nature neuroscience

Resolving human object recognition in space and time

Radoslaw Martin Cichy[1], Dimitrios Pantazis[2] & Aude Oliva[1]

A comprehensive picture of object processing in the human brain requires combining both spatial and temporal information about brain activity. Here we acquired human magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI) responses to 92 object images. Multivariate pattern classification applied to MEG revealed the time course of object processing: whereas individual images were discriminated by visual representations early, ordinate and superordinate category levels emerged relatively late. Using representational similarity analysis, we combined human fMRI and MEG to show content-specific correspondence between early MEG responses and primary visual cortex (V1), and later MEG responses and inferior temporal (IT) cortex. We identified transient and persistent neural activities during object processing with sources in V1 and IT. Finally, we correlated human MEG signals to single-unit responses in monkey IT. Together, our findings provide an integrated space- and time-resolved view of human object categorization during the first few hundred milliseconds of vision.

The past decade has seen much progress in the unraveling of the neuronal mechanisms supporting human object recognition, with studies corroborating each other across species and methods[1–5]. Object recognition involves a hierarchy of regions in the occipital and temporal lobes[1,4,6–8] and unfolds over time[9–11]. However, comparing data quantitatively from different imaging modalities, such as magneto- and electroencephalography (MEG/EEG) and functional magnetic resonance imaging (fMRI) within and across species[3,12–15] remains challenging, and we still lack fundamental knowledge about where and when in the human brain visual objects are processed.

Here we demonstrate how the processing of objects in the human brain unfolds in time, using MEG, and space, using fMRI, within the first few hundred milliseconds of neural processing[1,16]. First, by applying multivariate pattern classification[17–20] to human MEG responses to object images, we show the time course with which individual images are discriminated by visual representations[19,21–23]. Whereas individual images were best linearly decodable relatively early, membership at the ordinate and superordinate levels became linearly decodable later and with distinct time courses. Second, using representational similarity analysis[19,24,25], we define correspondences between the temporal dynamics of object processing and cortical regions in the ventral visual pathway of the human brain. By comparing representational dissimilarities across MEG and fMRI responses, we distinguished MEG signals reflecting low-level visual processing in V1 from signals reflecting later object processing in IT. Further, we identify V1 and IT as two differentiable cortical sources of persistent neural activity during object vision. This suggests that the brain actively maintains representations at different processing stages of the visual hierarchy. Lastly, using previously reported single-cell recordings in macaque[26], we extend our approach across species, finding that human MEG responses to objects correlated with the patterns of neuronal spiking in monkey IT. This work resolved dynamic object processing with a fidelity that has, to our knowledge, previously not been shown by offering an integrated space- and time-resolved view of the occipitoventral visual pathway during the first few hundred milliseconds of visual processing.

RESULTS

Human participants (n = 16) viewed images of 92 real-world objects[3,26] while we acquired MEG data (**Fig. 1a** and **Supplementary Fig. 1a**). The image set comprised images of human and nonhuman faces and bodies, as well as natural and artificial objects. Images were presented for 500 ms every 1.5–2s. To maintain attention, participants performed an object-detection task on a paper clip image shown on average every four trials. Paper clip trials were excluded from further analysis.

We extracted and preprocessed peri-stimulus MEG signal from −100 ms to 1,200 ms (1 ms resolution) with respect to stimulus onset. For each time point, we used a support vector machine (SVM) classifier to classify pairwise between all conditions (object images) (**Fig. 1b**). The results of the classification (percentage decoding accuracy, 50% chance level) were stored in a 92 × 92 decoding matrix, indexed by the 92 conditions. Thus, each cell in the matrix indicates the decoding accuracy with which the classifier distinguishes between two images. This matrix is symmetric across the diagonal, with the diagonal undefined. Results were averaged across two independent sessions. **Figure 1c** shows example matrices averaged across participants (see also **Supplementary Movie 1**). For all time points, significance was determined non-parametrically at the participant level by a cluster-based randomization approach[27,28] (cluster-defining threshold $P < 0.001$, corrected significance level $P < 0.05$). 95% confidence intervals for mean peak latencies and onsets (reported in parentheses throughout the Results) were determined by bootstrapping the participant sample.

MEG signals allow pairwise decoding of individual images

What is the time course with which individual images of objects are discriminated by visual representations? We found that MEG signals

[1]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [2]McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. Correspondence should be addressed to R.M.C. (rmcichy@mit.edu).

Received 12 October 2013; accepted 23 December 2013; published online 26 January 2014; doi:10.1038/nn.3635

could resolve brain responses on the single-image level[19], for up to 92 different objects. For every time point, we averaged across all cells of the MEG decoding matrix, yielding a time course of grand average decoding accuracy across all images (**Fig. 1d**). We calculated the onset of significance, the time point for which objects were first discriminated by visual representations, and the peak latency, the time point when visual representations of each individual image were most distinct in terms of linear separability. We report onset and mean peak latency with 95% confidence intervals in parentheses (for overview with all data, see **Supplementary Table 1**).

Before and just after stimulus presentation, grand average decoding accuracy fluctuated around chance level (50%). The curve rose sharply and reached significance at 48 ms (45–51 ms), followed by a peak at 102 ms (98–107 ms) and a gradual decline (**Fig. 1d**). Notably, we observed significant decoding accuracy of individual images within each of the six subdivisions of the image set (human and nonhuman faces and bodies, natural and artificial objects; **Fig. 1e**) 51–61 ms after stimulus onset, followed by peaks at 99–112 ms (**Supplementary Table 1b**). Thus, multivariate analysis of MEG data revealed the temporal dynamics of visual content processing in the brain even at the level of individual images[19].

Time course of category decoding

To determine when visual representations discriminate object membership at superordinate, ordinate and subordinate categorization levels, we compared decoding accuracy within and between the relevant partitions of the MEG decoding matrix (**Fig. 2**). The resulting

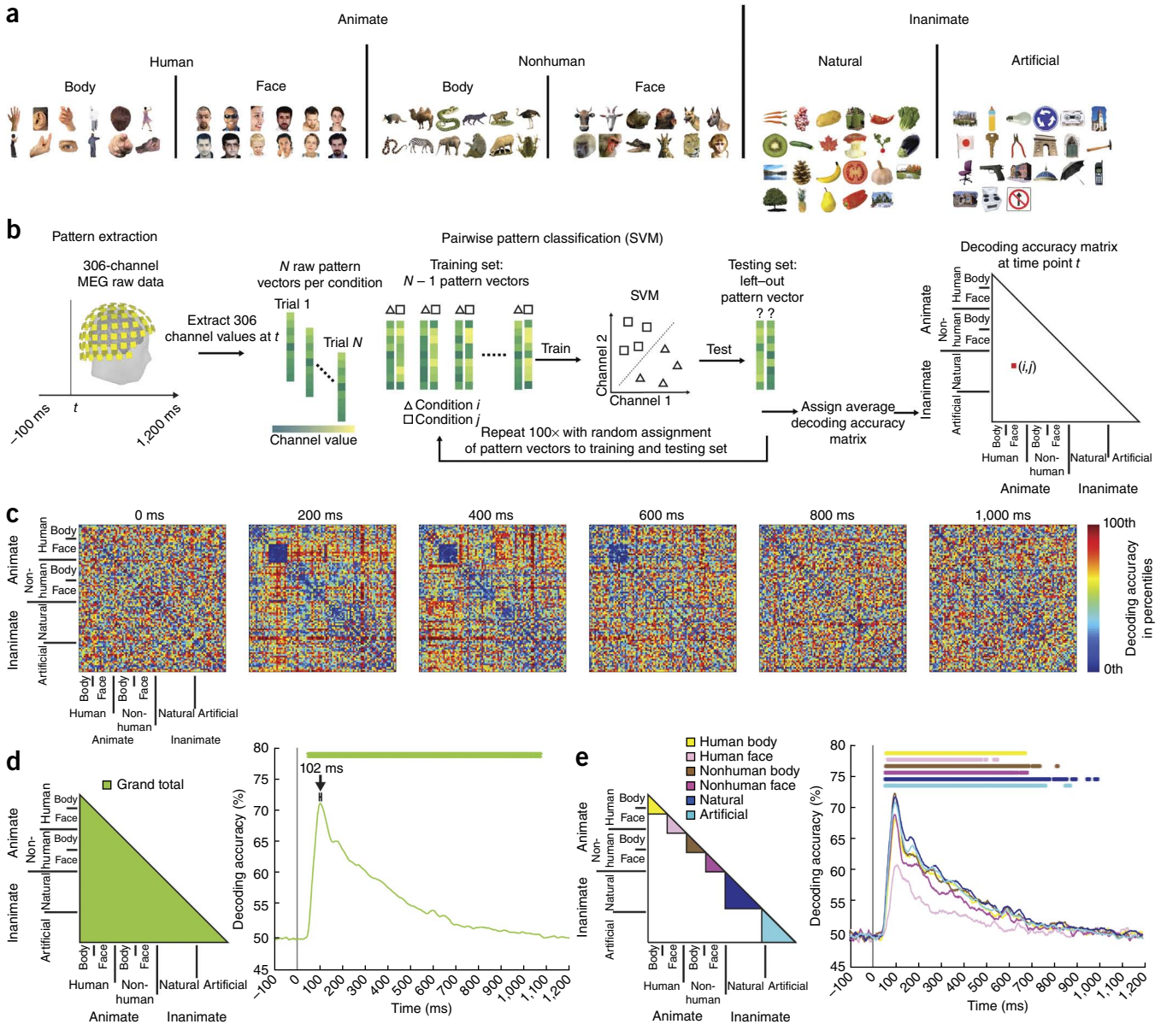

Figure 1 Decoding of images from MEG signals. (**a**) Image set of 92 images[3,26] of different categories of objects. (**b**) Multivariate analysis of MEG data. (**c**) Examples of 92 × 92 MEG decoding matrices (averaged over participants, $n = 16$). (**d**) Time course of grand total decoding was significant at 48 ms (45–51 ms), with a peak at 102 ms (98–107 ms; horizontal error bar above peak shows 95% confidence interval). (**e**) Time course of object decoding within subdivisions. The left panel illustrates the separately averaged sections of the MEG decoding matrix (color-coded), the right panel the corresponding decoding time courses. Peak latencies and onsets of significance are listed in **Supplementary Table 1b**. Rows of asterisks indicate significant time points ($n = 16$, cluster-defining threshold $P < 0.001$, corrected significance level $P < 0.05$). The gray vertical line indicates onset of image presentation.

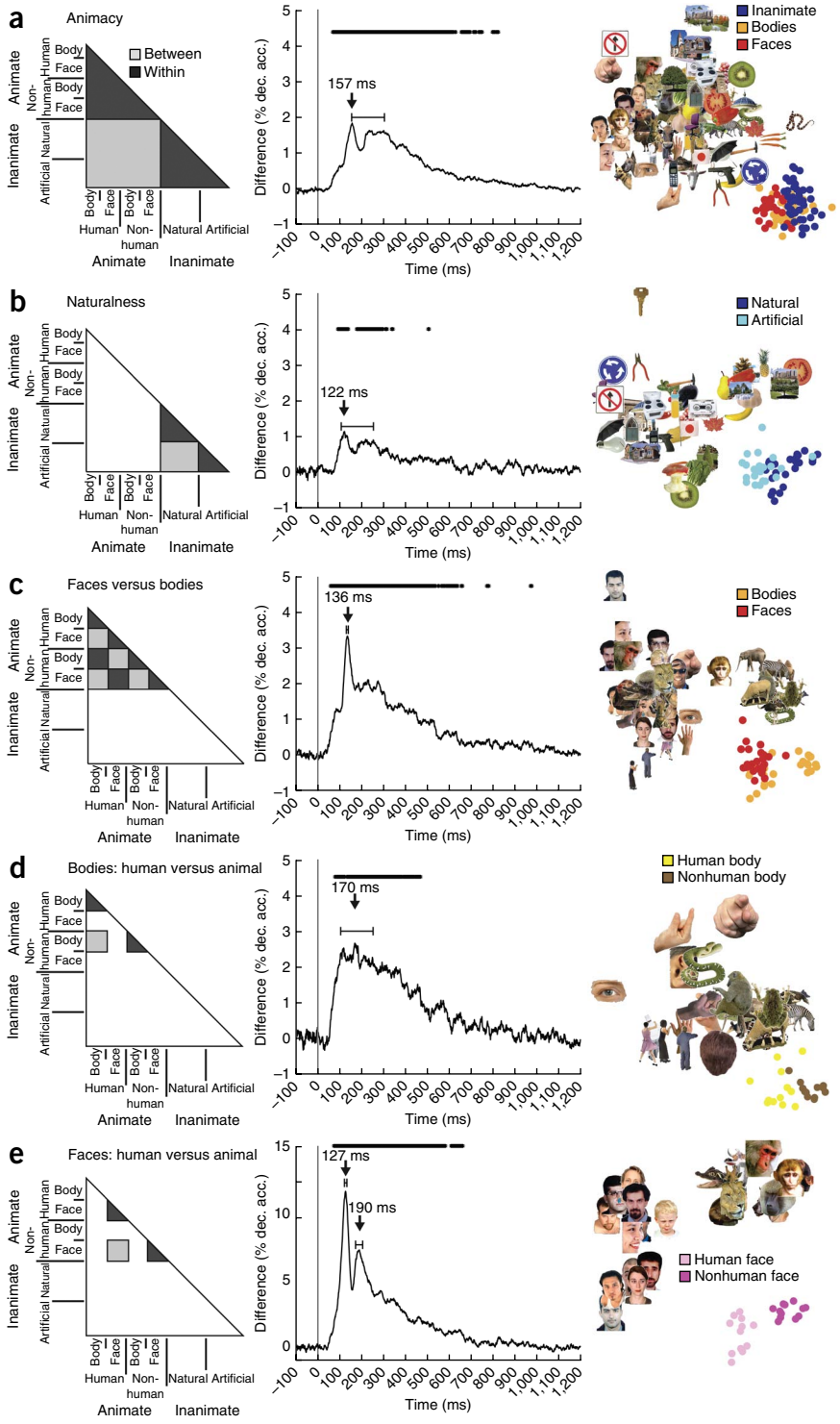

Figure 2 Time course of decoding category membership of individual objects. (**a**–**e**) We decoded object category membership for (**a**) animacy, (**b**) naturalness, (**c**) faces versus bodies, (**d**) human versus nonhuman bodies and (**e**) human versus nonhuman faces. The difference of within-subdivision (dark gray, left panel) minus between-subdivision (light gray, left panel) averaged decoding accuracies is plotted in the middle panel over time. Peaks in decoding accuracy differences indicate time points at which the ratio of dissimilarity within a subdivision to dissimilarity across subdivision is smallest. $n = 16$; asterisks, vertical gray line and error bars same as in **Figure 1**. Statistical details in **Supplementary Table 1c**. The right panel illustrates the structure in the MEG decoding matrix at peak latency revealed by the first two dimensions of the MDS (criterion: metric stress, 0.24 for **a**–**c**,**e**, 0.27 for **d**). "Dec. acc." indicates decoding accuracy.

measure (difference in decoding accuracy) indicates both linear separability and clustering of objects according to subdivision membership. Peaks in this measure represent time points at which the brain has untangled visual input such that relevant information about object membership is explicitly encoded. We determined significance as before by sign permutation tests ($n = 16$, cluster-defining threshold $P < 0.001$, corrected significance level $P < 0.05$) and 95% confidence intervals for peak latencies and onsets by bootstrapping the participant sample ($n = 16$).

We found that visual representations discriminated objects by animacy[21,23] with a peak at 157 ms (152–302 ms) (**Fig. 2a**). Similarly, visual representations discriminated objects by naturalness with a peak at 122 ms (107–254 ms) (**Fig. 2b**). Multidimensional scaling (MDS)[29,30] illustrated the main structure in the MEG decoding matrix at peak latency: clustering of objects into animate and inanimate, as well as natural and artificial.

Within the animate division, faces and bodies clustered separately. This suggested that membership to categorical divisions below the animate/inanimate distinction might be discriminated by visual representations[22,31–35]. Indeed, we found that the distinction between faces and bodies was significant at 56 ms (46–74 ms), with a clear peak at 136 ms (131–140 ms) (**Fig. 2c**). MDS at the 136-ms peak (**Fig. 2c**) showed a division between faces and bodies, dominated by nonhuman bodies versus the other conditions. At the subordinate level, we found that visual representations distinguished bodies by species with an onset at 75 ms (64–113 ms) and a peak at 170 ms (104–252 ms), and the MDS showed a clear species-specific clustering of bodies (**Fig. 2d**). We also observed a significant difference in decoding accuracy for human versus nonhuman faces starting at 70 ms (54–74 ms), followed by two prominent peaks at 127 ms (122–133 ms) and 190 ms (175–207 ms, calculated on the time window starting at the trough between the two peaks at 156 ms and 1,200 ms) (**Fig. 2e**). An MDS at the first peak illustrated the effect, with a perfect separation of faces along the species border.

For photographs of real-world objects, category membership is often associated with differences in low-level image properties. Thus, linear separability of objects by category may be achieved on the basis of low-level image property representations. However, an analysis testing the degree to which classifiers generalized across particular object images showed that the discrimination of category

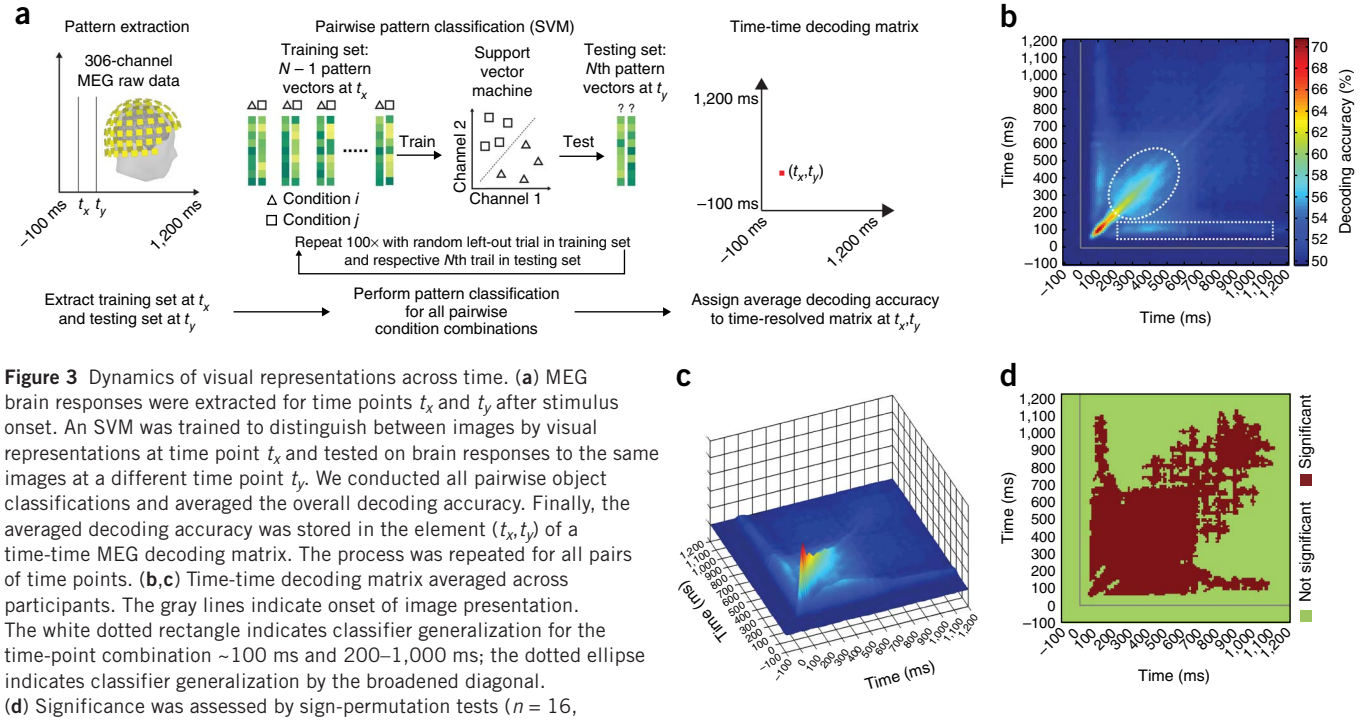

Figure 3 Dynamics of visual representations across time. (**a**) MEG brain responses were extracted for time points t_x and t_y after stimulus onset. An SVM was trained to distinguish between images by visual representations at time point t_x and tested on brain responses to the same images at a different time point t_y. We conducted all pairwise object classifications and averaged the overall decoding accuracy. Finally, the averaged decoding accuracy was stored in the element (t_x, t_y) of a time-time MEG decoding matrix. The process was repeated for all pairs of time points. (**b**,**c**) Time-time decoding matrix averaged across participants. The gray lines indicate onset of image presentation. The white dotted rectangle indicates classifier generalization for the time-point combination ~100 ms and 200–1,000 ms; the dotted ellipse indicates classifier generalization by the broadened diagonal. (**d**) Significance was assessed by sign-permutation tests ($n = 16$, cluster-defining threshold $P < 0.0001$, corrected significance level $P < 0.05$). Dark red indicates elements within the significant cluster.

membership was not solely determined by image-specific properties (**Supplementary Fig. 2**).

Comparing peak-to-peak latency differences (**Figs. 1** and **2**) using bootstrapping ($n = 16$, $P < 0.05$, Bonferroni-corrected), we found that images were discriminated earlier at the level of individual images than at higher categorization levels (all $P < 0.001$, except for human versus nonhuman body; for details, see **Supplementary Table 2**). In a behavioral experiment, new participants were asked to perform same-different image classification in the context of different categorization levels (identity, subordinate to superordinate classification) (**Supplementary Fig. 1c**). We found significant Pearson's correlations by bootstrapping ($n = 16$) between peak decoding accuracy and reaction times ($R = 0.53$, $P = 0.003$) and correctness ($R = -0.49$, $P = 0.012$) (**Supplementary Fig. 3**).

Taken together, multivariate analysis of MEG signals revealed the time course with which membership of individual objects at different categorical levels was linearly decodable. Our results complement previous work on rapid object detection in go/no-go tasks[21,23] and provide a content-based analysis of the time course with which object information is processed[19].

Transient and persistent neuronal activity

The dynamics of the decoding time courses above suggested highly variable and transient neural activity as their source. As neuronal signals propagate along the ventral visual stream, different image properties appear to be processed at subsequent time points. However, the previous analyses did not allow us to determine the existence of persistent neural activity during the course of object processing. Such persistent neural activity could maintain the results of a particular neural processing stage for later use.

Intuitively, if neuronal activity persists over time, MEG signals should be similar across time as well. To search for such similarities, we trained an SVM classifier at one time point (t_x) and tested at other time points (t_y) (**Fig. 3**). Conducting all pairwise discriminations between objects, we obtained a 92 × 92 MEG decoding matrix for every pair of time points (t_x,t_y). We then repeated the process across all pairs of time points, resulting in a four-dimensional image-image-time-time decoding matrix. Averaging across the first two dimensions yielded a time-time decoding matrix (**Fig. 3b**,**c**).

As expected, some neural activity during object processing was transient: the classifier generalized best to neighboring time points and performed poorly for distant time points. This is illustrated by the highest decoding accuracy along the diagonal and the sharp drop of decoding accuracy away from the diagonal (**Fig. 3b**), depicted as a high and steep crest of decoding accuracy (**Fig. 3c**). Notably, we also found evidence for persistent neural activity. First, the classifier generalized well for the time-point combination of ~100 ms and ~200–1,000 ms. As this effect was clearly circumscribed in time and persisted beyond the offset of image presentation at 500 ms, it is unlikely that it merely reflected constant passive influx of information during image presentation. This suggests that the brain actively maintains visual representations in early stages of the visual processing hierarchy, potentially as memory for low-level visual features[36]. Second, between ~250 ms and ~500 ms, the classifier produced a broader diagonal. This indicated that neural activity was similar across these time points, suggesting that a stable representation of objects in later stages of visual processing hierarchy is kept online.

Statistical testing (**Fig. 3d**, sign permutation test, $n = 16$, cluster-defining threshold $P < 0.0001$, corrected significance level $P < 0.05$) indicated widespread similarity of neural activity. However, the fact that this is not limited to particular time-point combinations may indicate either neural activity actively maintained at all cortical processing levels or a passive response of the brain to the constant influx of visual information during the presence of the stimulus.

In sum, across-time analysis of the dynamics in visual representations revealed both transient and persistent neuronal processing of objects. The presence of persistent neuronal activity at well-delineated time point combinations may indicate active maintenance of visual representations at different processing stages.

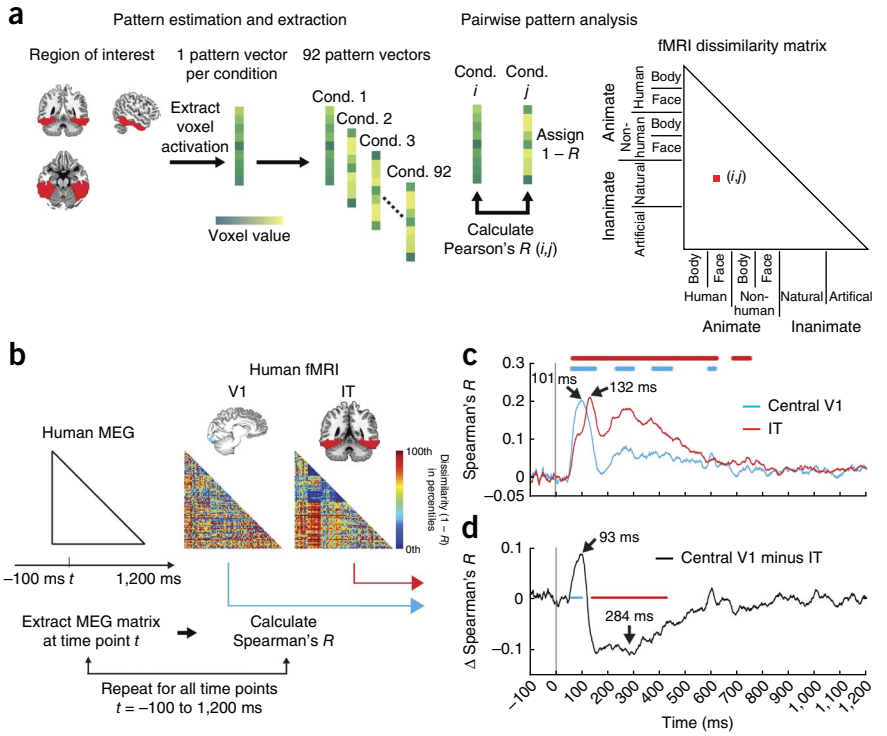

Figure 4 Relating MEG and fMRI signals in V1 and IT. (**a**) fMRI analysis. We selected voxels in two regions of interest: V1 and IT. For each condition (cond.), we extracted voxel activation values, yielding 92 pattern vectors. Then we calculated the pairwise Pearson's correlation (R) for all combinations of experimental conditions (i,j). The dissimilarity measure $1 - R$ was assigned to a 92 × 92 fMRI dissimilarity matrix indexed by the experimental conditions (i,j). This analysis was conducted independently for each region of interest. (**b**) For each time point t, we correlated (Spearman's rank-order correlation) the MEG decoding matrix to the fMRI dissimilarity matrices of V1 and IT. (**c**) MEG signals correlated with the fMRI dissimilarity matrix of central V1 earlier than with the fMRI dissimilarity matrix of IT. Blue and red asterisks indicate significant time points for V1 and IT. (**d**) Difference between the two curves in **c**. MEG correlated early more with V1 than with IT, and late more with IT than with V1. Blue and red asterisks in the plots indicate significant time points for positive and negative clusters respectively. For details, see **Supplementary Table 1d**. $n = 16$; gray line and statistical procedure as in **Figure 1**.

Resolving object recognition in space and time

What are the cortical sources of the MEG signals that discriminate objects? Given that V1 and IT process different aspects of the images[6], we expected MEG signals originating from these two cortical areas to differ: in other words, V1 and IT responses to individual objects should differ in their dissimilarity relations[3], resulting in distinct patterns over time in the MEG decoding matrices. Here we used representational similarity analysis[24,25] to show when representations extracted with MEG were comparable to those extracted with fMRI in human V1 and IT (**Fig. 4a**).

After adapting the MEG stimulation protocol to the specifics of fMRI (**Supplementary Fig. 1b**), we repeated the experiment with the same images in the same participants while acquiring fMRI data. We estimated individual (92) object image-related brain responses by fitting a general linear model. We then extracted voxel values from a region of interest (V1 or IT) to form a pattern vector (**Fig. 4a**). The resulting 92 pattern vectors were subjected to pairwise Pearson's correlation and then ordered into a 92 × 92 similarity matrix indexed by image condition. We converted matrix elements from R (similarity) to $1 - R$ (dissimilarity) to make the matrices directly comparable to the MEG decoding accuracy matrices. The above process produced two fMRI dissimilarity matrices, one for V1 and one for IT, for each participant.

Using these two fMRI matrices, we first successfully reproduced a main previous finding[3]: stronger representation of animacy in IT than in V1, shown by MDS, hierarchical clustering and quantitative testing (**Supplementary Fig. 4**). For further analysis, we computed participant-averaged fMRI matrices for both human V1 and IT. We then evaluated the extent of similar representations between fMRI and MEG by computing Spearman's rank-order correlations between fMRI dissimilarity matrices (separately for V1 and IT) and participant-specific MEG decoding accuracy matrices (separately for each time point) (**Fig. 4b**). We found that MEG signals correlated with fMRI dissimilarity matrices in human V1 and IT with different time courses (**Fig. 4c**, sign-permutation test, cluster-defining threshold $P < 0.0001$, corrected significance level $P < 0.05$, 95% confidence intervals by bootstrapping). The V1 correlation time course peaked early, at 101 ms (84–109 ms), whereas the IT time course peaked later, at 132 ms (129–290 ms) (for onset of significance see **Supplementary Table 1d**). The difference in peak-to-peak latency was significant ($n = 16$, sign-permutation test, $P = 0.016$). Importantly, comparing the V1 and IT time course directly (**Fig. 4d**), we found that MEG signals correlated more with V1 than with IT early (peak at 93 ms (79–102 ms)) and more with IT than with V1 later (peak at 284 ms (152–303 ms)).

Notably, the correlation of MEG with human IT was also present within each of the six subdivisions of the image set (**Supplementary Fig. 5**). Correlating MEG with previously reported fMRI data from human IT[3] yielded comparable effects (**Supplementary Fig. 6**)—namely, a peak at 158 ms (152–300 ms)—reinforcing the validity and generalizability of our results. Additionally, the correlation of MEG to V1 was specific to the stimulated portion of V1: an immediately adjacent V1 region corresponding to an unstimulated portion of the visual field (3–6° visual angle) showed significantly weaker correlation (sign permutation test, $n = 16$, cluster-defining threshold $P < 0.001$, corrected significance level $P < 0.05$) than central V1 (**Supplementary Fig. 7**). In summary, we have demonstrated that temporal dynamics as measured by MEG can be mapped onto distinct early and late human cortical regions along the ventral visual stream using representational similarity analysis.

Relating MEG and fMRI object signals across time

The above MEG-fMRI representational similarity analysis naturally extends to include the MEG time-time decoding matrices constructed earlier (**Fig. 3**). This analysis allows identifying the cortical sources that have persistent neural activity. We therefore correlated the fMRI dissimilarity matrices of V1 and IT with the MEG 92 × 92 decoding matrices obtained for each pair of time points (t_x,t_y) (**Figs. 3a** and **5a**). The results (**Fig. 5b**,**c**, sign permutation test, $n = 16$, cluster-defining threshold $P < 0.0001$, corrected significance level $P < 0.05$) demonstrated that neural activity for the time point combinations of ~100 ms

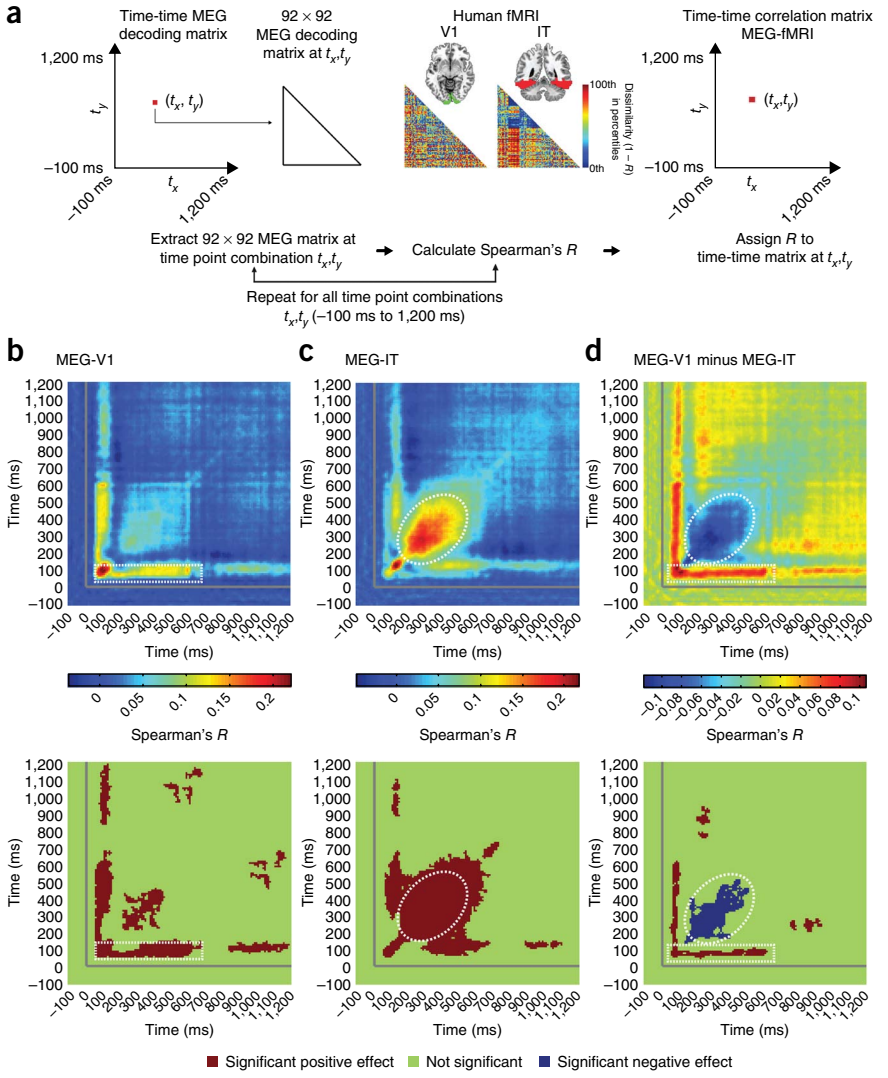

Figure 5 Relating MEG and fMRI signals across time. (**a**) Decoding matrices (92 × 92) were extracted for each combination of time points (t_x,t_y) and were correlated (Spearman's rank-order correlation) with the fMRI dissimilarity matrices for V1 and IT. The resulting correlation was assigned to a time-time MEG-fMRI correlation matrix at t_x,t_y. (**b**,**c**) Top, the time-time MEG and fMRI correlation matrix for V1 and IT. Bottom, significant cluster results (n = 16, cluster-defining threshold P < 0.0001, corrected significance level P < 0.05). Neural activity for the time point combinations of ~100 ms and ~200–1,000 ms (dotted white rectangle) correlated with V1. (**c**) Neural activity between ~250 ms and ~500 ms (dotted white ellipse) correlated with IT. (**d**) Difference between V1 and IT. Gray lines as in **Figure 1**.

and ~200–1,000 ms correlated with V1 dissimilarity matrices. In contrast, neural activity between ~250 ms and ~500 ms correlated with IT dissimilarity matrices. Notably, this was also true when comparing the correlations directly (**Fig. 5d**).

In sum, by combining fMRI and MEG, we identified V1 and IT as distinct cortical sources of persistent neural activity during visual object perception. This suggests that the visual system actively maintains neural activity at different levels of visual processing.

Relating human MEG to spiking activity in monkey IT

Previous research[3] has shown that object representations in IT are comparable in monkeys and humans. Here, using representational similarity analysis, we related the dynamics in human MEG to the pattern of activity in monkey IT (as measured electrophysiologically for the same image set[26]) (**Fig. 6a**). Brain responses in human MEG and monkey IT were significantly correlated (sign permutation test, n = 16, cluster-defining threshold P < 0.001, corrected significance level P < 0.05, 95% confidence intervals by bootstrapping), first at 66 ms (56–71 ms) and peaking at 141 ms (132–292 ms) (**Fig. 6b**). MDS at peak latency (**Fig. 6c**) revealed an arrangement strongly dominated by human faces. However, significant correlations (sign permutation test, n = 16, cluster-defining threshold P < 0.001, corrected significance level P < 0.05) were

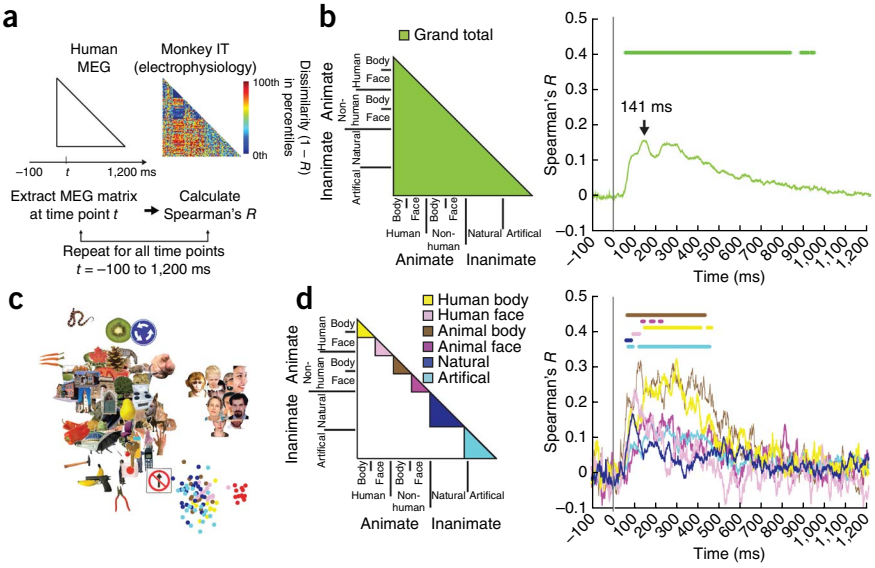

Figure 6 Relating human MEG to electrophysiological signals in monkey IT. (**a**) The MEG decoding matrix at time t was compared (Spearman's rank-order correlation R) against the monkey dissimilarity matrix in IT. The lower triangular monkey IT matrix is shown as percentiles of 1−R. (**b**) Representational dissimilarities in human MEG and monkey IT correlated significantly starting at 54 ms (52–64 ms), with a peak latency of 141 ms (132–292 ms). (**c**) MDS at peak-latency. (**d**) Results at the fine-grained level of the image set. Representational dissimilarities were similar across species and methods even for the finest categorical subdivision of the image set. n = 16; asterisks and gray line as in **Figure 1**.

also present within all subdivisions of the image set (**Fig. 6d** and **Supplementary Table 1g**). These results corroborated and extended the evidence for a common representational space for objects in monkeys and humans[3].

DISCUSSION

Using multivariate pattern classification methods[18–20] and representational similarity analysis[3,24,25] on combined human MEG-fMRI data, we have demonstrated how object recognition proceeds in space and time in the human ventral visual cortex. First, we found that whereas individual images were discriminated early, membership to ordinate and superordinate levels was discriminated later[19]. Notably, we identified neural activity that was either persistent or transient during the first few hundred milliseconds of object processing. Second, using representational similarity analysis, we combined human fMRI and MEG to show content-specific correspondence between early MEG responses and V1 fMRI responses, and later MEG responses and IT fMRI responses. Extending this analysis, we located the sources of differentiable persistent neural activity in V1 and IT. Last, we extended the representational similarity analysis across species[3,25] by showing that human MEG signals can be related to spiking activity in macaque IT. We thus extended the evidence for a common representational space for objects in monkeys and humans to the time domain.

The time course of object processing

Applying multivariate pattern classification to MEG data, we found that visual representations discriminated individual images[19] (peak at 102 ms) and then proceeded to classify them into categories. We found the peak latencies for classification of naturalness (122 ms) and animacy (157 ms) to match previous reports of neural response latencies in human and monkey IT[2,11,23,37]. The broad confidence intervals for peak latency for animacy and naturalness may indicate that object information is sustained online for more in-depth analysis after discrimination is first possible[21,23]. At the subordinate level, the body-specific peak (170 ms) and the two face-specific peaks (127 ms and 190 ms) concur with previous work (for bodies, 170–260 ms (refs. 31,38,39); for faces, first peak at 100–120 ms (ref. 35) and second peak at 170 ms (refs. 22,40)). It remains controversial whether the two face peaks have different cortical sources[41–43], a question that future studies comparing representational dissimilarities in MEG and fMRI may resolve. While our MEG results confirm a body of work[19], they generalize previous findings to a large image set and show when neural activity is transient or persistent during object analysis. This goes beyond what was previously possible with standard analysis techniques of the evoked response, allowing us to dissect the evoked response into functionally distinct neural components.

When comparing peak latencies of decoding accuracy at different levels of categorization, we observed that individual images were discriminated by visual representations early, whereas ordinate and superordinate levels emerged relatively late. However, onset latencies of significance for the various categorization levels were early (48–70 ms, except naturalness at 93 ms). Therefore, our results support models of object categorization that suggest processing of object information to begin at all levels of categorization simultaneously, with differential time courses of evidence accumulation for different levels of categorization[44,45]. Our results might seem to be at odds with previous research suggesting clear multistage processing in IT, with a stage of global processing followed by local processing[35,46,47]. However, our approach captures signals from the whole of the human brain simultaneously, not only IT. Thus, although we can determine which region predominates in shaping the MEG response at a specific time, we cannot distinguish between early and late phases of a response of a particular region.

The method and results of this work provide a gateway to resolving the time course of visual processing to a variety of other visual stimuli. In effect, it may permit a denser sampling of object space than previously achieved[4,5]. For example, a description of the temporal dynamics of face representation in humans might be feasible with a large and rich parametrically modulated stimulus set and comparison to monkey electrophysiology[48]. Similarly, most MEG or EEG studies based on event-related potential analysis investigating content-specific modulation of brain activity by cognitive factors like memory or attention must rely on a handful of categorical markers in the event-related potential waveform—for example, the M100 and N170 for faces. In contrast, by applying multivariate methods, potentially any kind of content and the modulation of its representation by cognitive factors may be tractable, increasing experimental flexibility and generalizability of results. Thus, the application of multivariate analysis techniques to MEG[19] might be as fruitful in the future study of object recognition as the introduction of these techniques was in fMRI[18,20].

Resolving object processing in time, space and species

Relating signals measured in different imaging modalities and combining the methods' respective advantages are challenges in systems neuroscience[25]. Using representational similarity analysis, we showed a content-selective correspondence between early MEG signals and fMRI responses in V1 and between later MEG signals and fMRI responses in IT. Our results match previously reported average onset latencies in the literature, ranging between 50 and 80 ms in V1 (ref. 9) and 80 and 200 ms in IT[11,37]. Thus, representational similarity analysis combining MEG and fMRI is a promising method for relating cortical activations across space and time.

Comparing visual representations across time, we differentiated transient from persistent neural activity during object processing and found evidence for persistent activity in both V1 and IT. Thus, during object viewing the brain maintained both low- and high-level feature representations. These effects are most likely actively controlled processes, as indicated by their limited temporal extent. They might form the basis of memory of images in representational formats that make explicit different properties of these images—for example, low-level features versus category membership.

An integrated theory of object recognition requires quantitative bridging of the gap not only across imaging modalities but also across species. Using representational similarity analysis, it has been shown that human and monkey IT share a similar object coding scheme[3]. Here we have extended this finding by taking first steps in linking the dynamics in human MEG to single-cell activity in monkey IT. Note that in this experiment all temporal variance came from MEG: the dissimilarity matrix of monkey IT is based on averaged activity in IT cells 71–210 ms after stimulus onset[26]. Future studies might compare the dynamics in human and monkey IT using monkey data resolved in time, thus potentially complementing spatial homologies with temporal ones. In the meantime, the linkage between the time course of individual object coding in humans and coding of these same objects in monkey IT, although predictable by previous research, is shown here to our knowledge for the first time.

Conclusion

Progress in understanding how object recognition is implemented in the brain is likely to come from the combination of advances in data analyses suitable for different imaging techniques and comparison

across species[25]. Here we have provided key advances on two fronts. First, by applying multivariate pattern classification to human MEG signals, we showed the dynamics with which the brain processes objects at different levels of categorization and actively maintains visual representation. Second, we proposed an integrated space- and time-resolved view of the human brain during the first few hundred milliseconds of visual object processing and showed that representational similarity analysis allows brain signals in space, time and species to be understood in a common framework.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was funded by US National Eye Institute grant EY020484 (to A.O.), US National Science Foundation grant BCS-1134780 (to D.P.) and a Humboldt Scholarship (to R.M.C.), and was conducted at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research, Massachusetts Institute of Technology.

AUTHOR CONTRIBUTIONS

R.M.C., D.P. and A.O. designed the research. R.M.C. and D.P. performed experiments and analyzed the data. R.M.C., D.P. and A.O. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Grill-Spector, K. & Malach, R. The human visual cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).
2. Hung, C.P., Kreiman, G., Poggio, T. & DiCarlo, J.J. Fast readout of object identity from macaque inferior temporal cortex. *Science* **310**, 863–866 (2005).
3. Kriegeskorte, N. *et al.* Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron* **60**, 1126–1141 (2008).
4. Kourtzi, Z. & Connor, C.E. Neural representations for object perception: structure, category, and adaptive coding. *Annu. Rev. Neurosci.* **34**, 45–67 (2011).
5. DiCarlo, J.J., Zoccolan, D. & Rust, N.C. How does the brain solve visual object recognition? *Neuron* **73**, 415–434 (2012).
6. Felleman, D.J. & Van Essen, D.C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
7. Ungerleider, L.G. & Mishkin, M. Two visual systems. In *Analysis of Visual Behavior.* (eds. Ingle, D.J., Goodale, M.A. & Mansfield, R.J.W.) 549–586 (MIT Press, 1982).
8. Milner, A.D. & Goodale, M.A. *The Visual Brain in Action* (Oxford Univ. Press, 2006).
9. Schmolesky, M.T. *et al.* Signal timing across the macaque visual system. *J. Neurophysiol.* **79**, 3272–3278 (1998).
10. Luck, S.J. *An Introduction to the Event-Related Potential Technique* (MIT Press, 2005).
11. Mormann, F. *et al.* Latency and selectivity of single neurons indicate hierarchical processing in the human medial temporal lobe. *J. Neurosci.* **28**, 8865–8872 (2008).
12. Baillet, S., Mosher, J.C. & Leahy, R.M. Electromagnetic brain mapping. *IEEE Signal Process. Mag.* **18**, 14–30 (2001).
13. Hari, R. & Salmelin, R. Magnetoencephalography: from SQUIDs to neuroscience: *Neuroimage* 20th anniversary special edition. *Neuroimage* **61**, 386–396 (2012).
14. Dale, A.M. *et al.* Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* **26**, 55–67 (2000).
15. Debener, S., Ullsperger, M., Siegel, M. & Engel, A.K. Single-trial EEG–fMRI reveals the dynamics of cognitive function. *Trends Cogn. Sci.* **10**, 558–563 (2006).
16. Logothetis, N.K. & Sheinberg, D.L. Visual object recognition. *Annu. Rev. Neurosci.* **19**, 577–621 (1996).
17. Carlson, T.A., Hogendoorn, H., Kanai, R., Mesik, J. & Turret, J. High temporal resolution decoding of object position and category. *J. Vis.* **11**(10): 9 (2011).
18. Haynes, J.-D. & Rees, G. Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* **7**, 523–534 (2006).
19. Carlson, T., Tovar, D.A., Alink, A. & Kriegeskorte, N. Representational dynamics of object vision: the first 1000 ms. *J. Vis.* **13**(10): 1 (2013).
20. Tong, F. & Pratte, M.S. Decoding patterns of human brain activity. *Annu. Rev. Psychol.* **63**, 483–509 (2012).
21. Thorpe, S., Fize, D. & Marlot, C. Speed of processing in the human visual system. *Nature* **381**, 520–522 (1996).
22. Bentin, S., Allison, T., Puce, A., Perez, E. & McCarthy, G. Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* **8**, 551–565 (1996).
23. VanRullen, R. & Thorpe, S.J. The time course of visual processing: from early perception to decision-making. *J. Cogn. Neurosci.* **13**, 454–461 (2001).
24. Edelman, S. Representation is representation of similarities. *Behav. Brain Sci.* **21**, 449–467, discussion 467–498 (1998).
25. Kriegeskorte, N. Representational similarity analysis – connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 4 (2008).
26. Kiani, R., Esteky, H., Mirpour, K. & Tanaka, K. Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J. Neurophysiol.* **97**, 4296–4309 (2007).
27. Nichols, T.E. & Holmes, A.P. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* **15**, 1–25 (2002).
28. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).
29. Kruskal, J.B. & Wish, M. Multidimensional scaling. University Paper Series on Quantitative Applications in the Social Sciences, Series 07-011 (Sage Publications, 1978).
30. Shepard, R.N. Multidimensional scaling, tree-fitting, and clustering. *Science* **210**, 390–398 (1980).
31. Allison, T. *et al.* Face recognition in human extrastriate cortex. *J. Neurophysiol.* **71**, 821–825 (1994).
32. Kanwisher, N., McDermott, J. & Chun, M.M. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* **17**, 4302–4311 (1997).
33. McCarthy, G., Puce, A., Belger, A. & Allison, T. Electrophysiological studies of human face perception. II: Response properties of face-specific potentials generated in occipitotemporal cortex. *Cereb. Cortex* **9**, 431–444 (1999).
34. Downing, P.E., Jiang, Y., Shuman, M. & Kanwisher, N. A cortical area selective for visual processing of the human body. *Science* **293**, 2470–2473 (2001).
35. Liu, J., Harris, A. & Kanwisher, N. Stages of processing in face perception: an MEG study. *Nat. Neurosci.* **5**, 910–916 (2002).
36. Harrison, S.A. & Tong, F. Decoding reveals the contents of visual working memory in early visual areas. *Nature* **458**, 632–635 (2009).
37. Liu, H., Agam, Y., Madsen, J.R. & Kreiman, G. Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron* **62**, 281–290 (2009).
38. Stekelenburg, J.J. & de Gelder, B. The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. *Neuroreport* **15**, 777–780 (2004).
39. Thierry, G. *et al.* An event-related potential component sensitive to images of the human body. *Neuroimage* **32**, 871–879 (2006).
40. Jeffreys, D.A. Evoked potential studies of face and object processing. *Vis. Cogn.* **3**, 1–38 (1996).
41. Halgren, E., Raij, T., Marinkovic, K., Jousmäki, V. & Hari, R. Cognitive response profile of the human fusiform face area as determined by MEG. *Cereb. Cortex* **10**, 69–81 (2000).
42. Sadeh, B., Podlipsky, I., Zhdanov, A. & Yovel, G. Event-related potential and functional MRI measures of face-selectivity are highly correlated: a simultaneous ERP-fMRI investigation. *Hum. Brain Mapp.* **31**, 1490–1501 (2010).
43. Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H. & Livingstone, M.S. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670–674 (2006).
44. Mack, M.L. & Palmeri, T.J. The timing of visual object categorization. *Front. Psychol.* **2**, 165 (2011).
45. Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G. & Mishkin, M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **17**, 26–49 (2013).
46. Sugase-Miyamoto, Y., Matsumoto, N. & Kawano, K. Role of temporal processing stages by inferior temporal neurons in facial recognition. *Front. Psychol.* **2**, 141 (2011).
47. Brincat, S.L. & Connor, C.E. Dynamic shape synthesis in posterior inferotemporal cortex. *Neuron* **49**, 17–24 (2006).
48. Freiwald, W.A. & Tsao, D.Y. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* **330**, 845–851 (2010).

ONLINE METHODS

Participants and experimental design. Sixteen right-handed, healthy volunteers with normal or corrected-to-normal vision (10 female, age: mean ± s.d. = 25.87 ± 5.38 years) participated in the experiment. The study was conducted according to the Declaration of Helsinki and approved by the local ethics committee (Institutional Review Board of the Massachusetts Institute of Technology). Fifteen participants completed two MRI and MEG sessions, and one participant participated in the MEG experiment only. The sample size is comparable to that used in previous fMRI and MEG studies. All participants provided written consent for each of the sessions. During the experiment participants saw images of 92 different objects presented at the center of the screen (2.9° visual angle, 500 ms duration) overlaid with a dark gray fixation cross. We chose this particular data set for two reasons. First, it allowed assessment of distinctions at three levels: superordinate, ordinate and subordinate categories. Second, it enabled direct comparison of our MEG and fMRI results with previous experiments using the same date set in monkey electrophysiology and human MRI[3,26]. The presentation parameters were adapted to the specific requirements of each acquisition technique (**Supplementary Fig. 1**). In detail, for each MEG session, participants completed 10 to 15 runs, each having duration 420 s. Each image was presented twice in each MEG run in random order, with a trial onset asynchrony (TOA) of 1.5 or 2 s. Participants were instructed to press a button and blink their eyes in response to a paper clip that was shown randomly every 3 to 5 trials (average 4). For each fMRI session, participants completed 10 to 14 runs, each having duration 384 s. Each image was presented once in each run in random order, with the restriction of not displaying the same condition on consecutive trials. Thirty null trials with no stimulus presentation were randomly interspersed, during which the fixation cross turned darker for 100 ms and participants reported the change with a button press. TOA was 3 s, or 6 s in the presence of a null trial.

Human MEG acquisition. We acquired continuous MEG signals from 306 channels (204 planar gradiometers, 102 magnetometers, Elekta Neuromag TRIUX, Elekta, Stockholm) at a sampling rate of 1,000 Hz, filtered between 0.03 and 330 Hz. Raw data were preprocessed using spatiotemporal filters (maxfilter software, Elekta, Stockholm) and then analyzed using Brainstorm[49]. MEG trials were extracted with 100 ms baseline and 1,200 ms post-stimulus (i.e., 1,301 ms length), the baseline mean of each channel was removed, and data were temporally smoothed with a 20-ms sliding window. A total of 20–30 trials were obtained for each condition, session and participant.

Multivariate analysis of MEG data. To determine the amount of object image information contained in MEG signals, we employed multivariate analysis in the form of linear support vector machines (SVMs; libsvm: http://www.csie.ntu.edu.tw/~cjlin/libsvm/)[50]. The SVM analysis was conducted independently for each participant and session (**Fig. 1a**,**b**). For each time point (100 ms before to 1,200 ms after image onset), MEG data were arranged in the form of 306 dimensional measurement vectors, yielding N pattern vectors per time point and condition (image). We used supervised learning, with a leave-one-out cross-validation approach, to train the SVM classifier to pairwise decode any two conditions. Namely, for each time point and pair of conditions, $N - 1$ pattern vectors comprised the training set and the remaining Nth pattern vectors the testing set, and the performance of the classifier to separate the two conditions was evaluated. The process was repeated 100 times with random reassignment of the data to training and testing sets, yielding an overall decoding accuracy of the classifier. The decoding accuracy was then assigned to a decoding accuracy matrix of size 92×92, with rows and columns indexed by the conditions classified. The matrix is symmetric across the diagonal, with the diagonal undefined. This procedure yielded one 92×92 matrix of decoding accuracies for every time point.

Visualization and exploration using multidimensional scaling. The 92×92 MEG decoding matrices contained complex high-dimensional structure that was difficult to visualize. To reveal any underlying patterns, we used multidimensional scaling (MDS)[29,30] to plot the data into a two-dimensional space of the first two dimensions of the solution, such that similar conditions were grouped together and dissimilar conditions far apart. MDS is an unsupervised method to visualize the level of similarity of individual objects contained in a distance matrix (here the decoding matrix), whereby objects are automatically assigned coordinates in space so that distances between objects are preserved.

MDS was applied to the whole or part of the decoding matrix, depending on the conditions explored. To avoid double-dipping[51] the data, we computed MDS at peak-latency time points with a leave-one-participant-out approach as follows: all but one participant were used to identify the peak latency time and the remaining participant provided the decoding matrix. The decoding matrix was averaged across all permutations, and only the overall decoding matrix was subjected to MDS.

Human fMRI acquisition. Magnetic resonance imaging (MRI) was conducted on a 3T Trio scanner (Siemens, Erlangen, Germany) with a 32-channel head coil. We acquired structural images using a standard T1-weighted sequence (192 sagittal slices, FOV = 256 mm^2, TR = 1,900 ms, TE = 2.52 ms, flip angle = 9°). For fMRI, we conducted 10–14 runs in which 192 volumes were acquired for each participant (gradient-echo EPI sequence: TR = 2,000 ms, TE = 31 ms, flip angle = 80°, FOV read = 192 mm, FOV phase = 100%, ascending acquisition, gap = 10%, resolution = 2 mm isotropic, slices = 25). The acquisition volume covered the occipital and temporal lobe and was oriented parallel to the temporal cortex.

Human fMRI analysis. fMRI data were processed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). For each participant and session separately, data were realigned and slice-time corrected, and then co-registered to the T1 structural scan acquired in the first MRI session. We neither normalized nor smoothed fMRI data. We then modeled the fMRI responses to the 92 images with a general linear model (GLM) in two independent models: one comprising only the first three runs of each session and one comprising the remaining runs. The onsets and durations of each image presentation, as well as those of the null trials, were entered into the GLM as regressors and convolved with a hemodynamic response function. Movement parameters entered the GLM as nuisance regressors. For each of the 92 image conditions, we converted GLM parameter estimates into t-values by contrasting each condition estimate against the explicitly modeled baseline. In addition, we assessed the effect of visual stimulation irrespective of condition in a separate t-contrast by contrasting the parameter estimates for all 92 images against baseline.

fMRI region of interest definition. We defined V1 separately for each participant on the basis of an anatomical eccentricity template[52]. The cortical surface of each participant was reconstructed with FreeSurfer on the basis of the T1 structural scan[53]. The right hemisphere was mirror-reversed and registered to the left hemisphere. This allowed us to register a V1 eccentricity template[52] to participant-specific surfaces and to define surface-based regions of interest (ROIs) corresponding to 0–3° and 3–6° visual angle, termed here central V1 and peripheral V1. These surface-based ROIs were resampled to the space of EPI volumes and combined in a common ROI for both cortical hemispheres.

For human IT, we used a mask consisting of bilateral fusiform and inferior temporal cortex (WFU Pickatlas, IBASPM116 Atlas[54]). Anatomical masks were reverse-normalized from MNI-space to single-participant space. To match the size of IT to the average size of central V1, we restricted the definition of IT for each participant and session: we considered only the 361 most strongly activated voxels in the t-contrast of all conditions against baseline in the GLM based only on the first three runs; we used only these voxels to extract t-values for each of the 92 images from the remaining runs for further analysis.

fMRI pattern analysis. We used a correlation-based method to determine the relations between brain fMRI responses to the 92 images (**Fig. 4**). Observations were formed from each ROI (central and peripheral V1, and IT), by extracting and concatenating the corresponding voxel fMRI activation values into pattern vectors. For every pair of the 92 conditions, we then computed the Spearman's rank-order correlation coefficient R between the corresponding pattern vectors of a given ROI and stored the result in a 92×92 symmetric matrix. We converted the correlations into a dissimilarity measure[3] $1 - R$, which is bounded between 0 (no dissimilarity) and 2 (complete dissimilarity). For further analyses, we averaged the dissimilarity matrices across sessions and participants, resulting in one matrix for each ROI.

Monkey electrophysiology. The details of monkey electrophysiological recordings and representational similarity analysis are described elsewhere[26]. In short, two awake macaque monkeys were presented with the same stimulus set as the

one used in our MEG and fMRI human experiments. Images spanned 7° visual angle and were presented in a rapid design (105 ms on, ISI = 0s) among a larger set of images while the monkeys maintained fixation. Single-cell responses in 674 neurons were recorded extracellularly from anterior inferior temporal cortex. Cell responses to each image were estimated as average spike rate in a 71–210 ms time window after stimulus onset. Representational dissimilarity matrices were created by pairwise-correlating responses to images across the 674 neurons (Pearson's product-moment correlation) and subtracting the resulting value from 1. The representational dissimilarity matrices were generously supplied to us by N. Kriegeskorte and R. Kiani[3,26].

Significance testing. We used non-parametric statistical inference[27,28], which does not make assumptions about the distribution of the data, for random-effects inference. Permutation tests were used for cluster-size inference, and bootstrap tests for confidence intervals on (1) maxima and cluster onsets/offsets and (2) peak-to-peak latency differences. The sample size (N) was always 16, and all tests were two-sided.

Permutation tests. The null hypothesis of no experimental effect differed throughout the paper depending on the analysis of interest: the MEG decoding time series was equal to 50% chance level; the within-subdivision minus between-subdivision portions of an MEG decoding matrix was equal to 0; the correlation of the MEG decoding matrices and fMRI (or monkey spiking activity) dissimilarity matrix was equal to 0. In all cases, under the null hypothesis we could permute the condition labels of the MEG data, which effectively corresponds to a sign permutation test that randomly multiplies the participant-specific data (for example, MEG decoding accuracies or correlations) with +1 or −1. For each MEG permutation sample, we recomputed the statistic of interest. Repeating this procedure 50,000 times, we obtained an empirical distribution of the data, which allowed us to convert our statistics (for example, MEG decoding time series, time-time decoding matrices, etc.) into one-dimensional or two-dimensional P-value maps.

Familywise error rate was then controlled with cluster-size inference. The P-value maps of the original data were thresholded at $P < 0.001$ for one dimension and $P < 0.0001$ for two dimensions to define suprathreshold clusters. These suprathreshold clusters were reported significant only if their size exceeded a threshold, estimated as follows: the previously computed permutation samples were also converted to P-value maps (relying on the same empirical distribution as the original data) and also thresholded to define resampled versions of suprathreshold clusters. These clusters were used to construct an empirical distribution of maximum cluster size and to estimate a threshold at 5% of the right tail of this distribution (that is, the corrected P value is $P < 0.05$).

Bootstrap tests. We calculated 95% confidence intervals for the onsets of the first significant cluster and the peak latency of the observed effects. To achieve this, we created 1,000 bootstrapped samples by sampling the participants with replacement. For each bootstrap sample, we repeated the exact data analysis as the original data (including the permutation tests), resulting in bootstrap estimates of onsets and peak latencies and thus the determination of their 95% confidence intervals.

To calculate confidence intervals on mean peak-to-peak latency differences, we created 50,000 bootstrapped samples by sampling the participant-specific latencies with replacement. This yielded an empirical distribution of mean peak-to-peak latencies. We set $P < 0.05$, Bonferroni-corrected. If the 95% confidence interval did not include 0, we rejected the null hypothesis of no peak-to-peak latency differences.

49. Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D. & Leahy, R.M. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* **2011**, 879716 (2011).
50. Müller, K.R., Mika, S., Rätsch, G., Tsuda, K. & Schölkopf, B. An introduction to kernel-based learning algorithms. *IEEE Trans. Neural Netw.* **12**, 181–201 (2001).
51. Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F. & Baker, C.I. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* **12**, 535–540 (2009).
52. Benson, N.C. *et al.* The retinotopic organization of striate cortex is well predicted by surface topology. *Curr. Biol.* **22**, 2081–2085 (2012).
53. Dale, A.M., Fischl, B. & Sereno, M.I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).
54. Maldjian, J.A., Laurienti, P.J., Kraft, R.A. & Burdette, J.H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* **19**, 1233–1239 (2003).

doi:10.1038/nn.3635