*, 2007). Indeed, the latter would seem to be in stark contrast to what is known from studies in perceptual learning. Discrimination performance typically improves with practice but does not transfer to new stimulus conditions (e.g., Fiorentini and Berardi, 1980; Karni and Sagi, 1991; Poggio *et al.*, 1992). Accordingly, perceptual learning in motion discrimination is highly specific to practised directions of motion for difficult discrimination tasks but transfers to new directions when the difficulty is reduced (Liu, 1999). Similarly, Jüttner *et al.* (2004) have shown that the effects of contrast reversal on visual pattern recognition are determined by task-dependent characteristics of learned representations.

Here we have observed that changes in rendering clearly impair visual object categorization but prior knowledge from active sensation may facilitate recognition and make it more invariant against changes in rendering. Furthermore, prior knowledge from active sensation may reduce continuous learning. To better understand these results, it is helpful to note that machine learning systems are typically based on the assumption of learning being restricted to a certain time interval (learning phase), whereupon systems are put into operation (working phase). Yet there exist systems that allow learning to take place all the time in an uninterrupted fashion, thus, being able to adapt to time-variant environments (Keith-Magee, 2001; Tagscherer, 2001; Tagscherer and Protzel, 1998). Moreover, there is reason to believe that perceptual categorization and generalization are adaptive *per se* (Edelman, 1987). Thus it is not surprising that more reliable information on the physical 3-D-structure of the environment from active sensation entails the construction of more invariant and robust internal representations for visual object categorization.

The present study differs from earlier work on visual object recognition mainly in two respects. First, a paradigm of category learning involving 2-D views from all over the viewing (half-) sphere of non-handed and handed structure-only objects was employed. This precluded recognition from non-accidental object parts by necessitating the construction and maintenance in working memory (see Baddeley, 1986) of object representations conserving at least aspects of 3-D-structure. By contrast, Nederhouser and co-workers (2007) used a matching-to-sample paradigm, where participants were presented simultaneously with three object views from same viewpoint of which two portrayed the same object. Thus, observers could rely on image discrimination by comparing contrast-invariant shape features. For instance, in Fig. 4(a) of Nederhouser *et al.*, the distractor and the right matching view show at their lowest central positions the same acute protuberance, which is lacking in the left matching view.

Second, the effects of changed rendering were most conspicuous for handed chiral objects. Concerning this observation, it is helpful to note that mirror-symmetric

Object 1 can be carried into itself by mirror reflections (3-fold mirror symmetry; see Weyl, 1952). There is evidence that such symmetries facilitate object recognition (Vetter *et al.*, 1994). Handed chiral Objects 2 and 3 can be carried into themselves only by movements enacted on rigid bodies such as pure rotations (see International Union of Crystallography, 2005), whereas mirror reflections carry them into each other, thus, contributing to their confusion. Indeed, physical models of handed chiral objects are disambiguated by spatial transformation in 3-D and alignment to a scene-based reference (Hinton and Parsons, 1988), whereas the recognition of handed objects from 2-D projections involves ‘mental rotation’ of internalized models (Shepard and Metzler, 1971). In brief, symmetry properties facilitate the recognition of non-handed Object 1 but impair the recognition of handed chiral Objects 2 and 3.

The possibility of non-handed Object 1 being recognized from about half of its test views by relying on projected symmetries and the three-fold connectivity of one of its sphere parts (Fig. 1(b), left) has to be acknowledged. Similarly, about half of the views of handed Objects 2 (Fig. 1(b), middle) and 3 (Fig. 1(b), right) display a ‘snake-like’ structure (introspective reports of participants), thus, allowing their distinction against non-handed Object 1 but not each other. However, for classifying the remaining ‘difficult’ (introspective reports of participants) views and in particular handedness, observers had to enact the full cycle of image understanding, i.e., the (mental) generation of candidate 3-D models from input images (inverse modelling), their rendering back into images (forward modelling) and the corresponding matching process (for formal solutions in the area of machine intelligence, see Caelli and Bischof, 1997 and Cheng *et al.*, 2006). Given the relative disadvantage of recognising handedness in the absence of prior knowledge and, to some extent, with active vision priming, it can be assumed that observers relied on attributes, or features, in image format under such conditions. The dependence of recognition on rendering conditions would be the consequence of such classification strategies.

We shall then turn to the question of how prior knowledge from active sensation may have affected the construction of object representations. It would have drawn the attention of subjects to 3-D structures, thus, encouraging them to generate corresponding models (Thoma *et al.*, 2004). They may have achieved this by encoding temporal sequences of object views (active vision) or exploratory finger and hand movements along the physical object models (active touch). As object palpation directly evokes mental imagery (Critchley, 1953, Chap. IV), we can conjecture that some sort of kinetic object traces were stored in multimodal representations (e.g., Zangaladze *et al.*, 1999). Subjects may then have inferred 3-D structures from linking object parts exposed in 2-D views to such representations. Related categorization strategies would have enabled more invariant recognition.

In conclusion, we agree with Nederhouser *et al.* (2007) in that effects of contrast reversal depend on the nature of internal object representations rather than physical stimulus attributes such as spectral composition (e.g., Hayes *et al.*,

1986). The difference in perspective resides in the fact that Nederhouser and co-workers attributed the effect of contrast reversal on recognition to fixed and invariant stimulus categories (face *vs.* non-face objects), whereas we found it task- or context-dependent. We propose that prior object knowledge from active sensation causes representational shifts as have been reported for face recognition (see Palmeri *et al.*, 2004) and classification of mirror-image patterns (Rentschler and Jüttner, 2007). Such shifts enable increased invariance of recognition due to an increased use of contrast- or illumination-invariant part-relational object attributes. Moreover, they render object representations more robust, i.e., less sensitive to modification through interfering input signals.

Acknowledgement

Study supported by Deutsche Forschungsgemeinschaft (DFG) and dedicated to Professor Terrence Michael Caelli on occasion of his retirement from the National Institute of Information Technologies (NICTA) in Canberra, Australia.

REFERENCES

- Ahmed, N. and Rao, K. R. (1975). *Orthogonal Transforms for Digital Image Processing*. Springer Verlag, Berlin, Germany.
- Albright, T. D. (1995). My most true mind thus makes mine eye untrue, *Trends Neurosci.* **18**, 331–333.
- Baddeley, A. (1986). *Working Memory*. Oxford University Press, Oxford, UK.
- Biederman, I. (1987). Recognition-by-components: a theory of human image under standing, *Psychol. Rev.* **94**, 115–147.
- Biederman, I. and Cooper, E. E. (1991). Priming contour-deleted images: evidence for intermediate representations in visual object recognition, *Cognit. Psychol.* **23**, 393–419.
- Biederman, I. and Kalocsai, P. (1997). Neurocomputational bases of object and face recognition, *Phil. Trans. Royal Soc. London: Biol. Sci.* **352**, 1203–1219.
- Bülthoff, H. H. and Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition, *Proc. Natl Acad. Sci. USA* **89**, 60–64.
- Caelli, T. and Bischof, W. F. (1997). *Machine Learning and Image Interpretation*. Plenum Press, New York, USA.
- Caelli, T. and Dreier, A. (1994). Variations on the evidence-based object recognition theme, *Pattern Recognition* **27**, 185–204.
- Caelli, T., Hübner, M. and Rentschler, I. (1986). On the discrimination of micropatterns and textures, *Human Neurobiol.* **5**, 129–136.
- Caelli, T. and Rentschler, I. (1986). Cross-correlation model for pattern acuity, *J. Optic. Soc. Amer.* **A3**, 1948–1956.
- Caelli, T., Rentschler, I. and Scheidler, W. (1987). Visual pattern recognition in humans. I. Evidence for adaptive filtering, *Biol. Cybernet.* **57**, 233–240.
- Cheng, L., Caelli, T. and Sanchez-Azofeifa, A. (2006). Component optimization for image understanding: a Bayesian approach, *IEEE Trans. Pattern Anal. Machine Intell.* **28**, 684–693.
- Cooper, L. A., Schacter, D. L., Ballesteros, S. and Moore, C. (1992). Priming and recognition of transformed three-dimensional objects: effects of size and reflection, *J. Exper. Psychol.: Learning, Memory and Cognition* **18**, 43–57.
- Critchley, M. (1953). *The Parietal Lobes*. Edward Arnold, London, UK.

- Edelman, G. M. (1987). *Neural Darwinism*. Basic Books, New York, USA.
- Fiorentini, A. and Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency, *Nature* **287**, 43–44.
- Galper, R. E. (1970). Recognition of faces in photographic negative, *Psychon. Sci.* **19**, 207–208.
- Hayes, A., Morrone, M. C. and Burr, D. C. (1986). Recognition of positive and negative bandpass-filtered images, *Perception* **15**, 595–602.
- Hinton, G. E. and Parsons, L. M. (1988). Scene-based and viewer-centred representations for comparing shapes, *Cognition* **30**, 1–35.
- International Union of Crystallography (2005). Chirality. <http://www.iucr.org/iucr-top/comm/cteach/pamphlets/21/node7.html>
- Jain, A. and Hoffman, D. (1988). Evidence-based recognition of objects, *IEEE Trans. Pattern Anal. Machine Intell.* **10**, 783–802.
- Jüttner, M., Caelli, T. and Rentschler, I. (1997). Evidence-based pattern recognition: a structural approach to human perceptual learning and generalization, *J. Math. Psychol.* **41**, 244–258.
- Jüttner, M., Langguth, B. and Rentschler, I. (2004). The impact of context on pattern category learning and representation, *Visual Cognition* **11**, 921–945.
- Karni, A. and Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity, *Proc. Natl Acad. Sci. USA* **88**, 4966–4970.
- Keith-Magee, K. Unpublished PhD dissertation, Curtin University of Technology, Perth, WA.
- King, R. B. (2003). Chirality and handedness. The Ruch ‘shoe-potato’ dichotomy in the right-left classification problem, *Annals New York Acad. Sci.* **988**, 158–170.
- Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions, *Proc. Natl Acad. Sci. USA* **96**, 14085–14087.
- Liu, C. H., Collin, C. A., Burton, A. M. and Chaudhuri, A. (1999). Lighting direction affects recognition of untextured faces in photographic positive and negative, *Vision Research* **39**, 4003–4009.
- Nederhouser, M., Yue, X., Mangini, M. C. and Biederman, I. (2007). The deleterious effect of contrast reversal on recognition is unique to faces, not objects, *Vision Research* **47**, 2134–2142.
- Osman, E., Pearce, A. R., Jüttner, M. and Rentschler, I. (2000). Reconstructing mental object representations: a machine vision approach to human visual recognition, *Spatial Vision* **13**, 277–286.
- Palmeri, T. J., Wong, A. C.-N. and Gauthier, I. (2004). Computational approaches to the development of visual expertise, *Trends Cognit. Sci.* **8**, 378–386.
- Phillips, R. J. (1972). Why are faces hard to recognize in photographic negative? *Percept. Psychophys.* **12**, 425–426.
- Poggio, T., Fahle, M. and Edelman, S. (1992). Fast perceptual learning in visual hyperacuity, *Science* **256**, 1018–1021.
- Rentschler, I., Gschwind, M., Brettel, H., Osman, E. and Caelli, T. (2008). Structural and view-specific representations for the categorization of three-dimensional object, *Vision Research* **48**, 2501–2508.
- Rentschler, I. and Jüttner, M. (2007). Mirror-image relations in category learning, *Visual Cognit.* **15**, 211–237.
- Rentschler, I., Jüttner, M. and Caelli, T. (1994). Probabilistic analysis of human supervised learning and classification, *Vision Research* **34**, 669–687.
- Shepard, R. N. and Metzler, J. (1971). Mental rotation of three-dimensional objects, *Science* **171**, 701–703.
- Subramaniam, S. and Biederman, I. (1997). Does contrast-reversal affect object recognition? *Investigat. Ophthalmol. Visual Sci.* **38**, 998.
- Tagscherer, M. (2001). Dynamische Neuronale Netzarchitektur für Kontinuierliches Lernen, *PhD Dissertation*, TU Chemnitz, <http://archiv.tu-chemnitz.de/pub/2001/0072/>
- Tagscherer, M. and Protzel, P. (1998). Adaptive input-space clustering for continuous learning tasks, in: *Proc. Artificial Intell. — FNS'98*, Munich, Germany, pp. 352–358.

- Thoma, V., Hummel, J. E. and Davidoff, J. (2004). Evidence for holistic representations of ignored images and analytic representations of attended images, *J. Exper. Psychol.* **30**, 257–267.
- Vetter, T., Poggio, T. and Bülthoff, H. H. (1994). The importance of symmetry and virtual views in three-dimensional object recognition, *Curr. Biol.* **4**, 18–23.
- Vuong, Q., Peissig, J., Harrison, M. and Tarr, M. (2005). The role of surface pigmentation for recognition revealed by contrast reversal in faces and Greebles, *Vision Research* **45**, 1213–1223.
- Weyl, H. (1952). *Symmetry*. Princeton University Press, Princeton, USA.
- Zangaladze, A., Epstein, C. M., Grafton, S. and Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation, *Nature* **401**, 587–590.