Exp Brain Res (2005) 166: 481–488
DOI 10.1007/s00221-005-2388-5

RESEARCH ARTICLE

Achille Pasqualotto · Ciara M. Finucane
Fiona N. Newell

Visual and haptic representations of scenes are updated with observer movement

Received: 20 July 2004 / Accepted: 21 February 2005 / Published online: 21 July 2005

Abstract Scene recognition has been found to be sensitive to the orientation of the scene with respect to the stationary observer. Recent studies have shown, however, that observer movement can compensate for changes in visual scene orientation, through a process of spatial updating. Here we investigated whether spatial updating in scene recognition is affected by the encoding or learning modality by examining whether observer movement can also compensate for orientation changes in haptic scene recognition. In experiment 1, we replicated previously reported effects of observer movement on visual scene recognition. In experiment 2, we used the same apparatus as in experiment 1 but here participants were required to learn and recognize the scenes using touch alone. We found a cost in recognition performance with changes in scene orientation relative to the stationary observer. However, when participants could move around the scene to recognize the new orientation, then this cost in recognition performance disappeared. Thus, we found that spatial updating applies to recognition in both the visual and haptic modalities, both of which intrinsically encode the spatial properties of a scene.

Keywords Scene perception · Vision · Haptics · Orientation-dependency · Egocentric and body-centred representations · Spatial updating

Introduction

Many studies on visual perception have found that the recognition of both objects and scenes is dependent on the view of the stimulus with respect to a stationary observer. Accordingly, the recognition of a scene from a novel viewpoint is less efficient than the recognition of a scene from a familiar or learned viewpoint (Diwadkar and McNamara 1997; Christou and Bülthoff 1999; Nakatani et al. 2002). However, our phenomenal experience of object or scene recognition in the real world is not that it is error-prone but, rather, that it is remarkably efficient. This is especially evident as we move around our world: while moving around a scene we would expect that the consequent changes in viewpoint of that surrounding scene and its constituent objects would be detrimental to visual recognition. Yet, this is clearly not the case. The question many researchers have posed, therefore, is why the effects of viewpoint changes are not obvious during real world perception.

To that end, Simons and colleagues embarked on a series of studies to test whether or not observer movement could compensate for the consequent change in viewpoint of an object (Simons et al. 2002) or scene (Simons and Wang 1998; Wang and Simons 1999) on recognition performance within the visual domain. Based on the literature from spatial updating in navigation studies (e.g. Rieser 1989; Easton and Sholl 1995), Simons and colleagues reasoned that observer motion should facilitate the updating of an egocentric representation of a scene in memory. Moreover, Simons and Wang (1998) argued that real world perception rarely involves the rotation of a scene relative to a passive or a stationary observer.

In their initial study, Simons and Wang (1998) tested the effects of changes in viewpoint on visual scene recognition under two conditions: passive viewing and active viewing (i.e. walking around) of a scene. They found that when the orientation of the scene changed relative to a passive or stationary observer, there was a cost in scene recognition performance (as reported previously by Diwadkar and McNamara 1997 among others). However, when the participant moved around to view the scene from a novel viewpoint, there was no cost in performance. In other words, when the change of view was caused by the scene rotation there was a cost in

A. Pasqualotto · C. M. Finucane · F. N. Newell (✉)
Department of Psychology, Institute of Neuroscience,
University of Dublin, Trinity College, Dublin 2, Ireland
E-mail: fiona.newell@tcd.ie
Tel.: +353-1-6083914
Fax: +353-1-6712006

recognition performance, but if the same extent of view change (47°) was due to a change in the observer's viewpoint then this cost disappeared. It seemed, therefore, that ''retinal projections'' (sic. Simons and Wang 1998) alone do not account for performance in scene recognition since performance should have been the same across the conditions where the observer viewed the same retinal projection between study and test. Instead, Simons and Wang argued that extra-retinal information is used to update egocentric representations of scenes in memory (Simons et al. 2002).

Many studies have tried to elucidate the mechanism behind spatial updating in scene perception. Generally, the sources of information used for updating during observer movement fall into two broad categories; external or internal cues (also called allothetic or idiothetic cues for navigation) (see, e.g. Avraamides 2003). External cues can be the spatial relationship between, for example, visual or auditory events in the environment and the effect of self-motion on these relationships. Internal cues are those changes that are generated as a consequence of self-movement such as changes in proprioception, optic flow or vestibular information. In a series of experiments Simons and Wang (1998) and Wang and Simons (1999) tried to elucidate the mechanism behind view independent performance in visual scene recognition when there was an active change in scene view. The findings from their studies suggest that internal cues, i.e. a mix of proprioceptive, vestibular and optic flow information, are necessary for scene updating to occur. When all these cues are absent (as in Simons and Wang 1998, experiment 3) then there is a cost due to scene orientation change regardless of the observer viewpoint change. Other studies based on human navigation have also suggested similar updating mechanisms during observer movement (e.g. Loomis et al. 1993; Rieser et al. 1994; Israel et al. 1999; Wang 1999; Wang and Spelke 2000).

Most of the research to date on spatial updating of scenes has, however, concentrated on the visual representation of scenes and largely ignored updating as a consequence of encoding information in other sensory modalities. Our aim here was to investigate if the encoding modality affected spatial updating or whether scene representations function equivalently irrespective of modality for the purpose of spatial updating. More specifically, we investigated whether haptic encoding and learning of scenes also benefits from spatial updating as visual encoding.

Some studies have found evidence that spatial perception is dependent on the encoding modality (Klatzky et al. 2002, 2003). For example, Klatzky et al. (2002) found that there was a performance cost when participants learned the position of objects specified through auditory spatial language compared with when the object positions were learned through vision or spatial audition. Moreover, Klatzky et al. (2003) reported that spatial updating was worse for spatial representations determined by language than representations learned either visually or through 3-D audition. Klatzky and colleagues argued that those sensory modalities that intrinsically encode spatial information (such as vision and 3-D audition as in their experiments) afford automatic updating spatial updating, whereas spatial information from language requires costly indirect processing.

In order for spatial updating to occur within the tactile domain, therefore, this modality should be able to directly encode and represent the spatial properties of a scene, at least in a similar functional manner to visual or auditory perception. Some studies have, however, suggested that the perception of tactile spatial layout in the absence of vision is distorted or poor (Worchel 1951; Hollins and Kelley 1988; Kappers and Koenderink 1999; but see Loomis et al. 1993). Lederman et al. (1985) suggest that the movement of the hand during tactile encoding may affect spatial distortion. Moreover, Kappers and Koenderink found that haptic space was increasingly distorted with increasing distance from the body midline. However, Millar and Al-Attar (2004) argue that it is the absence of an informative external reference frame during tactile spatial perception which results in relatively poor spatial encoding and consequent representation of haptic space (see also Rossano and Warren 1989 for effects of spatial misalignment in map reading). In their study, participants were required to learn and recall the location of landmarks on a tactile map. Millar and Al-Attar found that performance on this task improved when participants were provided with an external tactile reference frame during learning. Thus, the external cue aided tactile representation of spatial layout in a similar manner to external cues affecting the representation of spatial layout in visual memory (e.g. Christou et al. 2003), perhaps by allowing for a more object-based representation of the tactile scene (Klatzky 1999).

Some other recent studies have suggested that vision and haptics are functionally equivalent for the purpose of spatial perception. For example, in an investigation of the processes involved in scene recognition across vision and haptics, Newell et al. (2005) found that scene orientation resulted in a cost in recognition performance for both sensory modalities. Newell et al. (2005) argued that both the visual and haptic modalities represent spatial information in terms of egocentric or body-centred representations. Furthermore, Millar and Al-Attar (2002) found that the Müller–Lyer illusion, which typically distorts judgements of distance, occurs for both vision and touch and is eliminated by the same factors in both senses. Together, these studies suggest that touch can encode and represent the layout of objects within peripersonal space in a manner similar to the visual system.

The aim of our study, therefore, was to determine whether spatial updating was affected by the encoding modality, by comparing scene recognition performance across vision and touch. We first tested whether the findings reported by Simons and Wang in visual scene recognition generalize to our, more complex, scenes. We then investigated whether spatial updating effects extend to the haptic modality, using the same apparatus and stimuli as in our first experiment. Specifically, we were interested in determining whether haptic scene recognition would also benefit from observer movement in the same way as visual scene recognition such that changes in viewpoint are compensated by movement.

Experiment 1

In this experiment we tested whether the findings of Simons and Wang (1998) generalize to other situations involving a change in experimental procedure and stimuli (see also Wang and Simons 1999). Here we used a scene apparatus which we had previously used in our other studies on scene perception (Newell 2004; Newell et al. 2005). Our apparatus differed from that used by Simons and Wang in that we used more objects (seven rather than five), all objects were of roughly the same size and all were made from the same material. We needed to control for object size and material in order to allow us to conduct the subsequent haptic scene recognition experiment. In all other ways our procedure and design was the same as that reported by Simons and Wang.

If the effects reported by Simons and Wang are robust enough to generalize across stimuli and experimental situations, as seems to be the case (e.g. Wang and Simons 1999; Simons et al. 2002), then we expected to replicate their findings. That is, we expected to find that the observer's movement would compensate for the cost in recognition performance due to changes in the orientation of a scene.

Materials and methods

Participants

Fifteen undergraduate students from Trinity College, Dublin, participated in this experiment for research credits (seven males, eight females). Their ages ranged from 18 to 31 years. All had normal or corrected-to-normal vision. All participants gave written consent to take part in the experiment. [1]

Apparatus

The stimulus set of objects included 15 wooden shapes of common animals (e.g. cat, rabbit, goose, cockerel, horse, etc.). In each trial seven objects were randomly chosen from the set and placed on a rotatable circular platform positioned in the centre of a fixed round table. The platform had 19 sunken position markers in which the individual objects could be randomly placed. Each position marker was equidistant from any of its neighbouring markers by a distance of 7 cm. The diameter of the platform measured 54 cm, and the diameter of the supporting table was 80 cm.

Both the table and platform were placed behind an opaque curtain that extended from the ceiling to just beneath the table and completely surrounded the table. We placed two small viewing windows on the curtain that could be opened: each window measured 6 cm wide×9 cm high and was positioned at a height of about 1 m from the ground. The viewing windows were approximately 90 cm from the centre of the table and 60° apart. A chair was placed in front of each viewing window. See Fig. 1 for an illustration of the apparatus used.

The object stimuli were toy wooden shapes of animals and figures painted white and each was placed on a plastic stand that could be inserted in any position marker on the platform. All of these objects were 1 cm wide, varied in height from the base to the top of the object between 6 and 8 cm and varied in length from between 3.5 and 5.5 cm. These stimuli were specifically chosen because they were distinguishable using either vision or touch alone. Prior to this experiment we ran a pilot study with a different group of participants to help us to determine which objects, from a set of 30, were most discriminable from each other using touch alone. We then chose the 15 most distinctive object stimuli from the set. In both our experiments, each scene consisted of objects that were easily distinguishable from each other in that we avoided placing two objects onto a scene stimulus that were similar (e.g. a cat and a dog were never used in either the same visual or haptic scene). A typical scene, therefore, might consist of the following shapes; dog, rabbit, cow, cockerel, duck, pig and sheep. The other objects were cat, bison, farmer, goat, goose, hen, horse and pigeon.

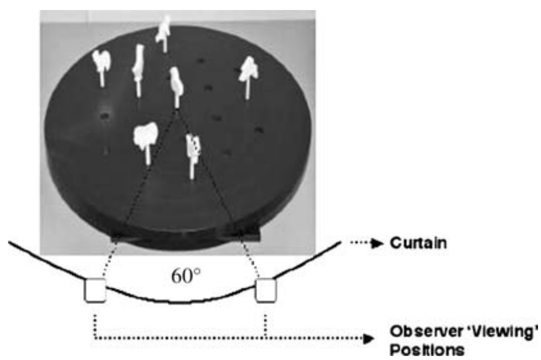

Fig. 1 An illustration of our experimental apparatus used in both experiments 1 and 2. Note that the egocentric projection of the scene is the same in the conditions where both the observer and the scene were unchanged (Ou–Su) and when both the observer and the scene were changed (Oc–Sc). See text for details

[1]All the studies reported here were approved by the Trinity College Department of Psychology Ethics Committee.

Design

The experiment was based on a 2×2 within-subjects design with scene orientation (0° or 45°), and observer viewpoint (unchanged, changed) as factors. Thus there were four experimental conditions; two where the observer remained stationary (i.e. unchanged viewing position) and the scene was either rotated or not, and two where the observer moved (i.e. changed viewing position) and the scene was again rotated or not. We coded these four conditions as follows: scene orientation unchanged [Su]; scene orientation changed [Sc]; observer viewpoint unchanged [Ou]; and observer viewpoint changed [Oc]. There was a total of 80 trials, with 20 trials in each condition. Trials were partially blocked in that every five trials were from the same condition, and every 20 trials contained four blocks of five trials across all conditions (see, Wang and Simons 1999 for a similar design). We blocked the trials so that the participants were clear about their task and our blocks were small in order to avoid block order effects. The order of the blocks was randomized across participants.

Procedure

During this experiment, a trial consisted of a learning phase followed by a test phase. Specifically, participants were required to first learn a scene and then to identify the object within that scene which was subsequently moved between learning and test. Prior to the experiment, participants were presented with two randomly chosen practice trials in order that they familiarize themselves with the task. The experimental trials followed the practice trials after a brief pause (to allow the participant to ask any questions).

During a trial, seven stimuli were randomly chosen from the full set and placed in random positions on the platform. Participants were required to learn each scene and were subsequently tested on their recognition of the scene. During learning the viewing window was opened by the experimenter and the participant viewed the scene for 3 s after which the viewing window was closed. The experimenter then displaced one of the seven objects in the scene. After an inter-stimulus interval of 20s,[2] the participant viewed the scene again and was required to identify the displaced object. There was no time limit for responding, although participants were requested to respond as fast and as accurately as possible and were probed for an answer after a lapse of about 1 min if none was forthcoming. The number of correct responses made was recorded. The experiment took about 90 min to complete.

According to the conditions of the experiment the scene was rotated by 45° in a clockwise direction, or it remained in the same orientation as learning. In turn, for half of these trials the participants moved to a new position (i.e. moved from one chair to another) or they remained in the same position. During the trials, where the viewing position was unchanged, the participants were requested to stand up, take a few steps, and sit down again on the same chair during the inter-stimulus period. This movement acted as a control for the body movement in the condition where there was a change in viewpoint. Participants were informed when the scene would change orientation or not prior to each trial. Performance was measured in terms of correct responses.

Results

The mean percentage of correct responses made during each condition is presented in Fig. 2. As can be observed in Fig. 2, performance on this task was well above chance in all conditions (chance level was 1/7 or 14.3% correct). We conducted a two-way, within-subjects ANOVA on the correct responses, using scene orientation (Su, Sc) and observer viewpoint (Ou, Oc) as factors. We found a main effect of scene orientation [$F(1,14) = 13.51$, $P < 0.01$] and no significant effect due to observer viewpoint [$F(1,14) < 1$]. As expected, we found a significant interaction between observer viewpoint and scene orientation [$F(1,14) = 20.13$, $P < 0.001$].

As shown in Fig. 2, there was a cost in performance when the scene was rotated relative to a stationary observer. However, when the observer could move to a new viewpoint, this cost in scene orientation disappeared. We

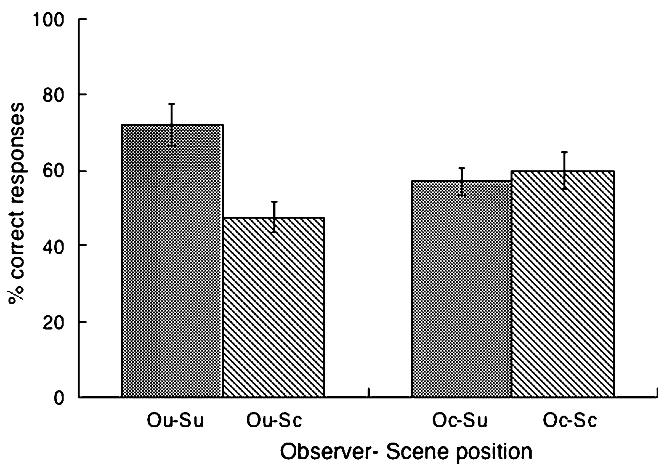

Fig. 2 Graph showing the mean percentage correct visual recognition responses to changes in scene orientation as a function of the observer's viewpoint in experiment 1. The four conditions in this experiment refer to whether or not the observer remained stationary (Ou) or changed his or her position (Oc) and whether or not the scene remained stationary (Su) or was rotated (Sc). The *error bars* represent ±1 SEM

[2]Although this ISI is longer than in previous studies (e.g. 7 s in Wang and Simon's 1999 study) this duration was necessary for practical reasons but mainly to allow for the same ISI across experiments 1 and 2. We found no evidence that this increase in ISI between learning and testing affected the spatial updating (see, Hollins and Kelley 1988) although that is not to say that any further increase might indeed disrupt updating.

conducted a post hoc Newman–Keuls analysis on the interaction between the factors. This analysis confirmed that when the observer's viewpoint remained stationary, a change in scene orientation resulted in significantly more errors than when the scene remained in the same orientation ($P < 0.05$). However, there was no effect of scene orientation on recognition performance when the observer changed his or her viewpoint.

Discussion

In summary, as in the previous studies reported by Simons and Wang (1998) and Wang and Simons (1999), we found a benefit with observer movement on the recognition of rotated scenes. Therefore, the results reported by Simons and Wang replicate across situations where the stimuli are less discriminable and the task more difficult (i.e. with more stimuli).

The overall interaction between observer's viewpoint and scene orientation suggests that participants built up a representation of the scene that could be updated when they moved to another position. This updating did not occur, however, when the scene was passively rotated in front of them. It is important to note that this updating was not due to motor cues, since participants were required to move in all experimental conditions.

One interesting observation can be made between the findings of Simons and Wang in their studies and our findings here. In the conditions where the observer moved, Simons and Wang sometimes found better recognition performance when the scene remained unchanged than when the scene changed. Furthermore, they also reported that performance to an unchanged scene was the same whether the observer moved or remained stationary (see, e.g. experiment 1, Simons and Wang 1998; Wang and Simons 1999). In our experiment here, however, performance between the changed and unchanged scene orientation with observer movement (Oc–Sc and Oc–Su) was not different. Moreover, performance in the Ou–Su condition was better than all other conditions. We are not completely clear why this difference between the studies has occurred although we can offer one suggestion. In their experiment 2, Simons and Wang (1998) found that the elimination of external visual cues was sufficient to reduce performance when the observer moved and the scene remained stationary than when the observer and scene remained in the same position. Moreover, when even a single external cue was available, such as a handle which moved during scene rotation (Wang and Simons 1999, experiment 1) this seemed to be sufficient to render performance in the condition where the observer moved and the scene remained stationary the same as when neither changed position. We think it might be likely that our experimental apparatus effectively removed external visual information which might aid in updating: that is, we surrounded our stimulus scene with a non-patterned curtain which may have effectively eliminated any external cues which may have been used for updating. In any case, the issue of the exact nature of the information mediating updating clearly warrants further investigation.

Our results, together with those previously reported in the literature, suggest that in the real world people recognize scenes and objects more efficiently when they move through the environment. Observer movement, therefore, allows for egocentric representations of scenes in visual memory to be updated (Wang and Spelke 2000). However, vision is not the only sense by which we encode the spatial layout of objects. It therefore remains an empirical question as to whether the haptic modality can benefit from spatial updating with observer movement. We attempted to address this question in the following experiment.

Experiment 2

This experiment was based on experiment 1 except that our participants could not view the stimuli or apparatus either before or during the experiment. Here the participants were required to explore the scene using touch alone. We predicted that if the spatial layout of scenes is encoded by touch in a functionally equivalent manner to vision, then scene updating should be independent of encoding modality. Consequently we expected to find that for haptic scene recognition the observer's movement should compensate for changes in scene orientation.

Materials and methods

Participants

Twenty-eight undergraduate students from Trinity College, Dublin, participated in this experiment for research credits (12 males, 16 females). Their ages ranged from 18 to 39 years. None reported any tactile impairment, and all were right-handed. All participants gave written consent to take part in the study.

Apparatus

See experiment 1 for a description of the apparatus and stimuli used. In this experiment, participants placed both their hands underneath the curtain to feel the scene. Before the experiment, the scene platform was moved closer to the participants if they were unable to comfortably reach the farthest point with their hands. The midpoint of the scene was always equidistant from each of the "viewing" [3] positions and the viewing positions were adjusted accordingly so that they were always

[3] Although the terms "viewing" and "viewpoint" are more synonymous with visual processing, we prefer to use the same terms of reference across experiments to minimize confusion.

placed at an angle of 45° apart from the centre of the scene. Participants used both their hands to explore a scene. The stimuli used in this experiment were the same as those used in experiment 1.

Design

The design was based on that described in experiment 1 with one exception: here we used a between-subjects design where one group of participants did not change their viewpoint between learning and test and the other group always changed viewpoint. A between-subjects design was used in order to reduce the time it took for the participant to complete the experiment. Participants were randomly assigned to each group. Each participant performed 40 trials in total, 20 during which no change in scene orientation occurred and another 20 where there was a change of 45° in scene orientation.

Procedure

The procedure largely followed that outlined in experiment 1 except that in this experiment participants had 60 s to learn the scene using touch alone. Learning time began as soon as any of the participant's hands touched an object stimulus in the scene. Participants were given unlimited time to respond although they were instructed to respond as fast as possible and were probed for an answer after a lapse of about 2 min. The participants who did not change their viewpoint before the test were required to stand up and walk a few steps after learning a scene. In all other ways the procedure followed that outlined in experiment 1. The experiment took about 60 min to complete.

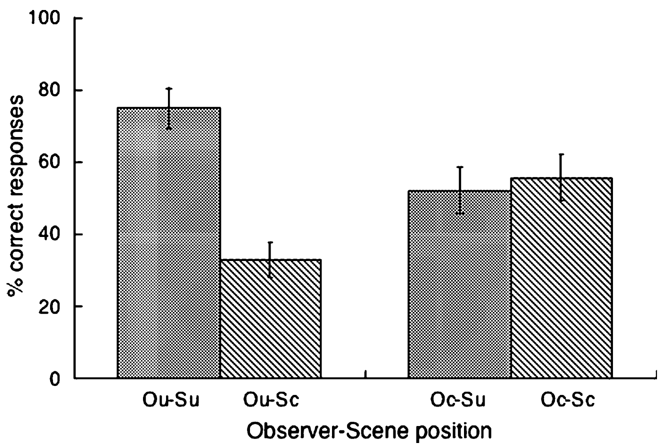

Fig. 3 Graph showing the mean percentage correct haptic recognition responses to changes in scene orientation as a function of the observer's position in experiment 1. The four conditions in this experiment refer to whether or not the observer remained stationary (*Ou*) or changed his or her position (*Oc*) and whether or not the scene remained stationary (*Su*) or was rotated (*Sc*). The *error bars* represent ±1 SEM

Results

The percentage of correct responses made across all conditions is shown in Fig. 3. As in experiment 1, performance was well above chance in all conditions (again chance level was 1/7 or 14.3% correct). We conducted a two-way mixed ANOVA on the correct responses with observer viewpoint as a between-subjects factor (Ou or Oc) and scene orientation as a within-subjects factor (Su or Sc). We found a main effect of scene orientation [$F(1,26) = 15.25$, $P < 0.001$], and no effect of observer viewpoint [$F(1,26) < 1$]. Interestingly, we found an interaction between observer's viewpoint and scene orientation [$F(1,26) = 21.42$, $P < 0.001$]. As in experiment 1, there was a cost with changes in scene orientation relative to the stationary observer but no cost in scene orientation with observer viewpoint changes (see Fig. 3).

We performed a post hoc Tukey HSD [4] analysis on the interaction between the factors and found that the recognition of a rotated scene was significantly worse than of a non-rotated scene when the observer remained stationary ($P < 0.05$). On the other hand, when participants moved to another viewpoint position their recognition performance was unaffected by scene orientation.

Discussion

The overall interaction between observer position and scene orientation suggests that the representation of the haptic scene could be updated as the observer moved to another position. This spatial updating with scene orientation did not occur when the observer remained stationary.

Our results for the haptic experiment are very similar to those found in experiment 1 using vision only. One noticeable difference between the experiments is that participants in the haptic task performed slightly worse than those in the visual task (i.e. experiment 1). This difference may be because of a general lack of familiarity with performing a haptic task: in the real world sighted people probably rely more on vision than on touch to explore their world. Furthermore, haptic exploration of space is limited to peripersonal space as opposed to the vast scale of visual space rendering visual scene exploration a more ubiquitous task. Therefore, more mistakes may have been made because the haptic task was unusual for the participants. Notwithstanding the differences in overall error rates, our results suggest that the haptic system updates representations of the world as the observer moves around in a similar manner to the updating of scenes in visual memory.

[4]This experiment was based on a between-subjects design, thus our post hoc analysis was conducted using Tukey HSD analysis rather than a Newman–Keuls as in experiment 1.

General discussion

In our studies, we replicated a previous finding that visual scene representations are updated with observer movement and we report a new finding that scene representations in the tactile domain are also updated with observer movement. For both vision and haptics, there was a cost in scene recognition performance when a scene was rotated relative to a stationary observer. However, this cost disappeared when the observer could move to the new viewpoint of the scene.

The finding that scene recognition was orientation dependent with respect to a stationary observer in both sensory modalities suggests that an egocentric or body-centred representation of the scene was constructed in visual and tactile memory. Other studies have argued that the representation of unfamiliar objects and scenes is egocentric in visual memory (e.g. Diwadkar and McNamara 1997; Newell and Findlay 1997; Tarr and Bülthoff 1998; Wang and Spelke 2000). Furthermore, more recent studies have suggested that haptic scenes are also encoded as body-centred representations, since changes in scene orientation with respect to the observer cause a cost in recognition performance (Newell et al. 2005). Our findings here provide further evidence that both the visual and haptic systems store information about scenes in terms of an egocentric reference frame.

If we accept that spatial information about objects in a scene is similarly represented across the visual and tactile sensory modalities (i.e. as egocentric representations), then it might be that the mechanism of updating egocentric representations is similar across these modalities. Indeed, by applying a similar mechanism to update information represented across vision and haptics, the sharing of information across these sensory systems may be more efficient. However, this raises the obvious question about whether the *same* updating mechanism is applied to information independent of sensory modality. It is conceivable that as the observer moves around the world optic flow, internal proprioceptive and vestibular information are used to update representations across all the sensory modalities together. Furthermore, the spatial relationships between external visual or auditory landmarks may also affect updating in both modalities if external reference frames are shared. In other words, instead of proposing that different information or cues are used by each of the sensory modalities for spatial updating, all available or relevant information may be used interchangeably by the different modalities. In some cases, such as in the case of tactile spatial perception in the absence of vision, some external cues may not be available for updating because they are outside the working range of haptics. However, once useful external cues are added to a haptic scene then efficient spatial perception can occur (Millar and Al-Attar 2004).

On the other hand, as is clear from other studies, spatial updating is affected by the encoding modality suggesting that the fidelity of the represented spatial information also affects updating. For example, some studies have found that the encoding of spatial information from verbal descriptions is not conducive to efficient updating (Klatzky et al. 2002, 2003). Moreover, direct perception of a scene seems to be necessary for updating to occur since imagined scenes are not updated as efficiently as real scenes (Wang 2004). These studies suggest that the encoding of real scenes in a modality that intrinsically encodes spatial distances and location (e.g. vision, audition, and touch) allows for updating to occur as the observer moves around that scene.

The fact that we find the similar spatial updating effects with observer movement on scene recognition might not, however, be a consequence of external or internal information that is shared across modalities to allow for updating. Instead, these similar effects across modalities may be because the representation of the scene itself may be modality independent. For example, it might be that haptic information is recoded into visual coordinates, therefore our data may reflect updating of direct visual and recoded haptic-to-visual representations. This might happen if the participants are visually imagining the haptic scene. In this case we might not expect any difference between the modalities in terms of updating performance although a general reduction in performance for the haptic modality might be expected if there is a cost in recoding to visual imagery. However, our recent research suggests that the haptic and visual representations of scenes are indeed discrete and that the recoding of haptic input to visual representations is not mandatory. For example, Newell et al. (2005) reported poor cross-modal recognition of scenes relative to within-modal recognition. They argued that because of differences between the representations of scene across the sensory modalities, this resulted in more error-prone crossmodal recognition (e.g. see Kappers and Koenderink 1999 for evidence of distorted representations in the haptic domain). Furthermore, we have recently started an investigation of spatial updating in persons without visual experience using the same apparatus as reported here and our data, although preliminary, suggest that updating of haptic scenes occurs as efficiently in blind as well as in sighted individuals (cf. Hollins and Kelley 1988). On the basis of these findings, we would argue that discrete representations are stored across the visual and haptic modalities.

In conclusion, our study provides further evidence that scenes are encoded by both the visual and haptic systems into egocentric representations in sensory memory. As such, any change in scene orientation caused a cost in recognition performance in both vision and haptics. Furthermore, we would argue that these representations, although similar, are discrete and specific to each sensory modality. Finally, our study provides evidence that representations of scenes can be updated in a similar manner in both visual and tactile memory. Whether the same external or internal information is used in a functionally equivalent manner for updating across all modalities is a matter for future research.

Acknowledgements This research was funded by a Higher Education Authority, PRTLI grant awarded to the Institute of Neuroscience, Trinity College Dublin, of which F.N.N. is a member.

References

Avraamides MN (2003) Spatial updating of environments described in texts. Cognit Psychol 47:402–431

Christou CG, Bülthoff HH (1999) View dependence in scene recognition after active learning. Mem Cognit 27:996–1007

Christou CG, Tjan BS, Bülthoff HH (2003) Extrinsic cues aid shape recognition from novel viewpoints. J Vis 3:183–198

Diwadkar VA, McNamara TP (1997) Viewpoint dependence in scene recognition. Psychol Sci 8:302–307

Easton RD, Sholl MJ (1995) Object-array structure, frames of reference, and retrieval of spatial knowledge. J Exp Psychol Learn Mem Cogn 21:483–500

Hollins M, Kelley EK (1988) Spatial updating in blind and sighted people. Percept Psychophys 43:380–388

Israel I, Ventre-Dominey J, Denise P (1999) Vestibular information contributes to update retinotopic maps. Neuroreport 10:3479–3483

Kappers AML, Koenderink JJ (1999) Haptic perception of spatial relations. Perception 28:781–795

Klatzky RL (1999) Path completion after haptic exploration without vision: implications for haptic spatial representations. Percept Psychophys 61:220–235

Klatzky RL, Lippa Y, Loomis JM, Golledge RG (2002) Learning directions of objects specified by vision, spatial audition, or auditory spatial language. Learn Mem 9:364–367

Klatzky RL, Lippa Y, Loomis JM, Golledge RG (2003) Encoding, learning, and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. Exp Brain Res 149:48–61

Lederman SL, Klatzky RL, Barber PO (1985) Spatial and movement-based heuristics for encoding pattern information through touch. J Exp Psychol Gen 114:33–49

Loomis JM, Klatzky RL, Golledge GR, Cicinelli GJ, Pellegrino WJ, Fry AP (1993) Nonvisual navigation by blind and sighted: assessment of path integration ability. J Exp Psychol Gen 122:73–91

Millar S, Al-Attar Z (2002) The Muller–Lyer illusion in touch and vision: implications for multisensory processes. Percept Psychophys 64:353–365

Millar S, Al-Attar Z (2004) External and body-centered frames of reference in spatial memory: evidence from touch. Percept Psychophys 66:51–59

Nakatani C, Pollatsek A, Johnson SH (2002) Viewpoint-dependent recognition of scenes. Q J Exp Psychol A 55:115–139

Newell FN (2004) Crossmodal object recognition. In: Calvert GA, Spence C, Stein BE (eds) The handbook of multisensory processes. MIT Press, Cambridge, Mass., pp 123–140

Newell FN, Findlay JM (1997) The effect of depth rotation on object identification. Perception 26:1231–1257

Newell FN, Woods AT, Mernagh M, Bülthoff HH (2005) Visual, haptic and cross-modal recognition of scenes. Exp Brain Res 161(2):233–242

Rieser JJ (1989) Access to knowledge of spatial structure at novel points of observation. J Exp Psychol Learn Mem Cogn 15:1157–1165

Rieser JJ, Garing AE, Young MF (1994) Imagery, action, and young children's spatial orientation: it's not being there that counts, it's what one has in mind. Child Dev 65:1262–1278

Rossano MJ, Warren DH (1989) Misaligned maps lead to predictable errors. Perception 18:215–229

Simons DJ, Wang RF (1998) Perceiving real-world viewpoint changes. Psychol Sci 9:315–320

Simons DJ, Wang RF, Roddenberry D (2002) Object recognition is mediated by extraretinal information. Percept Psychophys 64:521–530

Tarr MJ, Bülthoff HH (1998) Image-based object recognition in man, monkey and machine. Cognition 67:1–20

Wang RF (1999) Representing a stable environment by egocentric updating and invariant representations. Spat Cogn Comput 1:431–445

Wang RF (2004) Between reality and imagination: when is spatial updating automatic?. Percept Psychophys 66:68–76

Wang RF, Simons DJ (1999) Active and passive scene recognition across views. Cognition 70:191–210

Wang RF, Spelke ES (2000) Updating egocentric representations in human navigation. Cognition 77:215–250

Worchel P (1951) Space perception and orientation in the blind. Psychol Monogr 65:1–28