Available online at www.sciencedirect.com
SCIENCE DIRECT
Cognitive Brain Research 17 (2003) 14–25
COGNITIVE BRAIN RESEARCH
www.elsevier.com/locate/cogbrainres
ELSEVIER

Research report

Tactile–visual cross-modal shape matching: a functional MRI study

Daisuke N. Saito[a,b], Tomohisa Okada[c], Yusuke Morita[a], Yoshiharu Yonekura[b], Norihiro Sadato[c,d,*]

[a]*Department of Physiology, School of Medicine, The University of Tokushima, Tokushima, Japan*
[b]*Biomedical Imaging Research Center, Fukui Medical University, Fukui, Japan*
[c]*National Institute for Physiological Sciences, Okazaki, Japan*
[d]*JST (Japan Science and Technology Corporation)/RISTEX (Research Institute of Science and Technology for Society), Kawaguchi, Japan*

Accepted 23 January 2003

Abstract

The process and location of integration of information from different sensory modalities remains controversial. We used functional MRI to investigate the neural representation of cross-modal matching between tactile and visual shape information in eleven normal volunteers. During the scan, patterns of 2D shapes were presented both tactually and visually, simultaneously. Four different matching tasks were performed: tactile–tactile with eyes closed (TT), tactile–tactile with visual input (TTv), visual–visual with tactile input (VVt), and tactile–visual (TV). The TT task activated the contralateral primary sensorimotor area, and the postcentral gyrus, superior parietal lobules, anterior portion of the intraparietal sulcus, secondary somatosensory cortex, thalamus, dorsal premotor area, cerebellum, and supplementary motor area bilaterally, without occipital involvement. Visual matching activated the primary visual cortex and the lingual and fusiform gyri bilaterally. A cross-modal area was identified by subtracting TTv images from TV images, subtracting VVt images from TV images, and then determining common active areas. There was one discrete area that was active bilaterally; the posterior intraparietal sulcus close to the parieto-occipital sulcus. These data suggest that shape information from different sensory modalities may be integrated in the posterior intraparietal sulcus during tactile–visual matching tasks.

Theme: Motor systems and sensorimotor integration

Topic: Cortex

Keywords: Tactile; Visual; Cross-modal shape matching; Integration; Intraparietal sulcus

---

1. Introduction

Many human behaviours necessitate the integration of information conveyed through anatomically distinct sensory pathways. For example, object recognition is based on the visual and/or tactile extraction of basic features such as contours and their spatial arrangement, and may ultimately involve tactile–visual integration [1]. It has been assumed that early information processing in the sensory areas is strictly modality specific, and integrated only after subsequent transmission to other brain areas (i.e. association cortex) [11]. The process and location of tactile–visual integration, however, remains controversial.

Previous studies of tactile–visual cross-modal performance have focused on the relation between the sensory-specific systems and so-called polysensory areas, defined as areas activated by stimuli from more than one sensory modality. It has been reasoned that the communication between unimodal representations requires some pathways of interaction, which should involve the sensory-convergent polysensory areas [8]. Early results of the effects of brain lesions on cross-modal performance suggested the involvement of these areas. It was reported that there were impairments in tactile–visual matching abilities after posterior temporal cortex removal [41] or lesions of the arcuate sulcus cortex [34] in nonhuman primates. These

---

*Corresponding author. Section of Cerebral Integration, Department of Cerebral Research, National Institute for Physiological Sciences, Myodaiji, Okazaki, Aichi, 444-8585, Japan. Tel.: +81-564-55-7874; fax: +81-564-52-7913.

E-mail address: sadato@nips.ac.jp (N. Sadato).

0926-6410/03/$ – see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0926-6410(03)00076-4

lesion studies, however, failed to demonstrate a specific and consistent role of the polysensory areas in cross-modal matching [19]. This raises the possibility that simple polymodal convergence is not sufficient for cross-modal matching [18].

To investigate the neural substrates for tactile–visual cross-modal matching, we utilised functional MRI (fMRI) during tasks involving tactile–visual matching of two-dimensional (2D) shapes. We hypothesised that the neural substrates for tactile and visual shape matching are more active during tasks requiring matching of information coming from two different sensory modalities than from within either modality. The location of such a cross-modal area is likely within or adjacent to the polymodal areas that are activated by multiple sensory modalities. The results of the present study indicate that the bilateral posterior portion of the intraparietal sulcus (PIP), known to be a polysensory area, integrates shape information from different modalities during tactile–visual matching.

2. Materials and methods

2.1. Subjects

Eleven healthy volunteers (six men, five women, mean age 29.2±6.5 years) participated in this study. Ten subjects were right handed and one was left handed by Edinburgh's handedness inventory [32]. All subjects were unfamiliar with Mah-Jong, a traditional Chinese tile game, the tiles of which were used as visual and tactile stimuli. There was no history of neurological or psychiatric illness in any of the subjects. The protocol was approved by the ethical committee of Fukui Medical University, and all subjects gave their written informed consent for the study.

2.2. MRI

A time-course series of 46 volumes was acquired using T2*-weighted, gradient echo, echo planar imaging (EPI) sequences using a 3.0 Tesla MR imager (VP, General Electric, Milwaukee, WI, USA). The raw data were transferred to a parallel supercomputer (ORIGIN2000, SGI, Mountain View, CA, USA) to reconstruct the consecutive 2D images using an algorithm of 2D fast Fourier transformation (General Electric). Each volume consisted of 34 slices, each 3.5-mm thick, with a 0.5-mm gap, to cover the entire cerebral and cerebellar cortex. The time interval between two successive acquisitions of the same image was 3000 ms, and the echo time was 30 ms. The field of view (FOV) was 19 cm. The digital in-plane resolution was 64×64 pixels with a pixel dimension of 2.97×2.97 mm. The magnetic shim was optimised such that a true in-plane resolution of 2.97×2.97 mm was realised. Tight but comfortable foam padding was placed around the subject's head to minimise head motion.

For anatomical reference, T2-weighted fast spin echo images were obtained from each subject with location variables identical to those of the EPIs. In addition, high-resolution whole-brain MR images were obtained using a conventional T2-weighted, fast spin echo sequence. A total of 112 transaxial images were obtained. The in-plane matrix size was 256×256, slice thickness was 1.5 mm, and pixel size was 0.859×0.859 mm.

2.3. Shape-matching tasks

For the tactile–tactile, visual–visual, or tactile–visual matching tasks, we used patterns of Mah-Jong tiles. Mah-Jong is a Chinese game similar to card games, involving plastic tiles with various marks carved on one side. For tactile stimuli, we made 80 pair-wise blocks of plastic Mah-Jong tiles (Taiyo, Wakayama, Japan, 26×18.5×11.6 mm) by gluing two circular or stick patterned tiles side-by-side. Forty blocks consisted of two tiles with identical patterns and the remaining 40 blocks consisted of two tiles of two different patterns (Fig. 1). Identical visual blocks of patterns were prepared as digital data.

An fMRI session consisted of two rest and two task periods, each 30 s in duration, with the rest and task periods alternating. The subjects performed four different tasks: a tactile–tactile matching task with no visual input (TT), a tactile–tactile matching task with visual input (TTv), a visual–visual matching task with tactile input (VVt), and a tactile–visual matching task (TV). One task was performed in each session. Each task session was repeated twice, and thus, eight sessions were completed by each subject. The presentation order of the eight sessions was counterbalanced across subjects in pseudorandom order. Prior to the fMRI session, the subjects were trained for the three tactile discrimination tasks until their performance exceeded 60% accuracy.

For the TT task, subjects were asked to place their right hand in a supine position. Their left hand was placed on the button box, which was connected to a microcomputer for recording their response. The subjects closed their eyes throughout the session. During the 30-s rest periods, the experimenter touched the subject's foot every 6 s to signal the subject to push buttons with the left index finger and the left middle finger alternately. During the task periods, a block (Fig. 1) was manually placed on the subject's right palm every 6 s. The blocks were placed so that the top of the patterns was toward the fingers. The subjects were required to explore the surface of the block with the right thumb for 4 s. When the experimenter touched the subject's foot, the subject responded by pushing a button with the left index finger if the pair-wise patterns were the same, or with the middle finger if the patterns were different. Then the subject dropped the block. Each task period contained five trials of matching tasks, resulting in a total of ten trials per session.

For the TTv task, the procedure was identical to the TT

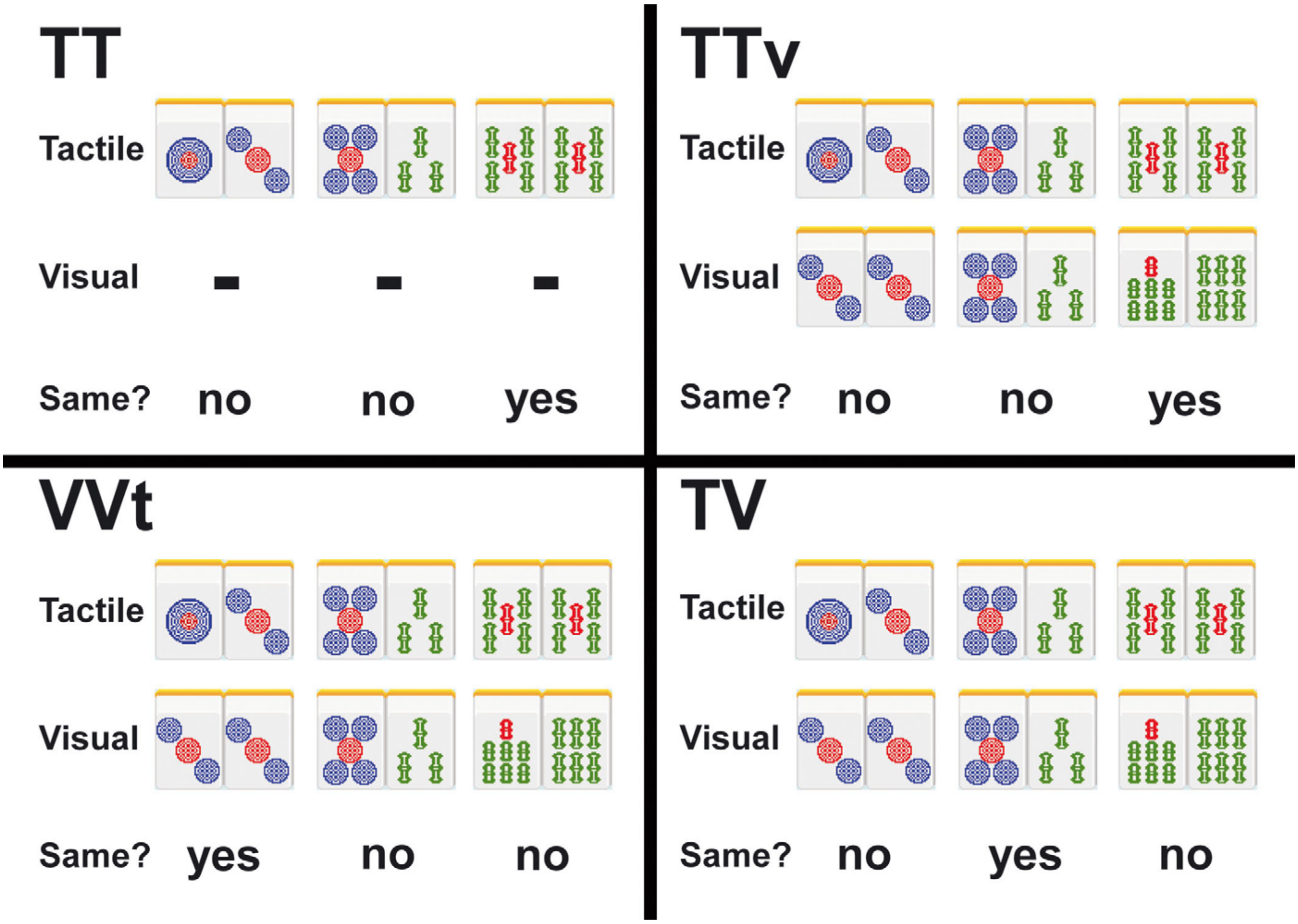

Fig. 1. Task design with correct responses. Tactile and visual stimuli of blocks of Mah-Jong tile patterns were presented simultaneously. During tactile–tactile matching without visual input (TT), the subjects closed their eyes. A block of two joined Mah-Jong tiles was placed in their right palm to feel with the right thumb (upper left). If the side-by-side patterns were the same, then the correct response was 'yes', signalled by pressing a button with the left index finger. If the patterns were different, then the correct response was 'no', signalled by pressing a different button with the left middle finger. Similarly, for tactile–tactile matching with visual input (TTv), the correct response was based on the tactile stimuli (upper right). For visual–visual matching with tactile input (VVt), the correct response was based on the visual stimuli (lower left). For tactile–visual matching (TV), the comparison was made between both in the pair of patterns of tactile and visual stimuli (lower right): if the first tile pattern of the tactile block matched that of the visual block and the second tile pattern of the tactile block matched that of the visual block, then the correct response was 'yes', otherwise 'no'.

task except that additional visual stimuli were presented simultaneously. The visual stimulation was projected using a LCD projector (ELP-7200L, Epson, Tokyo, Japan) connected to a personal computer (Endeavor Pro-600L, Epson), which generated visual stimuli using Presentation software (Neurobehavioral Systems, CA, USA), onto a half-transparent screen hung approximately 0.6 m from the subject's eye. The screen was viewed by the subjects through a mirror. It was confirmed that the subjects were not able to see their right hand during the task. During rest periods, the subjects were required to fixate on a set of cross-hairs on the screen. The experimenter touched the subject's foot every 6 s to signal the subject to push buttons with the left index finger and the left middle finger alternately. During task periods, the tactile and visual stimuli were presented simultaneously for 4 s, followed by a response cued by a touch on the foot. The visual stimuli were pair-wise blocks of Mah-Jong patterns as were the tactile stimuli, but not necessarily of the same patterns. The visual angle of each block pattern was 19°. The subject responded by pushing a button with the left index finger if the tactile pair-wise patterns were the same, or with the middle finger if the patterns were different, irrespective of the visual stimuli. Through the task periods, the subjects were required to keep their eyes fixed on the cross-hairs superimposed on the visual stimuli.

For the VVt task, the procedure was identical to the TTv task except that the response was based on the visual stimuli, irrespective of the tactile stimuli.

Finally, for the TV task, the procedure was identical to the TTv and VVt tasks except that the response was based on both the tactile and visual stimuli. The subject responded by pushing a button with the left index finger if the patterns on both tiles of the tactile and visual blocks matched across sensory modalities (i.e. the first tactile tile matched the first visual tile and the second tactile tile matched the second visual tile), or with the middle finger if either or both of the patterns were different.

2.4. Data analysis

The first six volumes of each fMRI session were discarded to allow for stabilisation of the magnetisation, and the remaining 40 volumes per session, a total of 320 volumes per subject, were used for analysis. The data were analysed using statistical parametric mapping (SPM99, Wellcome Department of Cognitive Neurology, London, UK) and implemented in MATLAB (Mathworks, Sherborn, MA, USA) [12–14]. Following realignment, all images were coregistered to the high resolution, 3D, T2-weighted MRI, in reference to the anatomical MRI with T2-weighted spin echo sequences from locations identical to those of the fMRI images. The parameters for affine and nonlinear transformation into a template of T2-weighted images that was already fit to a standard stereotaxic space (MNI template) [9] were estimated with the high-resolution, 3D, T2-weighted MR images by least square means [13,14]. The parameters were applied to the coregistered fMRI data. The anatomically normalised fMRI data were filtered using a Gaussian kernel of 8 mm (full width at half maximum) in the x, y and z axes.

2.5. Statistical analysis

Statistical analysis in the present study was conducted at two levels. First, individual task-related activation was evaluated. Second, the summary data of each individual were incorporated into the second level analysis using a random effect model [16] to make inferences at a population level.

2.6. Individual analysis

The signal was proportionally scaled by setting the whole-brain mean value to 100 arbitrary units. The signal time course for each subject was modelled using a box-car function convolved with a hemodynamic response function, session effect, and high-pass filtering (120 s). The explanatory variables were centred at 0. To test hypotheses about regionally specific condition effects, the estimates for each model parameters were compared with the linear contrasts. First, we delineated the areas that were active during the TT, TTv, VVt, and TV tasks compared with those active during the rest periods of the same session. The images taken during the TT task were subtracted from those taken during the VVt task to delineate the neural substrates involved in visual shape matching. The resulting set of voxel values for each contrast constituted a statistical parametric map (SPM) of the t statistic (SPM$\{t\}$). The SPM$\{t\}$ was transformed to the unit normal distribution (SPM$\{Z\}$). The threshold for SPM$\{Z\}$ was set at $Z>3.09$ and $P<0.05$ with a correction for multiple comparisons at the cluster level for the entire brain [15]. The areas active during visual–visual matching were calculated by subtracting the images taken during the TT task from those taken during the VVt task. Visual–visual discrimination is normally performed extremely well, so well that the discrimination process may be performed even during the TTv task, thus, subtraction of images taken during the TTv task from those taken during the VVt task may not reveal task-related effects.

To identify cross-modal areas that were more active when information coming from two different sensory modalities was matched than when information from within either modality was matched, first active areas remaining after subtracting VVt from TV were defined at a statistical threshold of $P<0.05$ corrected for multiple comparisons at the cluster level for the entire brain. Then, from within these areas, active areas after subtracting TTv from TV were defined ($P<0.05$ corrected for multiple comparisons at the voxel level for the searched volume).

2.7. Group analysis with random effect model

The weighted sum of the parameter estimates in the individual analysis constituted ‘contrast’ images, which were used for the group analysis [16]. Contrast images obtained via individual analysis represent the normalised task-related increment of the MR signal of each subject. For each contrast, an unpaired Student’s t-test was performed for each and every voxel within the brain to obtain population inferences. The resulting set of voxel values for each contrast constituted a statistical parametric map of the t statistic (SPM$\{t\}$). The SPM$\{t\}$ was transformed to the unit normal distribution (SPM$\{Z\}$). The threshold for SPM$\{Z\}$ was set at $Z>3.09$ and $P<0.05$ with a correction for multiple comparisons at the cluster level for the entire brain [15]. Similar to the individual analysis, to depict cross-modal areas that are more prominently activated when information coming from two different sensory modalities is matched than when the matching is performed within either modality, activated regions by both TV–TTv and TV–VVt were depicted: First the areas activated by the comparison of TV–VVt was defined with statistical threshold of $P<0.05$ corrected at multiple comparisons searched over whole brain, at cluster level. Within these areas, activated regions by TV–TTv were searched ($P<0.05$ corrected for multiple comparisons at voxel level for the searched volume).

3. Results

3.1. Task performance

The mean (±S.D.) percentage of correct responses was

72.9±11.1% for the TV task, 63.6±9.9% for the TTv task, 99.3±1.9% for the VVt task, and 72.9±11.5% for the TT task. Performance on the VVt task was significantly better than those of the other tasks ($P<0.0001$, one-way analysis of variance followed by Fisher's PLSD).

3.2. Group analysis with random effect model

Several brain areas were active specifically during the tactile and visual discrimination tasks, as shown by second level analysis with a random effect model (Tables 1 and 2, Fig. 2). During the TT task, the left primary sensorimotor area (SM1) was active, and the postcentral gyrus (GPoC), superior parietal lobule (LPs), anterior portion of the intraparietal sulcus (AIP), secondary somatosensory area (SII), dorsal premotor area (PMd), thalamus, cerebellum, and supplementary motor area (SMA) were active bilaterally (Table 1, Fig. 2). During the TTv task, the left SM1, LPs, SII and thalamus were active, and the GPoC, AIP, fusiform gyrus (GF), PMd, dorsolateral prefrontal cortex (DLPFC), cerebellum, and SMA were active bilaterally (Table 1, Fig. 2). During the VVt task, the left SM1 extending to the AIP was active, the right AIP was active, and the cuneus, GF, cerebellum, and SMA were active bilaterally (Table 1, Fig. 2). During the TV task, the left SM1 and thalamus were active, the right PMd was active, and the GPoC, AIP, the posterior portion of intraparietal sulcus (PIP), cuneus, GF, DLPFC, cerebellum, and SMA were active bilaterally (Table 1, Fig. 2).

The lingual gyrus and GF were active bilaterally during visual shape matching, identified by subtracting images taken during the TT task from those taken during the VVt task (Table 2, Fig. 2). These areas were active during the task with visual involvement but not without visual involvement. The task-related increase in the MR signal in GF was suppressed when the subjects engaged in tactile–tactile matching with irrelevant visual input (TTv; Fig. 3).

The PIP, close to the parieto-occipital sulcus, was more prominently active bilaterally during the TV task compared with both the TTv and VVt tasks (Table 2, Figs. 2 and 3). The task-related increase in the MR signal in PIP was enhanced during cross-modal matching to a greater extent than during intra-modal matching tasks. These foci were located posterior to the areas active during tactile–tactile matching, that is, the AIP (Figs. 2 and 3), which was suppressed during visual–visual matching with irrelevant tactile input (Fig. 3).

Individual analysis consistently revealed more prominent activity in the PIP close to the parieto-occipital sulcus during the TV task than during the TTv and VVt tasks. Talairach's coordinates of the local maximum in the left PIP was $x=-24.7\pm4.8$ mm, $y=-69.1\pm6.5$ mm, and $z=40.0\pm6.8$ mm (mean±S.D., $n=11$), and that in the right PIP was $x=29.6\pm4.7$ mm, $y=-66.0\pm8.1$ mm, and $z=42.0\pm7.1$ mm (mean±S.D., $n=11$). Fig. 4 shows two representative examples.

4. Discussion

The results of the present study show that tactile–tactile matching involved parietal but not occipital cortex, while visual–visual matching involved occipital but not parietal cortex. Also, the results suggest that PIP close to the parieto-occipital sulcus is involved in the integration of visual and tactile sensory information.

4.1. Task design

In the present study, we adopted a block design rather than an event-related design because a block design was statistically more efficient and the task instruction was considerably more straightforward for these tasks [17]. Our task design was similar to that of Banati et al. [2] who examined tactile–visual cross-modal shape matching using PET imaging and a modified version of the 'arc-circle' test, a 2D extraction of shape in visual or tactile sensory modalities. A major difference was that we included TT, TV, TTv and VVt tasks so as to identify active areas common to both the TV-VVt subtraction and the TV− TTv subtraction, whereas Banati et al. [2] examined activation during TV and VV tasks without including a TT task. Our study design also included rest periods to detect task-related activity within each task session, whereas Banati et al. [2] did not.

Shapes were limited to 2D in the present study, because in tactile–visual matching [19] of 3D objects, the tactile information is of a different nature than that of the visual information. While feeling an object, only a limited portion of the surface is touched by the fingers at any one time. Likewise, only a part of the surface of a 3D object is seen, if viewed at a stationary angle [19]. Thus, both bottom-up (image-based) and top-down (knowledge-based) components are involved in the recognition of 3D objects [30]. To minimise the difference in input information and to eliminate the top-down component of the process, 2D engraved patterns were utilised in the present study. Although the tactile stimuli are engraved, and thus, 3D, the essential component of the task was the discrimination of the 2D alignment of the engraved patterns. In this respect, the tactile tasks used in the present study were more similar to Braille discrimination tasks [37,38] than object recognition tasks [1,19]. Target stimuli for the matching tasks, both visual and tactile, were presented simultaneously to minimise the involvement of working memory. All the tasks except the TT task, were also designed to match the sensory input across the tasks. Finally, the tasks were also designed specifically to avoid the need for object naming.

4.2. Task performance

Performance of the VVt task was significantly better than the tactile-relevant tasks (TV, TTv, or TT), whereas there was no significant performance difference among the

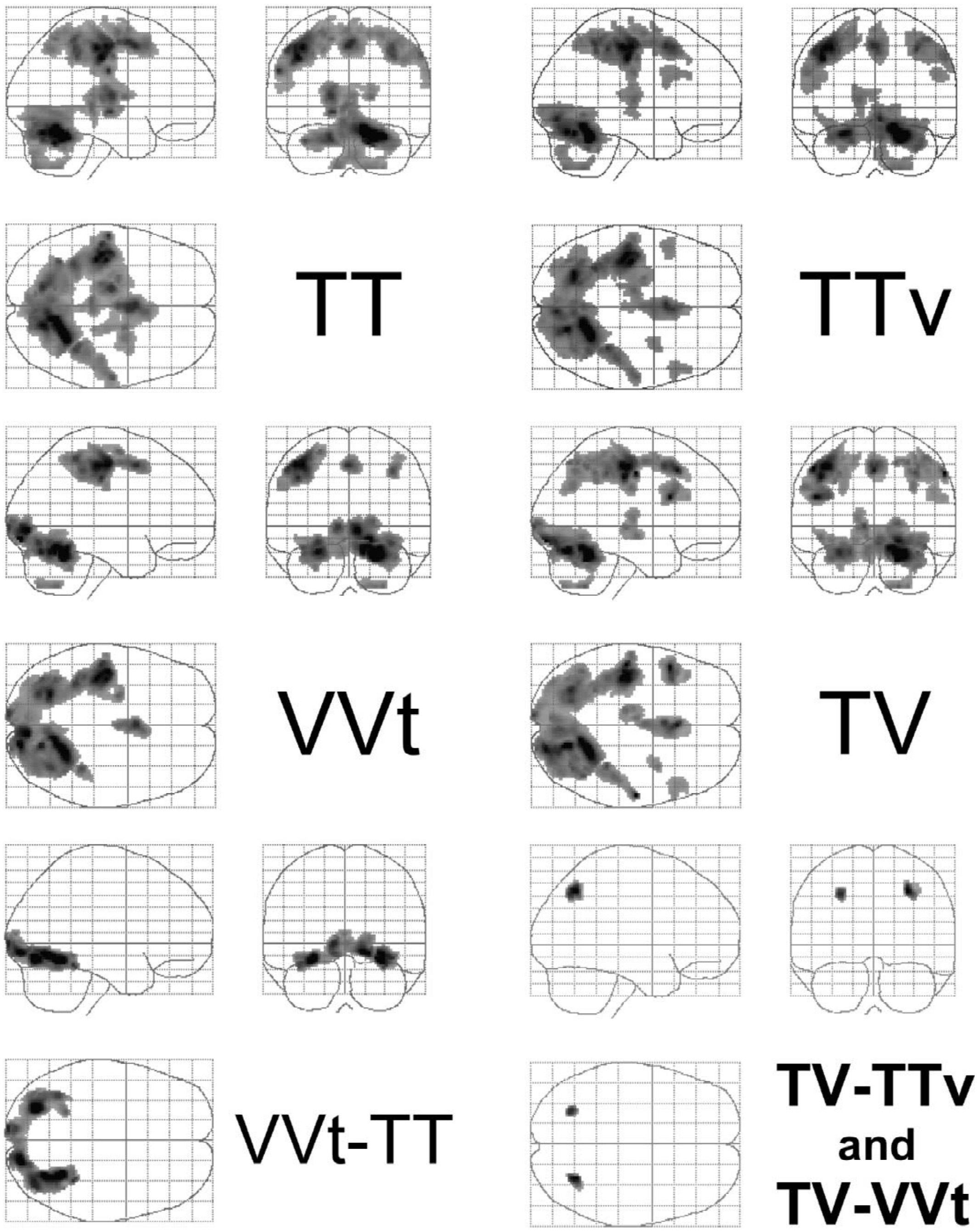

Fig. 2. Task-related activity recorded during the sessions and their subtraction. VVt-TT (lower left) is the subtraction of the images taken during the TT task from those taken during the VVt task with a statistical threshold of $P<0.05$ with correction for multiple comparisons at the cluster level with a height threshold of $Z=3.09$ [15]. The statistical parametric maps are shown in standard anatomical space. The 3D information was collapsed into 2D sagittal, coronal, and transverse images (i.e. maximum intensity projections viewed from the right, back, and top of the brain). Neural substrates for tactile–visual cross-modal matching (TV–TTv and TV–VVt) (lower right) were the bilateral posterior intraparietal areas, which was calculated by means of the TV−VVt subtraction with a statistical threshold of $P<0.05$ corrected for multiple comparisons over the searched volume within the active areas defined by the TV− TTv subtraction with a statistical threshold of $P<0.05$ at the cluster level. Activity was observed in the occipital cortex during visual–visual matching (VVt–TT), whereas the occipital cortex was not active during tactile–tactile matching (TT).

latter. Thus, the performance difference between the VVt and TV tasks may be due to the complexity of the tactile discrimination processes common to the tactile-relevant tasks, not due to cross-modal comparison of the overall tactile patterns with the visual one, which is only required for the TV task.

Table 1
Task-related activation ($n=11$)

Task	Cluster level		x (mm)	y (mm)	z (mm)	Z value	Location	
	P*	Size					Side	Area
TT	<0.001	7080	−36	−56	−28	5.06	Left	Cerebellum
			22	−54	−24	6.49	Right	Cerebellum
			−16	−14	10	5.39	Left	Thalamus
			20	−18	14	4.06	Right	Thalamus
	<0.001	5326	−38	−18	50	5.71	Left	SM1
			−38	−26	48	5.77	Left	GPoC
			−34	−44	56	5.16	Left	AIP
			−14	−60	56	4.69	Left	LPs
			−48	−18	18	3.48	Left	SII
			0	6	50	5.47		SMA
			−22	−8	62	4.70	Left	PMd
			26	−2	60	4.43	Right	PMd
	<0.001	1768	46	−22	40	4.46	Right	GPoC
			36	−44	48	5.15	Right	AIP
			16	−58	60	4.28	Right	LPs
			66	−12	20	3.62	Right	SII
TTv	<0.001	5816	−26	−66	−22	5.72	Left	Cerebellum
			20	−54	−22	6.86	Right	Cerebellum
			−36	−82	−12	3.69	Left	GF
			40	−74	−12	3.65	Right	GF
	<0.001	2744	−36	−20	50	5.79	Left	SM1
			−38	−26	50	5.79	Left	GPoC
			−32	−48	52	5.19	Left	AIP
			−16	−60	54	4.16	Left	LPs
			−48	−18	18	3.86	Left	SII
			−24	−8	62	4.25	Left	PMd
	<0.001	848	0	10	50	5.31		SMA
	<0.001	862	54	−18	44	4.48	Right	GPoC
			34	−38	48	5.19	Right	AIP
	<0.001	308	−14	−16	6	4.42	Left	Thalamus
	0.009	147	34	4	56	3.93	Right	PMd
	0.009	145	52	18	26	3.84	Right	DLPFC
	0.002	187	−48	12	24	3.78	Left	DLPFC
VVt	<0.001	5105	−26	−64	−22	5.48	Left	Cerebellum
			24	−52	−24	6.32	Right	Cerebellum
			−10	−98	0	5.04	Left	Cuneus
			20	−96	4	4.64	Right	Cuneus
			−42	−58	−12	3.45	Left	GF
			34	−64	−14	4.97	Right	GF
	<0.001	1820	−44	−18	50	5.54	Left	SM1
			−40	−24	50	5.55	Left	GPoC
			−34	−42	56	4.97	Left	AIP
	<0.001	362	0	6	50	5.16		SMA
	0.003	185	36	−38	46	4.82	Right	AIP
TV	<0.001	5522	−28	−68	−22	5.34	Left	Cerebellum
			20	−54	−24	6.41	Right	Cerebellum
			−8	−98	0	5.00	Left	Cuneus
			18	−94	2	4.76	Right	Cuneus
			−44	−58	−10	4.62	Left	GF
			42	−58	−16	4.82	Right	GF
	<0.001	2355	−38	−22	50	5.62	Left	SM1
			−56	−18	32	4.55	Left	GPoC
			−32	−42	58	4.24	Left	AIP
			−26	−62	42	4.09	Left	PIP
	<0.001	1158	46	−24	40	3.76	Right	GPoC
			34	−38	44	4.91	Right	AIP
			30	−64	46	4.45	Right	PIP
	<0.001	686	0	20	46	5.49		SMA
	<0.001	489	−48	14	24	5.61	Left	DLPFC
	<0.001	242	56	18	24	3.74	Right	DLPFC
	<0.001	282	−14	−16	6	4.22	Left	Thalamus
	0.015	111	34	4	56	4.29	Right	PMd

Abbreviations: AIP, anterior portion of the intraparietal sulcus; DLPFC, dorsolateral prefrontal cortex: GF, fusiform gyrus; GPoC, postcentral gyrus; LPs, superior parietal lobule; PMd, dorsal premotor area; PIP, posterior portion of the intraparietal sulcus; SII, secondary somatosensory area; SM1, primary sensorimotor area; SMA, supplementary motor area. All P values are corrected for multiple comparisons. Height threshold, $Z=3.09$, $P=0.001$. Extent threshold, $P=0.05$, corrected.

Table 2
Differential cortical activation ($n=11$)

Task	Cluster level		x (mm)	y (mm)	z (mm)	Z value	Location	
	*, P	Size					Side	Area
VVt − TT	<0.001	2296	−10	−92	−8	4.20	Left	GL (17)
			18	−88	−8	5.00	Right	GL (17)
			−26	−76	−12	5.06	Left	GF (19)
			30	−66	−10	4.96	Right	GF (19)
TV − TTv and	<0.001	63	−28	−66	42		Left	PIP
TV − VVt	<0.001	121	28	−64	44		Right	PIP

Abbreviations: GF, fusiform gyrus; GL, lingual gyrus; PIP, posterior portion of the intraparietal sulcus. Numbers in parentheses are Brodmann's areas. All P values are corrected for multiple comparisons. Height threshold, $Z=3.09$, $P=0.001$. Extent threshold, $P=0.05$, corrected. *, P values for the subtraction of TTv from TV corrected for multiple comparisons within the search volume defined by the subtraction of VVt from TV (corrected $P<0.05$ over the entire brain at the cluster level with a height threshold of $Z>3.09$).

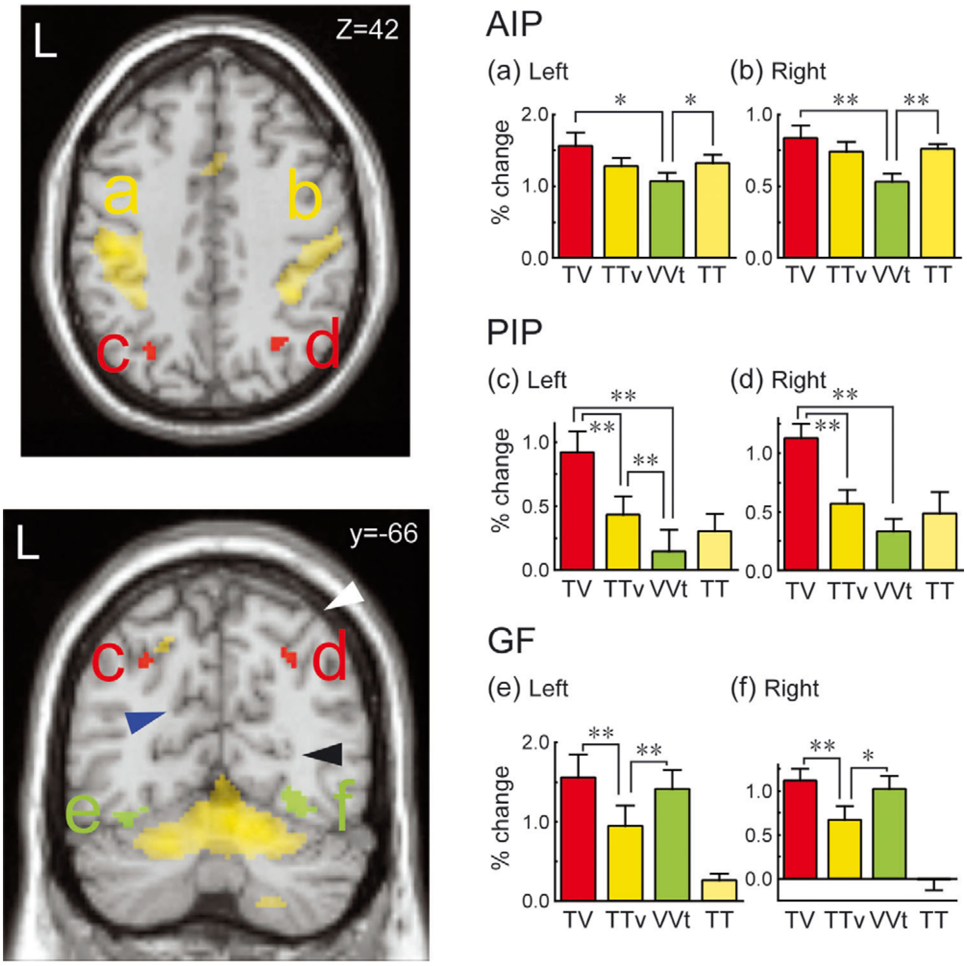

Fig. 3. Group analysis of task-related activation during cross-modal and intra-modal shape matching. Areas active during cross-modal matching (red, TV–TTv and TV–VVt), visual–visual matching (green, VVt–TT), or tactile–tactile matching (yellow, TT), are superimposed on a high-resolution transaxial MRI taken at Talairach's coordinate $z=42$ (upper left) and on a high-resolution coronal MRI taken at $y=-66$ (lower left). The statistical thresholding was as in Fig. 2. Task-related MR signal changes (%) in the (a) left AIP, (b) right AIP, (c) left PIP, (d) right PIP, (e) left GF, and (f) right GF. *, $P<0.05$, **, $P<0.001$ (paired Student's t-test, $n=11$). Task-related MR signal changes (%) in the AIP in each subject relative to the rest period was calculated using a spherical volume of interest (VOI) with a diameter of 12 mm centred on the local maximum in SPM{Z} during the TT task. Data points represent the means ± the standard error of the mean (SEM) of 11 subjects. Similarly, the task-related MR signal changes in the PIP (using the TV− TTv and TV−VVt subtractions: small discrete clusters were used as VOIs) and the GF (VVt− TT: using a spherical VOI with a diameter of 12 mm in the GF) were calculated. The white arrowhead indicates the right intraparietal sulcus, the blue indicates the left parieto-occipital sulcus, and the black indicates the right calcarine sulcus.

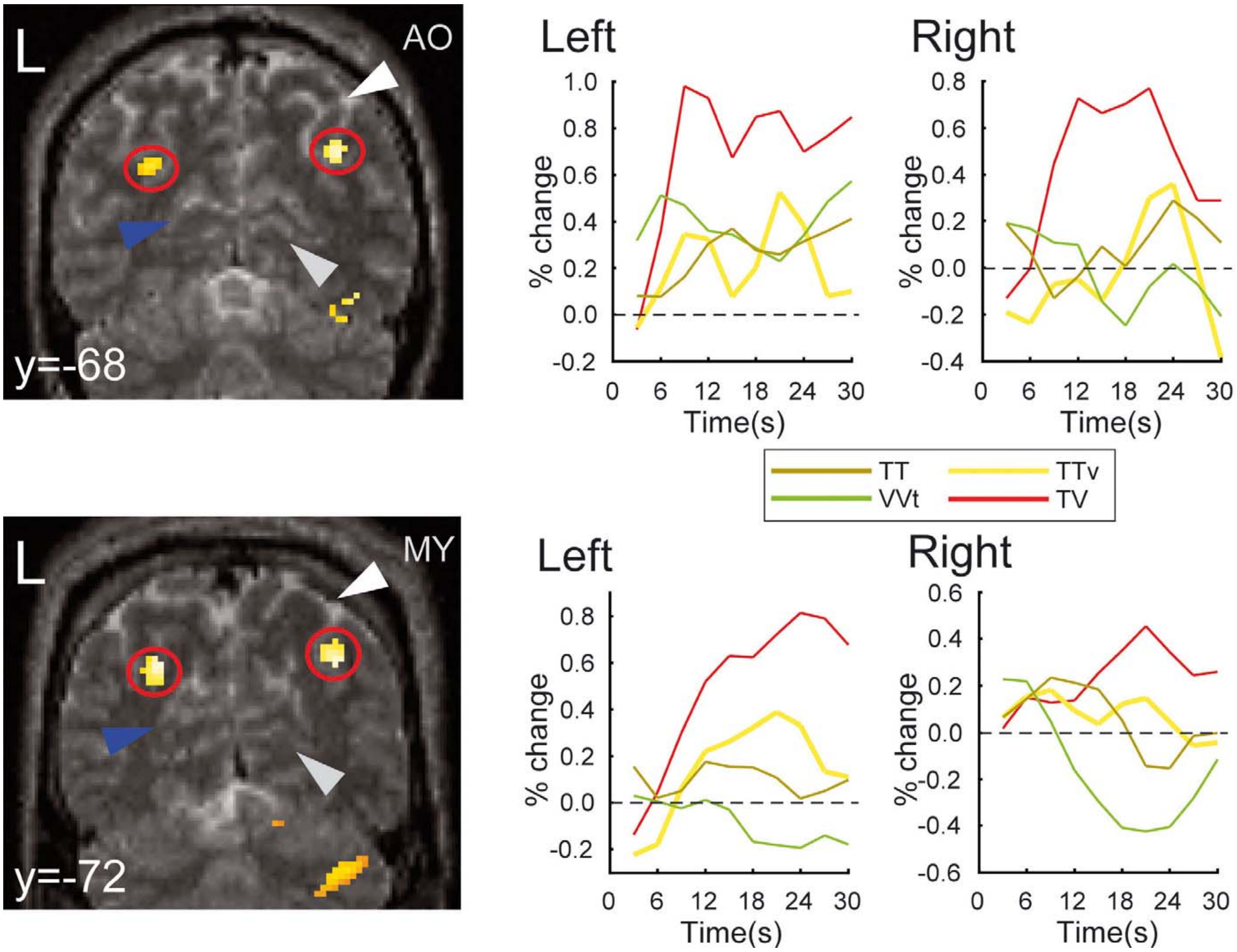

Fig. 4. Statistical parametric maps of individual analysis of neural activity during cross-modal matching (the TV task) compared with that during the TTv and VVt tasks. Representative cases are shown (left column, top: subject AO, bottom: subject MY). The task-related increase in MR signal (activation shown in yellow) was superimposed on a T2-weighted high-resolution coronal MRI of each individual. The upper left image was taken at Talairach's coordinate $y = -68$ mm, and the lower left image was taken at $y = -72$ mm. Statistical thresholding as in Fig. 2. The white arrowhead indicates the right intraparietal sulcus, the blue indicates the left parieto-occipital sulcus, and the grey indicates the right calcarine sulcus. Task-related MR signal changes (%) in the left (middle column) and right (right column) PIP during the 30-s task period, averaged across four task periods in two sessions for each task. The Talairach's coordinates of the left PIP were [−28, −68, 30] (AO) and [−26, −76, 34] (MY) and the right PIP were [34, −66, 34] (AO) and [36, −74, 34] (MY).

4.3. Task-related activation

Intra-modal shape matching tasks activated the modality-dependent cortical areas. Tactile–tactile shape matching revealed the neural substrates for tactile discrimination with active exploration. The activity in the parietal cortex including the bilateral GPoC extending to the LPi and LPs was likely due to the initial processing of the tactile stimuli. This may reflect the fact that somatotopic representation is contralateral in BA 3 and 1, and bilateral in 2, 5, and 7 [22]. A PET study [36] showed that AIP activity is related to perception of the somatosensory form (length and shape), and that SII activity is related to perception of roughness. This finding suggests the existence of separate processing streams for the different somatosensory submodalities of microgeometry and macrogeometry. Other neuroimaging studies have shown that the bilateral postcentral gyrus, LPs, and the cortex lining the AIP are activated specifically during tactile processing of the shape and length of objects [19,33,35,42] or during non-Braille tactile shape discrimination [38,39]. Thus, activity in the AIP may be due to tactile shape discrimination.

Visual–visual shape matching, as calculated by the VVt–TT subtraction, involved activation of the ventral visual cortices, including GF. This is consistent with the results of previous studies of visual object or shape recognition [44,47,49].

In the present study, there was no occipital activation during the TT task, which is consistent with the results of previous studies of tactile shape discrimination [38–40]. Recent functional neuroimaging studies, however, suggest the recruitment of the visual cortices during tactile object recognition [1,7]. Using a 3D object recognition task, Amedi et al. [1] showed that the lateral occipital cortex (LOC) was activated by both visual and tactile stimuli. They suggested that the LOC may process bimodal (visual–tactile) object-selective information in humans, and its activation may reflect the accessing of stored object-related visual information via tactile cues (top-down processing) [1]. The present study utilised 2D stimuli, which do not require top-down processing to reconstruct a whole image, but require only discrimination of the shape, without recognition of the object. Therefore, it may be that during the TT task in the present study, there was no need to refer to the stored object-related visual information [1], and, thus, no activation of the visual cortex.

There was a suppression of the task-related increase in the MR signal during tactile shape matching and visual shape matching. During the tactile-related task (TTv) compared to the visual-related tasks (TV and VVt), the signal was suppressed in the GF while sensory input was equal across the different task conditions. GF was not activated during the TT task. Human neuroimaging studies with PET indicate that selective attention to one sensory modality suppresses activity in cortical areas that process input from other sensory modalities [21,23]. Therefore, it is likely that the GF is a unimodal area involved in processing visual information. Similar suppression was observed in the AIP, which was less active during the VVt task compared with other tasks (especially tactile-related tasks), probably due to attentional shifts towards vision. Interestingly, the AIP was active during all tasks (TV, TTv, VVt, and TT). This is consistent with the idea that the AIP is a polysensory area [4], although the present study did not include a visual-only task, which would have tested this possibility further.

4.4. Tactile–visual matching

The bilateral PIP close to the parieto-occipital sulcus was activated more prominently during the TV task than either the TTv or VVt tasks. In the PIP, the increase in the MR signal observed during the TV task was larger than the sum of those during each intra-modal matching task.

In macaque monkeys, the ventral intraparietal area (VIP), located in the fundus of the intraparietal sulcus (IPS) is known to contain cells with distinct polysensory receptive fields, responsive to visual and somatosensory stimuli [6], and the VIP to have significant and consistent connection between somatosensory areas 2, 1, and 3a, primarily related to digit or digit/face representation [26]. The fundus of the IPS has reciprocal connections with visual-related areas: visual area 3, the ventral posterior area [10], the parieto-occipital area [5,43], and the middle temporal area [27,46]. The lateral portion of the VIP tends to have stronger connections with visual-related areas, whereas the medial portion is more strongly connected to sensorimotor-related areas [26]. The lateral bank of the IPS is extensively interconnected with known visual areas [3,5]. Sereno and Maunsell [45] found that neurones in the lateral intraparietal area (LIP), a part of the IPS, exhibited shape selectivity equivalent to those in the ventral pathways. The LIP receives projections from areas V4, and the inferior temporal cortex (areas TEO and TE) that are the higher stations of the visual information stream for object recognition (ventral ‘what’ pathways) [48]. Furthermore, using fMRI, Bremmer et al. [4] revealed that the depth of the human IPS was equivalent to that of monkey area VIP using polysensory (visual, auditory, and somatosensory) stimuli conveying motion information. These results strongly indicate that the cortical areas in and around the PIP are polysensory areas [26], as stated in our hypothesis.

The cross-modal TV task was expected to be more difficult than the other tasks included in the present study, because four patterns had to be evaluated and matched in a cross-modal fashion whereas in the TT, TTv, or VVt tasks, only two patterns were to be matched. Nevertheless, the performance on the TV task was equivalent to that on the TT and TTv tasks and significantly worse than that on the VVt task. Considering that visual matching is quite easy and that the TT, TTv, and TV tasks were performed equally well, the tactile component is likely the limiting factor for the performance of tactile-relevant tasks. Performance was equivalent, nevertheless, it is possible that the tasks required different or additional mental processes with attentional or task-demand differences that may have been revealed in differences in duty cycle or reaction times, however, these were not measured in the present study.

When the visual and tactile shape information to be matched is presented with different position, orientation, or size, it is expected that additional mental processes will be used to complete the matching compared with those used for intra-modal matching of shapes presented simultaneously with the same position, orientation, and size. The fact that the PIP was more active during the TV task than during the other tasks, is consistent with its suggested role in mental imagery or mental rotation of shape information [20,29]. On the other hand, intra-modal matching tasks (TTv and VVt) did not require such transformation because the shape information was presented side by side. Taken together, these data suggest that PIP activity may be related to amodal transformation of the shape information.

4.5. Comparison with previous neuroimaging studies

Previous PET studies [2,19] concluded that the claustrum was specifically involved in cross-modal matching. Additionally, Banati et al. [2] showed anterior parietal activation during cross-modal matching tasks. The insula, claustrum, or anterior parietal cortices were not active during cross-modal matching tasks in the present study, but the PIP was, an observation not reported in the previous studies. The differences in the results between the previous results and the present study likely arise from differences in the methods and task designs. First, the previous studies utilised PET imaging, which has limited sensitivity for cross-modal and intra-modal recognition comparison compared with fMRI [25]. Second, Hadjikhani and Roland [19] used 3D ellipsoids as tactile stimuli whereas 2D carved patterns were used in the present study. Third, Hadjikhani and Roland [19] calculated the common active areas to TV−TT and TV−VV subtractions, in which the TV task contained a short-term memory component, whereas their TT and VV tasks did not. In the present study, the common active areas to TV−TTv and TV−VVt subtractions were calculated, and none of the tasks contained a memory component. Fourth, Banati et al. [2] observed inferior parietal activity by means of a TV− VVt subtraction. They did not include a tactile–tactile control because of the difficulty in designing such a task. Without a tactile–tactile control task, it is difficult to identify a cross-modal activation area. Additionally, a TV−VVt subtraction of our data reveals an activation pattern similar to that presented by Banati et al. [2], except for activity in the middle and superior temporal gyri and insula. As Banati et al. [2] suggested, the inferior parietal activity may reflect higher-order tactile processing, that is, tactile shape information processing [7,24]. Finally, the statistical threshold used by Banati et al. [2] was more lenient than that used in the present study, without correction for multiple comparisons. In sum, the difference of the patterns in activity across the studies may be due to different methods and task designs.

Using fMRI, Grefkes et al. [18] found that the left anterior part of the intraparietal sulcus is specifically involved in the cross-modal transfer of object information. They used a cross-modal and intra-modal delayed match-to-sample task that required both object encoding and recognition processes, and, thus, working memory. Furthermore, Grefkes et al. [18] explicitly included a mental rotation aspect in each task in each matching session. All tasks (TV, VT, VV, and TT) required the mental transformation (in this particular case, mental rotation in three dimensional space) of information from either tactile or visual modalities, and, thus, the activity associated with this mental transformation process would be expected to be subtracted out in the TV−VV or TV−TT subtractions. In fact, they did not report any PIP activation. On the other hand, because of their a priori hypothesis, they focused on the AIP and PIP activity was not explicitly tested. Another difference is that the task-design of Grefkes et al. [18] included a working-memory component, whereas the design of the tasks used in the present study excluded it. Considering that the AIP may be related to short-term memory processes for 3D object features [31] and that the AIP is involved in the final steps of tactile shape processing, the AIP may be related to the amodal 3D shape process, which requires short-term memory function [18]. Although the IPS is known to be a polymodal area, simple polymodal convergence may not be sufficient for the cross-modal transfer of shape information [18]. The results of the present study and those of Grefkes et al. [18] suggest that different aspects of cross-modal shape processing occur in discrete areas of the polymodal IPS: the anterior portion is related to the amodal 3D shape processes requiring short-term memory and the posterior portion is involved in amodal transformation of shape information.

The results of the present study demonstrate bilateral, cross-modal activation of the PIP, despite the fact that the tactile tasks were performed using only the right hand. This may explain why previous lesion studies failed to produce consistent deficits in cross-modal matching [28]. Tactile somatotopic representation is contralateral in BA 3, and 1, and bilateral in 2, 5, and 7 [22]. Given bilateral cross-modal nodes in the PIP, a functional deficit would be unclear or absent with unilateral lesions, as shown in a previous report [19].

In conclusion, the results of the present study indicate that the PIP near the parieto-occipital sulcus may be involved in the integration of relevant shape information from different sensory modalities during tactile–visual matching.

Acknowledgements

The authors are appreciative of the subjects’ participation and Dr. Jennifer Kahle for skilful editing. This study was supported in part by a research grant for the ‘Research for the Future’ Program from the Japan Society for the Promotion of Science (JSPS-RFTF97L00203).

References

[1] A. Amedi, R. Malach, T. Hendler, S. Peled, E. Zohary, Visuo–haptic object-related activation in the ventral visual pathway, Nat. Neurosci. 4 (2001) 324–330.

[2] R.B. Banati, G.W. Goerres, C. Tjoa, J.P. Aggleton, P. Grasby, The functional anatomy of visual–tactile integration in man: a study using positron emission tomography, Neuropsychologia 38 (2000) 115–124.

[3] D. Boussaoud, L.G. Ungerleider, R. Desimone, Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque, J. Comp. Neurol. 296 (1990) 462–495.

[4] F. Bremmer, A. Schlack, N.J. Shah, O. Zafiris, M. Kubischik, K. Hoffmann, K. Zilles, G.R. Fink, Polymodal motion processing in posterior parietal and premotor cortex: a human fMRI study strongly implies equivalencies between humans and monkeys, Neuron 29 (2001) 287–296.

[5] C.L. Colby, R. Gattass, C.R. Olson, C.G. Gross, Topographical organization of cortical afferents to extrastriate visual area PO in the macaque: a dual tracer study, J. Comp. Neurol. 269 (1988) 392–413.

[6] C.L. Colby, J.R. Duhamel, Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey, Neuropsychologia 29 (1991) 517–537.

[7] E. Deibert, M. Kraut, S. Kremen, J. Hart, Neural pathways in tactile object recognition, Neurology 52 (1999) 1413–1417.

[8] G. Ettlinger, W.A. Wilson, Cross-modal performance: behavioural processes, phylogenetic considerations and neural mechanisms, Behav. Brain Res. 40 (1990) 169–192.

[9] A.C. Evans, M. Kamber, D.L. Collins, D. MacDonald, in: S.D. Shorvon, D.R. Fish, F. Andermann, G.M. Bydder (Eds.), An MRI-based probabalistic atlas of neuroanatomy, Magnetic Resonance Scanning and Epilepsy, Plenum Press, New York, 1994, pp. 263–274.

[10] D.J. Felleman, A. Burkhalter, D.C. Van Essen, Cortical connections of areas V3 and VP of macaque monkey extrastriate visual cortex, J. Comp. Neurol. 379 (1997) 21–47.

[11] J.A. Fodor, The Modularity of Mind, an Essay on Faculty Psychology, MIT press, Cambridge, 1983.

[12] K.J. Friston, K.J. Worsley, R.S.J. Frackowiak, J.C. Mazziotta, A.C. Evans, Assessing the significance of focal activations using their spatial extent, Hum. Brain Mapp. 1 (1994) 210–220.

[13] K.J. Friston, J. Ashburner, C.D. Frith, J.D. Heather, R.S.J. Frackowiak, Spatial registration and normalization of images, Hum. Brain Mapp. 2 (1995) 165–189.

[14] K.J. Friston, A.P. Holmes, K.J. Worsley, J.B. Poline, C.D. Frith, R.S.J. Frackowiak, Statistical parametric maps in functional imaging: a general linear approach, Hum. Brain Mapp. 2 (1995) 189–210.

[15] K.J. Friston, A.P. Holmes, J.B. Poline, C.J. Price, C.D. Frith, Detecting activations in PET and fMRI: levels of inference and power, NeuroImage 4 (1996) 223–235.

[16] K.J. Friston, A.P. Holmes, K.J. Worsley, How many subjects constitute a study?, NeuroImage 10 (1999) 1–5.

[17] K.J. Friston, E. Zarahn, O. Josephs, R.N. Henson, A.M. Dale, Stochastic designs in event-related fMRI, Neuroimage 10 (1999) 607–619.

[18] C. Grefkes, P.H. Weiss, K. Zilles, G.R. Fink, Crossmodal processing of object features in human anterior intraparietal cortex: an fMRI study implies equivalencies between humans and monkeys, Neuron 35 (2002) 173–184.

[19] N. Hadjikhani, P.E. Roland, Cross-modal transfer of information between the tactile and the visual representations in the human brain: a positron emission tomographic study, J. Neurosci. 18 (1998) 1072–1084.

[20] I.M. Harris, G.F. Egan, C. Sonkkila, H.J. Tochon-Danguy, G. Paxinos, J.D. Watson, Selective right parietal lobe activation during mental rotation: a parametric PET study, Brain 123 (2000) 65–73.

[21] J. Haxby, B. Horwitz, L.G. Ungerleider, J. Maisog, P. Pietrini, C. Grady, The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations, J. Neurosci. 14 (1994) 6336–6353.

[22] Y. Iwamura, Hierarchical somatosensory processing, Curr. Opin. Neurobiol. 8 (1998) 522–528.

[23] R. Kawashima, B.T. O'Sullivan, P.E. Roland, Positron-emission tomography studies of cross-modality inhibition in selective attentional tasks: closing the 'mind's eye', Proc. Natl. Acad. Sci. USA 92 (1995) 5969–5972.

[24] R. Kawashima, J. Watanabe, T. Kato, A. Nakamura, K. Hatano, T. Schormann, K. Sato, H. Fukuda, K. Ito, K. Zilles, Direction of cross-modal information transfer affects human brain activation: a PET study, Eur. J. Neurosci. 16 (2002) 137–144.

[25] M. Lepage, A.R. McIntosh, E. Tulving, Transperceptual encoding and retrieval processes in memory: a PET study of visual and haptic objects, Neuroimage 14 (2001) 572–584.

[26] J.W. Lewis, D.C. Van Essen, Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey, J. Comp. Neurol. 428 (2000) 112–137.

[27] J.H. Maunsell, D.C. Van Essen, The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey, J. Neurosci. 3 (1983) 2563–2586.

[28] C. McNally, G. Ettlinger, A.R. Smith, Cross-modal performance in patients with cerebral lesions, Cortex 18 (1982) 91–104.

[29] E. Mellet, L. Petit, B. Mazoyer, M. Denis, N. Tzourio, Reopening the mental imagery debate: lessons from functional anatomy, Neuroimage 8 (1998) 129–139.

[30] C. Moore, S.A. Engel, Neural response to perception of volume in the lateral occipital complex, Neuron 29 (2001) 277–286.

[31] A. Murata, V. Gallese, M. Kaseda, H. Sakata, Parietal neurones related to memory-guided hand manipulation, J. Neurophysiol. 75 (1996) 2180–2186.

[32] R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory, Neuropsychologia 9 (1971) 97–113.

[33] B.T. O'Sullivan, P.E. Roland, R. Kawashima, A PET study of somatosensory discrimination in man. Microgeometry versus macrogeometry, Eur. J. Neurosci. 6 (1994) 137–148.

[34] M. Petrides, S.D. Iversen, Cross-modal matching and the primate prefrontal cortex, Science 192 (1976) 1023–1024.

[35] P.E. Roland, B. Larsen, Focal increase of cerebral blood flow during stereognostic testing in man, Arch. Neurol. 33 (1976) 551–558.

[36] P.E. Roland, B.T. O'Sullivan, R. Kawashima, Shape and roughness activate different somatosensory areas in the human brain, Proc. Natl. Acad. Sci. USA 95 (1998) 3295–3300.

[37] N. Sadato, A. Pascual-Leone, J. Grafman, V. Ibanez, M.P. Deiber, G. Dold, M. Hallett, Activation of the primary visual cortex by Braille reading in blind subjects, Nature 380 (1996) 526–528.

[38] N. Sadato, A. Pascual-Leone, J. Grafman, M.P. Deiber, V. Ibanez, M. Hallett, Neural networks for Braille reading by the blind, Brain 121 (1998) 1213–1229.

[39] N. Sadato, V. Ibanez, M.P. Deiber, M. Hallett, Gender difference in premotor activity during active tactile discrimination, NeuroImage 5 (2000) 532–540.

[40] N. Sadato, T. Okada, M. Honda, Y. Yonekura, Critical period for cross-modal plasticity in blind humans: a functional MRI study, NeuroImage 16 (2002) 389–400.

[41] A. Sahgal, M. Petrides, S.D. Iversen, Cross-modal matching in the monkey after discrete temporal lobe lesions, Nature 257 (1975) 672–674.

[42] R.J. Seitz, P.E. Roland, C. Bohm, T. Greitz, S. Stone-Elander, Somatosensory discrimination of shape: tactile exploration and cerebral activation, Eur. J. Neurosci. 3 (1991) 481–492.

[43] B. Seltzer, D.N. Pandya, Converging visual and somatic sensory cortical input to the intraparietal sulcus of the rhesus monkey, Brain Res. 192 (1980) 339–351.

[44] M.I. Sereno, A.M. Dale, J.B. Reppas, K.K. Kwong, J.W. Belliveau, T.J. Brady, B.R. Rosen, R. Tootell, Borders of multiple visual areas in humans revealed by functional magnetic-resonance-imaging, Science 268 (1995) 889–893.

[45] A.B. Sereno, J.H. Maunsell, Shape selectivity in primate lateral intraparietal cortex, Nature 395 (1998) 500–503.

[46] L.G. Ungerleider, R. Desimone, Cortical connections of visual area MT in the macaque, J. Comp. Neurol. 248 (1986) 190–222.

[47] D.C. Van Essen, Behind the optic nerve: an inside view of the primate visual system, Trans. Am. Ophthalmol. Soc. 93 (1995) 123–133.

[48] M.J. Webster, J. Bachevalier, L.G. Ungerleider, Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque, Cereb. Cortex 5 (1994) 470–483.

[49] S. Zeki, J.D. Watson, C.J. Lueck, K.J. Friston, C. Kennard, R.S.J. Frackowiak, A direct demonstration of functional specialization in human visual cortex, J. Neurosci. 11 (1991) 641–649.