Supplementary Figure 1  fMRI responses separated to account for between-subjects variability in cue weights. fMRI data is expressed as a rebound index for three groups of subjects (Supplementary Table 1). Rebound index equals peak fMRI response in a condition divided by the response to the Identical condition. Graphs in the inserts show mean psychophysical response index. Subjects who weight disparity and perspective similarly (high disparity group) have lowest fMRI responses in the Balanced Change condition where the perceived 3D shape of the test stimulus (PSE = 125°) is very close to the reference (S_p = S_d = 126°), suggesting adaptation to the perceived shape. Complementing this, similar responses are observed in the Disparity and Perspective Change conditions where the perceived 3D shape change is comparable (11° and 9° respectively). For the Medium Disparity group, fMRI responses are high in the Perspective Change condition ( = 15°), whilst responses in Balanced ( = 5°) and Disparity Change ( = 5°) conditions are similar. Finally, subjects with low disparity weight, showed the lowest responses in the Disparity Change ( = 2°) condition but high responses in the Balanced ( = 8°) and Perspective Change ( = 18°) conditions.

Supplementary Figure 2  Eye movement controls. (a) Event-related eye movement traces for three subjects. Mean horizontal right eye position during stimulus presentation (950ms) in the inconsistent-cue experiment is shown. Eye position was measured whilst subjects (n = 3) performed the discrimination task inside the magnet by limbus tracking under infra-red illumination (CRS monocular Eyetracker: temporal resolution 1kHz, quoted spatial resolution < 0.2°). Eye traces were processed by first removing blinks and saccades (> 0.5°) by visual inspection. There were no significant differences between conditions in the number of blinks (mean = 5.4; per subject Chi-sqr, uncorrected, P > 0.05) or saccades (mean = 4.6; per subject Chi-sqr, uncorrected, P > 0.25). Traces for individual trials were brought to a common baseline (to remove drift) by subtracting the mean eye position over the first 100ms of each trial. Eye traces for each trial were averaged (14 < n ≤ 24, depending on the number of blinks and saccades) to produce an event-related eye trace for each subject. Six conditions are shown: the 5 experimental conditions and fixation. Separate plots show results for each observer. Two of the observers show a small (< 0.5°) deviation in mean eye position for the experimental conditions compared with the fixation baseline condition. This deviation cannot account for the experimental results as it is not systematically different between experimental conditions. (b) We replicated Experiment 1 (n = 4) when a pair of nonius lines were presented each side of the fixation point to help promote correct vergence alignment. fMRI responses are expressed as a rebound index. Error bars show normalized SEM; they are large because we used a conservative method (Taylor) that incorporates the error
estimates of both the numerator and the denominator. Insert shows between-subjects mean psychophysical behavior expressed as a response index (difference between response proportion and 0.5). Error bars ± SEM.

Supplementary Figure 3 Results from an experiment on consistent-cue stimuli. We examined fMRI responses when observers (n = 8) discriminated consistent-cue stimuli using an event-related adaptation design. We used a test stimulus with consistent cues ($S_p = S_q = 126^\circ$), and examined four experimental conditions: Identical ($Test = 126^\circ; Ref = 126^\circ$); Plus 10 ($Test = 126^\circ; Ref = 136^\circ$); Minus 10 ($Test = 126^\circ; Ref = 116^\circ$); Minus 20 ($Test = 126^\circ; Ref = 106^\circ$). Data correspond to the averaged peak time points of the fMRI time course for each condition (hemispheres combined). Error bars depict ± SEM. The insert shows mean between-subjects psychophysical response index for each experimental condition; error bars ± SEM. fMRI responses in the higher visual areas (LOC and hMT+/V5) showed a similar pattern to the observers' psychophysical responses. Specifically, responses in the high (psychophysical response > 0.25: Minus 20) discriminability condition were significantly higher (LOC: $F_{2,147} = 1.10, P = 0.05$; hMT+/V5: $F_{2,147} = 2.34, P = 0.05$) than low discriminability conditions (plus 10 and minus 10), whereas no differences (LOC: $F_{1,147} < 1, P = 0.14$; hMT+/V5: $F_{1,147} < 1, P = 0.10$) were observed between fMRI responses obtained in equally discriminable conditions (plus 10 vs. minus 10). No significant differences were observed in the early visual areas either between (i) high and low discriminability conditions (V1: $F_{2,147} < 1, P = 0.14$; V2: $F_{2,147} < 1, P = 0.13$; V3: $F_{2,147} < 1, P = 0.11$; V3a: $F_{2,147} < 1, P = 0.12$; Vp: $F_{2,147} < 1, P = 0.11$; V4: $F_{2,147} < 1, P = 0.15$) or between (ii) equally discriminable conditions (V1: $F_{1,147} < 1, P = 0.10$; V2: $F_{1,147} < 1, P = 0.18$; V3: $F_{1,147} < 1, P = 0.10$; V3a: $F_{1,147} < 1, P = 0.25$; Vp: $F_{1,147} < 1, P = 0.17$; V4: $F_{1,147} < 1, P = 0.34$). These data support the hypothesis that responses in higher visual areas relate to the perceived 3D shape, and indicate that this finding is not limited to inconsistent-cue stimuli.

Supplementary Figure 4 Additional example flatmaps showing examined ROIs. Functional activation maps for two subjects showing the early retinotopic regions, hMT+/V5 and the LOC (comprising the regions labeled pFs and LO). The functional regions are superimposed on flattened cortical surfaces of the right and left hemispheres. The sulci are coded in darker grey than the gyri and the Anterior-Posterior orientation is noted by A and P. Major sulci are labeled: STS: superior temporal sulcus, ITS: inferior temporal sulcus, OTS: occipitotemporal sulcus, CoS: collateral sulcus. The LOC was defined as the voxels in ventral occipitotemporal cortex showing significantly stronger activation ($P < 10^{-4}$, corrected) to intact than to scrambled images of objects. hMT+/V5 was defined as the contiguous voxels in the ascending limb of the inferior temporal sulcus.