

The Coding of Color, Motion, and Their Conjunction in the Human Visual Cortex

Kiley Seymour,¹ Colin W.G. Clifford,¹ Nikos K. Logothetis,^{2,3} and Andreas Bartels^{2,*}

¹School of Psychology

University of Sydney

Sydney NSW 2006

Australia

²Max Planck Institute for Biological Cybernetics

72076 Tuebingen

Germany

³Imaging Science and Biomedical Engineering

University of Manchester

Manchester

UK

Summary

Background: Color and motion serve as the prime examples of segregated processing in the visual brain, giving rise to the question how color-motion conjunctions are represented. This problem is also known as the “binding problem.”

Results: Human volunteers viewed visual displays containing colored dots rotating around the center. The dots could be red or green and rotate clockwise or counterclockwise, leading to four possible stimulus displays. Superimposed pairs of such stimuli provided two additional displays, each containing both colors and both directions of motion but differing in their feature conjunctions. We applied multivariate classifiers to voxel-activation patterns obtained while subjects viewed such displays. Our analyses confirm the presence of directional-motion information across visual cortex and provide evidence of hue coding in all early visual areas except V5/MT⁺. Within each cortical area, information on color and motion appeared to be coded in distinct sets of voxels. Furthermore, our results demonstrate the explicit representation of feature conjunctions in the primary visual cortex and beyond.

Conclusions: The results show that conjunctions can be decoded from spatial activation patterns already in V1, indicating an explicit coding of conjunctions at early stages of visual processing. Our findings raise the possibility that the solution of what has been taken as the prime example of the binding problem engages neural mechanisms as early as V1.

Introduction

One of the most prominent examples of functional specialization in the primate brain is the processing of color and motion. These two features have been shown to be processed in largely distinct, yet mutually connected processing streams [1, 2]. Color is primarily processed in the blobs of V1, in the thin stripes of V2, in the human V4 complex, and in regions anterior to it, whereas motion is primarily processed in layer 4B of V1, in the thick stripes of V2, and in the V5/MT⁺ complex and regions beyond it [3–6]. Although information on both features is present in V1, V2, and V4, it appears to be largely segregated

at the cellular level [4, 7, 8]. There is seemingly no evidence for chromatically selective neurons in V5/MT⁺, although the area does have reciprocal connections with V4 and some of its neurons can respond to moving isoluminant edges [3, 9].

The physiological evidence for this segregation is confirmed in a causal way by patient studies, showing that lesions in the vicinity of V4 impair color perception but spare motion perception, whereas the opposite is true for lesions to the V5/MT complex [10, 11]. The two systems not only differ in their anatomical locations and functional properties but also exhibit different conduction velocities and processing latencies [12]. A perceptual consequence of this segregation may be the temporal asynchrony of color and motion perception [13, 14], as well as the observed slow binding across compared to within these features [15].

Functional segregation is directly linked to the fundamental question of how the brain combines different attributes, leading to an apparently coherent percept of our visual world—the so-called binding problem [16, 17]. Some observations from psychophysics, such as the existence of monocular color-contingent motion aftereffects and the fast combination of color and orientation information, suggest that some visual features may be coded in combination as early as in area V1 where neurons with various functional properties are closely spaced and thus potentially well suited for functional integration [18–20]. However, differential delays in perception [13], slow across-feature binding compared to within feature binding [15], and slow performance in visual conjunction search [16] suggest that slower, attention-driven mechanisms may also be involved in combining features. It is therefore still unclear at which levels in the visual system information about feature conjunctions can be found. It also remains a matter of debate whether visual-feature binding is mediated by a temporal code [17], by communication between visual areas [21], by feedback connections to early visual areas [22–24], or by representations at higher, cognitive stages [16].

Here, we ventured to ask a question that has not yet received a clear answer in physiology or imaging, namely which levels of the visual system contain information regarding the conjunctions of features. In other words, we not only asked whether areas code for the color and the direction of motion of a visual object but also whether they code explicitly for the particular color-motion pairing defining the object. We used high-resolution imaging and multivariate statistics, allowing us to retrieve information present in cortical regions from consistent patterns of voxel activations [25]. This multivariate approach in human fMRI has confirmed physiological findings of direction selectivity in V5/MT⁺ [26] and of orientation selectivity in V1 and elsewhere [27–29]. We explored the coding of visual-feature conjunctions using two transparent motion stimuli that each contained the same two colors and two motion directions but differed exclusively in their unique feature pairings. Using linear classifiers, we found that information about color and motion was present to varying extents across the visual cortex, with the exception that no significant color information was present in V5/MT⁺. Notably, we found evidence for conjunction coding as early as V1 and distributed across the entire visual cortex.

*Correspondence: andreas.bartels@tuebingen.mpg.de

Stimulus conditions and coding logic

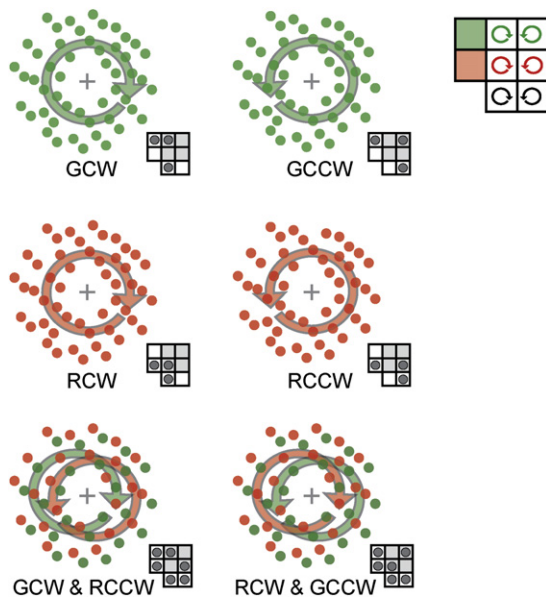


Figure 1. Schematic Representation of the Six Stimulus Conditions Used in the Experiment, Along with a Coding Logic of Potential Voxel-Wise Responses. The top four conditions are single conjunctions, containing one color and one motion direction each. The bottom two conditions are double conjunctions, each containing both colors and both motion directions. The diagram at the top right represents all possible feature and conjunction specific responses: the two units on the left represent one color each, the two on the bottom one motion direction each, and the remaining respond to one feature conjunction each. The gray-scale diagrams next to each stimulus indicate the resulting responses. Note that the double-conjunction stimuli would be indistinguishable without conjunction-specific responses, as all four feature specific units are active in both conditions.

Results

Univariate Analysis

A conventional univariate analysis (consisting of stimulus responsive voxels in V1, V2, V3, V3A/B, V4, and V5/MT⁺) of each subject revealed no significant response bias to a particular color, motion direction, or conjunction stimulus at a Bonferroni corrected threshold of $p < 0.05$ (see [Figure 1](#) for stimuli). [Figure S1](#) (available online) shows for each ROI of a subject the voxel-wise distribution of the t value differences between the two double-conjunction stimuli. This provides a measure for potential biases to the double-conjunction stimuli. The distribution of these biases across all voxels followed an approximately Gaussian distribution and showed no significant ($p < 0.05$) bias toward either stimulus. In particular, no bias was observed when this analysis was performed separately for each of the four quadrants of the visual-field representations of the visual areas. The above analyses therefore revealed that neither large-scale spatial inhomogeneities nor strong voxel-wise biases at the univariate level could explain subsequent classification performance at the multivariate level. Therefore, we continued with a multivariate approach testing whether voxel-activation patterns within ROIs discriminated between the features.

Decoding of Motion Directions

First, we tested whether linear support vector machines (SVMs) could discriminate between voxel-activity patterns

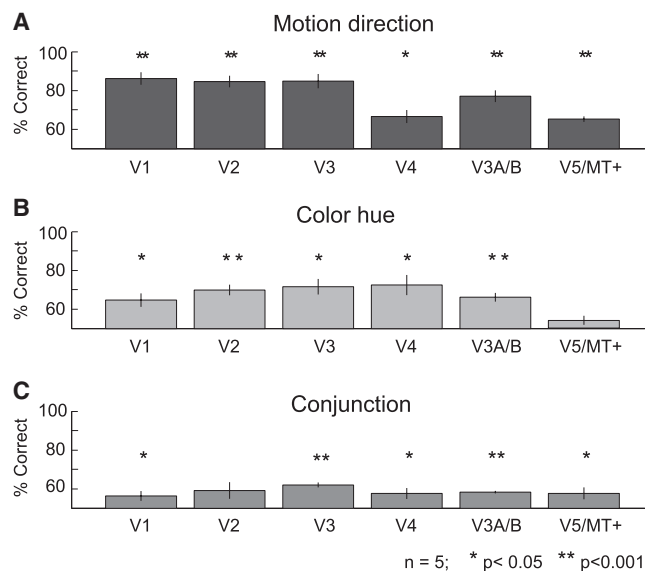


Figure 2. Discrimination Performance of a Linear Classifier in Distinguishing Motion Directions, Colors, or Double Conjunctions on the Basis of Voxel Patterns Evoked in Visual Areas

(A) Decoding of motion directions (CW versus CCW) by a linear SVM was significantly above chance (chance = 50%) in every ROI, averaged across five subjects.

(B) Decoding of color hue (red versus green) was significant in all areas except for V5/MT⁺.

(C) Decoding of double-conjunction stimuli (GCW + RCCW versus RCW + GCCW) reached significance in most areas, including V1. Error bars represent between subjects' standard error (n = 5 subjects).

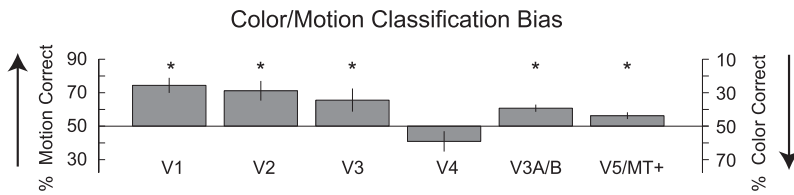
evoked by stimuli containing clockwise or counterclockwise motion, regardless of stimulus color (i.e., between blocks of RCW and GCW on the one hand and RCCW and GCCW on the other). This analysis was conducted separately for each ROI of each subject. Across all subjects, motion direction was decoded in V1, V2, V3, V3A/B, V4, and V5/MT⁺ with performance ranging from 57% to 93%. Each area performed better than the chance level of 50% at a significance of $p < 0.005$ when tested across n = 5 subjects ([Figure 2A](#)). Best performance was achieved in V1 with a mean of 86% ($p = 0.0002$). Area V5/MT⁺ achieved 65% ($p = 0.0004$), and V4 achieved 67% ($p = 0.004$). Because of differences in the size of each ROI (range 233 to 2189 voxels), and the potential differences underlying functional architecture in each region (which would affect voxel biases and thus classifier performance), we shall avoid making any direct comparison between ROIs beyond the inference of presence of motion information. When we restricted our analysis to the best (or a random selection of) 233 voxels in each ROI, a similar pattern of results was obtained (see [Figures S2 and S3](#)). Single-subject performances are reported in [Figure S4](#).

Decoding of Color Hues

When SVMs were trained to distinguish red (blocks of RCW and RCCW) from green (blocks of GCW and GCCW), significant color discrimination was achieved in all areas of the visual cortex with the exception of V5/MT⁺ ([Figure 2B](#)). In V5/MT⁺, color discrimination still failed to reach significance when only trials of one motion direction were used (e.g., RCW against GCW) (see below and [Figure 4](#)). Across all ROIs of all subjects, prediction accuracy in areas V1–V4 ranged from

Cortical Coding of Color-Motion Conjunctions

3



information). All areas except for V4 showed a significant bias toward motion classification. Classification based on V4 (biased toward color) differed significantly from that based on V5/MT+ (biased toward motion) ($p < 0.05$). Error bars represent between subjects' standard error ($n = 5$ subjects).

57% to 86%. Mean prediction performance in V1 was 65% ($p = 0.007$), whereas V4 achieved 73% ($p = 0.006$) and V5/MT+ achieved 54% ($p = 0.072$). When we restricted our analysis to the best (or a random set of) 233 voxels in each ROI, a similar pattern of findings was observed (Figures S2 and S3).

Color and Motion Interactions: Biases for and Generalization across Features

To determine whether color or motion signals dominated voxel-activation patterns and thus classification, and whether distinct ROIs differed in this bias toward motion or color coding, we trained an SVM to discriminate between RCW and GCCW trials but tested it on GCW and RCCW trials (and vice versa). In this case, the SVM could either get the motion direction correct but color incorrect, or vice versa. Figure 3 reports percent-correct motion classification (50% indicating no bias, greater than 50% signifying a bias toward motion, and less than 50% signifying a bias toward color). With the exception of V4, all areas of the visual cortex favored a correct classification of motion rather than color (e.g., V1: 74%, $p = 0.003$; V5/MT+: 56%, $p = 0.018$). Area V4 was the only region observed with a trend, albeit a nonsignificant one, to favor the classification of color over motion (41%, $p = 0.103$, $n = 5$). A paired t test revealed a significant difference in bias between area V4 (in favor of color) and area V5/MT+ (motion) ($n = 5$, $p = 0.036$). Note that in this analysis, absolute values (including the 50% line) are less informative than the differences between regions. For example, a bias toward classifying motion may not necessarily imply a stronger selectivity for motion in a region but could equally well be considered to reflect the relative signal strength of color and motion in the stimuli. Therefore, comparisons between regions are meaningful here for the assessment of differential feature biases.

Finally, we tested to what extent motion discrimination performance could generalize over different test colors (Figure 4A) and, conversely, the extent to which color decoding was affected by motion directions (Figure 4B). For instance, we trained a classifier to discriminate motion only on red dots (RCW versus RCCW) and then measured its performance testing first on red then on green dots (GCW versus GCCW). Our analysis showed that color did not have a significant influence on motion-decoding performance in any visual area (Figure 4A). Similarly, we found no significant difference in color-decoding performance when an SVM was forced to generalize color predictions across motion directions (e.g., train GCW versus RCW and test GCCW versus RCCW) (Figure 4B).

Motion and Color Conjunction Coding

We directly assessed conjunction coding using stimuli that each contained two conjunctions (RCW/GCCW versus RCCW/GCW trials). This way each stimulus contained both clockwise and counterclockwise motion and both red and

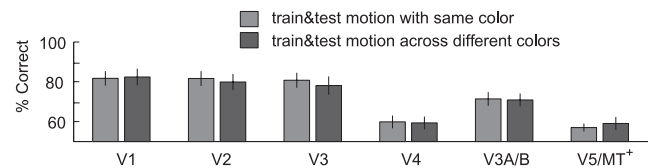
Figure 3. Classification Bias to Color or Motion in Visual Cortical Regions

The classifier was trained on voxel responses to one set of single-conjunction stimuli, e.g., RCW versus GCCW, and tested on another, e.g., GCW versus RCCW. The classifier thus faced the dilemma of making a correct motion classification at the expense of an incorrect color classification or vice versa (50% = chance, i.e., the classifier relied equally often on color and motion

green color, with the only difference being the two unique feature conjunctions that were specific to each stimulus (see Figure 1). We found that SVMs could discriminate responses to distinct conjunctions in all areas of visual cortex, with V1, V3, V4, V3A/B, and V5/MT+ reaching significance across all subjects ($p < 0.05$, $n = 5$) (Figure 2C). V1 decoded color-motion conjunctions with a mean prediction accuracy of 56% ($p = 0.036$). Areas V3 and V3A/B reached highest significance (mean 62%, $p = 0.0004$, and 59%, $p = 0.0001$, respectively), area V4 achieved 59%, ($p = 0.02$), and V5/MT+ achieved 56% ($p = 0.03$). The pattern of results was similar when a random set or the best 233 voxels were used for each ROI (Figures S2 and S3).

To confirm that classification performance could not be biased by luminance differences between the two colors (that we had attempted to exclude by individual isoluminance settings and the pseudo-random $\pm 10\%$ luminance offsets between blocks, see Experimental Procedures), we also tested a conjunction-trained classifier on discriminating motion directions. If the classifier were to use motion information to decode conjunction stimuli based on the stronger luminance of one motion direction, we would expect that when tested on the two motion directions (pooled across color), the SVM should

A Motion decoding: invariance to color



B Color decoding: invariance to motion

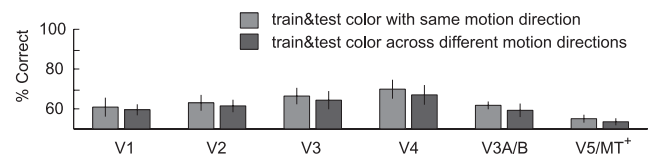


Figure 4. Invariance of Motion-Related Voxel Patterns to Color and Vice Versa

(A) The classifier was trained to discriminate voxel responses to motion directions from stimuli of one color and then tested on responses to motion stimuli of the same or different color. Classifier performance was not affected when training and test stimuli had different colors.

(B) For color discrimination, the classifier was trained to discriminate voxel responses to red or green stimuli moving in one direction and then tested on responses to color stimuli moving in the same or the opposite direction. Classifier performance was not affected when training and test stimuli had different motion directions. Error bars represent between subjects' standard error ($n = 5$ subjects).

produce above chance performance. We found that all areas gave chance performance at decoding motion direction in this analysis (mean across ROIs = $50.1\% \pm 1.8\%$, $n = 5$). The same analysis also rules out the conceivable scenario of biased activations in response to a perceptual dominance of one set of dots in the double-conjunction stimuli (even though red and green dots randomly occluded each other, and no subject—nor the authors—perceived a saliency asymmetry between features). Finally, this analysis also shows that no artifactual above-chance performance results from our methodology in our data. In addition to this, we also created an “artificial” double-conjunction data set by averaging responses of the single-conjunction stimuli (GCW + RCCW) and (RCW + GCCW). A classifier trained and tested on this data set discriminated the two stimulus classes with similar performance to that obtained with the real double-conjunction stimuli, as shown in supplementary figure S5. This demonstrates the presence of conjunction information also in single-conjunction conditions, and again reveals conjunction decoding in the absence of any potential attentional confounds.

Separate Voxels Code for Motion and Color

Linear support vector machines assign a weight that indicates its importance in the given discrimination task to each voxel. We were interested to learn to what extent voxels with high weights in SVMs trained on color discrimination overlapped with those obtaining high weights in motion discrimination and how these weights related to those in conjunction-trained SVMs. A map of V1 illustrating the spatial distribution of voxels coding well for color, motion, and conjunctions in subject EF can be found in Figure S6. We quantified this overlap by calculating the angles between pairs of weight maps obtained from the three classifier types, for each ROI for each subject (see [Experimental Procedures](#)). We found that the weight maps between SVMs trained for color and motion differed by larger angles than expected on the basis of bootstrap resampling, indicating a negative relationship between color and motion information carried among voxels: in each ROI, the better a voxel coded for color, the worse it coded for motion (in radians: mean: 0.19 ± 0.02 SD, max = 0.22, min = 0.16, $n = 24$, $p < 0.05$ in every ROI). We also found either no or negative relationships between weight maps of conjunction-trained SVMs with those trained for color or for motion. Specifically, V1 (mean: 0.06 ± 0.01 SD, $n = 4$, $p < 0.05$) and V3a (mean: 0.04 ± 0.02 SD, $n = 4$, $p < 0.05$) showed significantly negative relationships between conjunction- and color-trained weight maps; V5/MT⁺ (mean: 0.13 ± 0.04 SD, $n = 4$, $p < 0.05$) showed a significantly negative relationship between conjunction- and motion-trained weight maps. The results therefore indicate that even though most areas contain information about color, motion, and their conjunctions, this information is segregated at the scale of voxels, with distinct sets of voxels carrying information for color, for motion, and for their conjunctions. The results are also in line with the above generalization results that showed that the decoding of one feature was not influenced by the other.

Discussion

We used linear multivariate-classifier analyses to investigate whether information about color, motion, and the conjunction of these two features could be decoded from patterns of BOLD responses recorded in functionally defined visual cortical areas of the human brain. As early as the primary visual cortex,

patterns of BOLD signals distributed across voxels were shown to contain reliable information not only about motion direction and color hue but also about the specific conjunction of these features. Whereas some areas showed better performance as well as biases for decoding one feature over the other (e.g., V5/MT⁺ for motion; V4 for color), information about both features and their conjunction was present in nearly every visual cortical region. Analysis of the patterns of activity learned by the multivariate classifiers suggests that largely distinct sets of voxels contain information about color and motion and their conjunction, indicating a considerable degree of segregation within each region. Furthermore, the conjunction of color and motion appears to be coded explicitly at every stage of cortical visual processing.

Motion Information Can Be Decoded throughout the Visual Cortex

Single-cell electrophysiology has shown that whereas the primary regions involved in motion processing involve V1, V2, and V5/MT⁺, directionally selective neurons can also be found in almost every visual area in the primate cortex [30]. The presence of directional-motion information across the human visual cortex has accordingly been confirmed in recent fMRI experiments using multivariate analyses [26]. The results of the current study support these findings. The relatively weak decoding performance obtained from V5/MT⁺ (even when voxel numbers are equated; [Figures S2 and S3](#)), an area with possibly the highest density of directionally selective units, is counterintuitive. It is, however, precisely this dense arrangement of directionally selective units that may account for such a finding. Importantly, decoding performance does not rely solely on the presence of feature information but also on the spatial layout of this information in the brain (see [31] for an illustration).

Color Information Can Be Decoded throughout Visual Cortex except V5/MT⁺

We found that BOLD signals in the human visual cortex contain information allowing for correct predictions of the color present in a stimulus; this is, to our knowledge, the first report of this kind. Prior studies reported decoding of colored oriented gratings [29] or of color percepts during rivalry [32], but in the former decoding relied equally on the grating orientation, and in the latter neither color-channel load nor luminances were precisely matched, perhaps explaining differences in the results. In our data, hue could be decoded from activity patterns within V1, V2, V3, V3A/B, and V4. This finding is unlikely to reflect luminance artifacts, given that colors were set to isoluminance in every subject and luminance was pseudorandomly offset by $\pm 10\%$ (see [Experimental Procedures](#)). Additionally, signals to the L-M (“red-green”) and S (blue-yellow) channels were matched in both colors such that the sum of red and green (or blue and yellow) was perceived as achromatic gray. In contrast to all other visual regions tested, the fMRI signals we measured within motion area V5/MT⁺ did not allow for significant color discrimination. Because classification performance is dependent on spatial layout as well as feature selectivity, interpretations of a null result must be met with caution and cannot prove a lack of information [31]. Nonetheless, in the case of V5⁺/MT⁺ the result of a lack of color coding is consistent with primate physiology [9]. In contrast, when a classifier was forced to either rely on motion or on color information, V4 was the only region to present a bias toward color decoding as opposed to motion decoding. This is consistent with

evidence from electrophysiology, imaging, and patient studies that have demonstrated a central role for the V4 complex in color processing and perception [1, 5, 7, 11]. In accord with this, decoding success for color was highest in V4 when the number of voxels was equated across ROIs. This finding is compatible with a high degree of hue information in V4, and the result may have been strengthened by a modular functional organization within V4, similar to the blobs or stripes in V1 or V2, such that distinct sites may be particularly specialized for color processing and distinct hues may be spatially organized to form cortical color maps with a spacing that favors voxel-wise biases [33].

Conjunctions, as Opposed to Joint Selectivity, Can Be Decoded throughout the Visual Cortex

We found that information related to a specific pairing of color and motion direction could be decoded from BOLD responses in V1, V2, V3, V3A/B, V4, and V5/MT⁺. This constitutes the first direct functional evidence of conjunction coding of color and motion in the human visual cortex. Because each double-conjunction stimulus contained the same basic color and motion information, a classifier could not rely on independent color and motion “feature maps” to distinguish the two double-conjunction stimuli. In such a case of joint selectivity, both stimuli would evoke the same activation. The only factor distinguishing our two double-conjunction stimuli was the unique pairing of features. Note that the same applies for potential nonlinearities arising through neurovascular coupling: any nonlinear mixing of color-specific with motion-specific responses could not account for responses specific to one of the double-conjunction stimuli, given that each double-conjunction stimulus would elicit identical neural responses unless there was a specific neural detector for feature conjunctions. Therefore, our data show evidence for specific feature-conjunction information throughout the visual cortex. Because the voxels most informative about color-motion conjunctions were largely distinct from those informative about color or motion alone, it is possible that separate functional units code for motion, color, and conjunctions of these features. The high degree of invariance in motion classification across color and vice versa also supports such an interpretation.

It is important to note that conjunction coding has not been directly demonstrated by the many studies simply showing joint selectivity for various visual features, such as color and orientation, color and depth, or color and motion, etc. [8, 29]. These studies have not disentangled whether cells (or voxels) respond exclusively to specific conjunctions of two features, have supralinear responses, or simply respond when either feature is present. Until now, conjunction coding of motion and color has primarily been suggested on the basis of psychophysical studies that have examined motion aftereffects [19, 20, 34]. Although the neural loci of the mechanisms mediating the color-contingent MAE are not certain, the effect shows little or no interocular transfer [20, 35] and can be elicited by adaptation to a locally paired opposite-motion dot display [19], suggesting that it may reflect changes in V1. The apparent processing asynchrony between color and motion evident in the color-contingent MAE [14] would be consistent with a role for feedback to V1 in generating the aftereffect. Specifically, if extrastriate areas use feedback to V1 to “gate” their inputs, then the timing of this gating might be important for adaptation of “double-duty” units in V1 that are selective both for the color and the direction of motion of the stimulus [24]. Although there

has been physiological evidence to suggest that at least some joint selectivity for motion and color is present in V1 [8], our present results provide direct evidence for early conjunction coding in the human brain. Indeed, our findings suggest that engagement of neural mechanisms at the level of V1 may be critical in solving what has been taken as a prime example of the binding problem [15, 18, 23, 36]. However, these results must be met with caution because the spatial and temporal resolutions of fMRI limit the interpretation of the precise neural machinery underlying such pattern activation [31, 37]. Cellular physiology may deduce whether conjunction information is explicitly represented in cells or columns of cells and whether it is generated within V1 or is a result of feedback from higher levels [23, 38].

Conclusions

Our fMRI study using multivariate-pattern analysis has provided new evidence about the functional architecture of the processing of color and motion in the human visual cortex. We confirmed the presence of directionally specific motion information across the visual cortex and also showed for the first time evidence of spatially structured hue representations in several visual regions. The cortical representation of a given feature, e.g., color, exhibited a high degree of invariance to changes in the corresponding secondary feature it was attached to, e.g., motion. This and our analysis of the weight maps within each visual area extend the prior evidence for a segregated processing of color and motion. Notably, however, we found evidence for the explicit coding of conjunctions of motion and color—two features that are the most common exemplars of segregated processing in the visual brain. Information on color-motion conjunctions was found throughout the visual system, including the primary visual cortex. Evidence of conjunction coding across the visual cortex has implications for theories concerning the binding problem and, to our knowledge, has not been obtained before in human or in monkey.

Experimental Procedures

Stimuli and Subjects

Five subjects (one author) participated in this study. All had normal, or corrected-to-normal, visual acuity and color vision. Each subject was familiarized with the task during one preliminary psychophysics session outside of the scanner.

Conjunction and Double-Conjunction Stimuli

In total we presented six stimulus conditions in a blocked design, all consisting of colored dots rotating around the center of the screen (see below). Four stimulus conditions resulted from the four possible combinations of pairing one of two colors with one of two motion directions: red clockwise (RCW), red counterclockwise (RCCW), green clockwise (GCW), and green counterclockwise (GCCW). Alone, each of these stimuli could be considered a conjunction stimulus, but they could be distinguished by an observer (or classifier) on the basis of independent recognition of color and motion. We therefore included two additional double-conjunction conditions. Each comprised two color-motion pairings: either RCW + GCCW or RCCW + GCW. These stimuli could only be distinguished by true conjunction coding, given that each contained both colors and both motion directions, differing only by their conjunctions (see Figure 1). These stimuli were perceived as transparent surfaces, but the individual dots were randomized in their depth ordering during occlusions to avoid any apparent depth ordering of the whole surfaces.

Blocks of each condition lasted 12 s and were presented five times throughout one run of the experiment in pseudorandom permutations of the six conditions. Each block type was preceded equally often by the others. There were a total of eight runs, each lasting 6 min.

In order to ensure that subjects were attending to the stimuli, we incorporated changes in luminance or dot speed. Each block contained two changes at random times, each lasting 500 ms. When subjects perceived a change, they indicated by pressing one of two buttons whether the change was in luminance or speed. For each stimulus condition, luminance decrements (halving) occurred equally as often as speed increments (doubling). Subjects fixated on a central fixation cross while carrying out this task.

Basic Stimulus Parameters

Each stimulus consisted of 1000 colored dots (mean luminance: 10 cd/m²) on a black background. The dots were randomly positioned within an annulus extending from 1.0 to 7.5 degrees of eccentricity and rotated around the center at a rate of 0.167 rotations per second. Dot size (range: 0.13 to 0.26 deg) scaled with eccentricity with an exponent of 1.3 to account for the magnification factor of receptive-field size in the visual cortex. Dot density followed the corresponding inverse scaling.

Colors were set to isoluminance with the minimum flicker technique inside the scanner for each subject [39]. Hues were adjusted such that they combined to gray, thus ensuring matched saturation and color-channel load. In addition to this, we presented each color at one of two luminance settings (9 or 11 cd/m²) in each block, with equal frequency across runs. Blocks of the four single-conjunction stimuli would be randomly either “high” or “low” in luminance, and double-conjunction blocks could be low-low, low-high, high-low, or high-high for the two colors. This minimized the possibility of luminance-based color classification. Dots within a block also had random luminance values within a range of $\pm 10\%$ of the mean luminance of the color in that block. All stimuli were presented with Cogent2000 version 1.27 (J. Romaya, Wellcome Dept. of Imaging Neuroscience, London; http://www.vislab.ucl.ac.uk/cogent_graphics.php) running under Matlab 2006b (Mathworks). They were presented at a resolution of 1280 \times 1024 pixels and at a screen refresh rate of 75 Hz from a viewing distance of 82 cm. Stimuli were projected onto a transparent screen at the end of the scanner bore and viewed through a tilted mirror fixed to the head coil.

Localizer Stimuli

For the purpose of phase-encoded retinotopic mapping, according to standard procedures we presented three rotating wedge runs and two expanding ring runs lasting 6 min each to each subject [6]. In addition, we localized regions generally responsive to our stimuli using a 6 min run of alternating blank fixation and double-conjunction blocks of 12 s each.

fMRI

Functional images were acquired in a 3 Tesla Siemens (Erlangen, Germany) TIM scanner with a gradient echo planar imaging (EPI) sequence and a 12-channel phased-array head coil. We collected 27 slices positioned over the visual cortex, using an interleaved sequence with the following parameters: repetition time (TR) 2.34 s; echo time (TE) 39 ms, 96 \times 96 matrix; and voxel size 2 \times 2 \times 2 mm. A high-resolution (1 mm isotropic) T1-weighted 3D-MDEFT image was acquired as an anatomical reference.

Preprocessing, Area Delineation and Voxel Selection

We carried out the minimum of preprocessing using BrainVoyagerQX. Data was coregistered in raw AC-PC space and not transformed to any standard coordinate system. We corrected for head motion and made a mean intensity adjustment (global scaling); no smoothing was applied.

For the delineation of retinotopically mapped visual areas with the phase-encoding method, cortical inflations of each subject were reconstructed from a high-resolution T1-weighted image. Gray and white matter was segmented and the cortex was reconstructed with BrainVoyagerQX (<http://www.brainvoyager.com/>). Border delineation with polar retinotopic mapping was achieved with established methods [6]. A linear correlation between wedge position and neural activity was performed for each voxel, and borders of areas V1–V4 and V3A/B were identified on the basis of field sign alternations. Delineation on the cortical representation was carried out manually. Areas V1–V3 were defined on the basis of their clear meridian borders. For areas V3A/B, V4, and V5/MT⁺, we used the following operational definitions: V3A/B was defined as the retinotopically organized region dorsal to V3d representing both lower and upper visual field—we did not attempt to separate V3A from V3B. V4 was operationally defined as voxels anterior to V3v in the fusiform gyrus representing both upper and lower visual fields. Its somewhat diffuse anterior boundaries may include regions anterior to it such as V4_v/VO-1 [5, 6]. V5/MT⁺ was evident in each subject as a separate cluster of voxels responsive to the (moving) stimulus localizer, centered at the ascending limb of the inferior temporal sulcus (ALITS) [40].

Multivariate analyses were carried out separately for each subject for each of the cortical areas defined above, with voxels that reached a liberal threshold of at least $p < 0.01$ (uncorrected) in the localizer scan. This ensured all voxels responsive to the size and location of our stimuli were considered. In addition, analyses matched for voxel number across ROIs were conducted, by selecting either the best N or a random N voxels from each ROI (N = 233). A size of 233 voxels was selected on the basis of the smallest ROI (V5/MT⁺) found across all five subjects.

For each ROI, the data from the main experiment were analyzed with a GLM containing a separate regressor for every stimulus block. Each regressor was a 12 s boxcar convolved with the canonical hemodynamic response function (hrf). For every voxel, this resulted in 40 beta-estimates for each of the six conditions.

Multivariate Analysis

The beta-estimates were used to train linear SVM algorithms to distinguish responses to motion directions, color hues, and double conjunctions [41]. We used the matlab implementation of SVMs provided by the Spider toolbox to achieve this (<http://www.kyb.tuebingen.mpg.de/bs/people/spider/>).

Prior to training SVMs, data from each voxel were normalized to a mean of 0 and a standard deviation of 1. Outliers were removed from the data by setting all values that were beyond 2 standard deviations from the mean to a fixed value of ± 2 SD. A leave-one-out approach was employed, in which a classifier was trained on voxel responses from N-1 blocks and tested on the remaining block, cycling through all blocks. During each training iteration, a 10-fold cross validation was run on the training set to select for the C parameter that optimized learning generalization (for C, a range of 2 to the power of -7 to 1 and infinity were used). For each ROI of each subject, the mean classification accuracy of the test blocks was determined (chance performance was always 50%). We determined significance for each ROI using a one-sample t test performed across the classification accuracies of the five subjects.

Examining Color and Motion Selectivity

Our four single-conjunction conditions allowed us to pool sets of blocks to train SVMs for motion-direction or color-hue discrimination, respectively. For discrimination of color we pooled blocks of RCW with RCCW and blocks of GCW with GCCW; for discrimination of motion we pooled RCW with GCW and RCCW with GCCW. Thus, when training a SVM for one feature (e.g., color discrimination), the SVM had to generalize over the other (e.g., motion direction).

We also examined the extent to which each visual area would bias classification to rely more on motion or color information. To achieve this, we first trained an SVM to discriminate RCW versus GCCW and then tested on an entirely new block type, namely GCW versus RCCW (and vice versa). The classifier thus had a dilemma: it could either achieve correct performance on color hue, thereby getting motion direction wrong, or vice versa. The result would reveal which of the two features was more “salient” to the classifier, for each region.

Finally, we examined whether motion direction could influence color-hue classification performance and, conversely, whether color had an influence on direction-classification performance. For the former, we trained a classifier to discriminate RCW versus GCW (39 blocks) and tested on the left-out block. We then also tested iteratively on each block from a data set consisting of RCCW versus GCCW. Equal performance in both cases would imply that color classification was unaffected by the motion direction. The same procedure was used to examine motion classification and its dependence on color hue.

Testing for Conjunction Selectivity

The primary aim of this study was to recover evidence for the presence of information coding feature conjunctions across visual regions. We examined this in two ways.

First, we used the two transparent double-conjunction conditions (RCW/GCCW versus RCCW/GCW) to train and test an SVM. Blocks of each condition contained both color hues and both motion directions. Successful classification had thus to rely on distinct voxel-pattern responses evoked by the two distinct conjunctions of color and motion in the stimuli.

Second, instead of using voxel responses evoked by the double-conjunction stimuli, we created two sets of “artificial” voxel responses, each containing information about both color hues and both motion directions, by averaging beta estimates from pairs of the first four conditions (i.e., RCW + GCCW versus RCCW + GCW). If a classifier could discriminate between these two data sets then, under the same rationale as above,

information about the conjunction of features present in the individual conditions had to be available to the learning algorithm.

Voxel-Wise Segregation of Color, Motion and Conjunction Information

In order to test whether information about color, motion, and conjunctions was conveyed by the same or different sets of voxels, we compared the absolute magnitudes of the classifier weights that had been assigned to each voxel during training on the distinct stimulus categories. Thus, a given weight map of n voxels can be considered a vector in n -dimensional space. The angle between two such weight vectors indicates their similarity. For each pair of weight maps, we calculated their angle. In addition, we also calculated the mean angle resulting from 1000 random permutations of their values and thus yielded the expected angle given the assumption of no relation between the two weights vectors. For each ROI, we then tested (t test) whether across the five subjects the actual angles between two weight maps differed significantly from the expected angles. We carried out the analysis by using only voxels whose weights exceeded ± 2 SD for either of the features, to test whether the highly informative voxels of two differently trained SVMs were more (or less) related than expected by chance. Significantly smaller angular differences than expected by chance would indicate a positive relation, in that voxels informing on, e.g., motion direction, were also informative about color, etc. Significantly larger angular differences than chance would indicate a negative relation, in that the better a given voxel coded for, e.g., motion, the worse it coded, e.g., for color. A lack of significant angular difference would imply that weights, e.g., for color were not related to those for, e.g., motion.

Supplemental Data

Supplemental Data include six figures and can be found with this article online at [http://www.current-biology.com/supplemental/S0960-9822\(09\)00544-2](http://www.current-biology.com/supplemental/S0960-9822(09)00544-2).

Acknowledgments

We thank Arthur Gretton for insightful discussions on and help with support vector machines. This work was supported by the Max Planck Society, Germany, an Australian Research Council Discovery Project, an Australian Research Fellowship to C.C., and a Campbell-Perry award to K.S.

Received: September 29, 2008

Revised: November 18, 2008

Accepted: December 10, 2008

Published online: January 29, 2009

References

- Zeki, S.M. (1978). Functional specialization in the visual cortex of the monkey. *Nature* 274, 423–428.
- Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47.
- Shipp, S., and Zeki, S. (1995). Segregation and convergence of specialized pathways in macaque monkey visual cortex. *J. Anat.* 187, 547–562.
- Sincich, L.C., and Horton, J.C. (2005). Input to V2 thin stripes arises from V1 cytochrome oxidase patches. *J. Neurosci.* 25, 10087–10093.
- Bartels, A., and Zeki, S. (2000). The architecture of the colour centre in the human visual brain: New results and a review. *Eur. J. Neurosci.* 12, 172–193.
- Brewer, A.A., Liu, J., Wade, A.R., and Wandell, B.A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat. Neurosci.* 8, 1102–1109.
- Desimone, R., and Schein, S.J. (1987). Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *J. Neurophysiol.* 57, 835–868.
- Horowitz, G.D., and Albright, T.D. (2005). Paucity of chromatic linear motion detectors in macaque V1. *J. Vis.* 5, 525–533.
- Seidemann, E., Poirson, A.B., Wandell, B.A., and Newsome, W.T. (1999). Color signals in area MT of the macaque monkey. *Neuron* 24, 911–917.
- Damasio, A.R. (1985). Disorders of complex visual processing: Agnosias, achromatopsia, Balint's syndrome, and related difficulties of orientation and construction. In *Principles of Behavioural Neurology, Volume 1*, M.M. Mesulam, ed. (Philadelphia: Davis), pp. 259–288.
- Vaina, L.M. (1994). Functional segregation of color and motion processing in the human visual cortex: Clinical evidence. *Cereb. Cortex* 4, 555–572.
- Schmolesky, M.T., Wang, Y., Hanes, D.P., Thompson, K.G., Leutgeb, S., Schall, J.D., and Leventhal, A.G. (1998). Signal timing across the macaque visual system. *J. Neurophysiol.* 79, 3272–3278.
- Moutoussis, K., and Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. *Proc. R. Soc. Lond. B. Biol. Sci.* 264, 393–399.
- Arnold, D.H., Clifford, C.W.G., and Wenderoth, P. (2001). Asynchronous processing in vision: Color leads motion. *Curr. Biol.* 11, 596–600.
- Bartels, A., and Zeki, S. (2006). The temporal order of binding visual attributes. *Vision Res.* 46, 2280–2286.
- Treisman, A. (1996). The binding problem. *Curr. Opin. Neurobiol.* 6, 171–178.
- Engel, A.K., Fries, P., Konig, P., Brecht, M., and Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Conscious. Cogn.* 8, 128–151.
- Holcombe, A.O., and Cavanagh, P. (2001). Early binding of feature pairs for visual perception. *Nat. Neurosci.* 4, 127–128.
- Blaser, E., Papathomas, T., and Vidnyanszky, Z. (2005). Binding of motion and colour is early and automatic. *Eur. J. Neurosci.* 21, 2040–2044.
- Favreau, O.E., Emerson, V.F., and Corballis, M.C. (1972). Motion perception: A color-contingent aftereffect. *Science* 176, 78–79.
- Bartels, A., and Zeki, S. (1998). The theory of multi-stage integration in the visual brain. *Proc. R. Soc. Lond. B. Biol. Sci.* 265, 2327–2332.
- Hochstein, S., and Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804.
- Clifford, C.W., Spehar, B., and Pearson, J. (2004). Motion transparency promotes synchronous perceptual binding. *Vision Res.* 44, 3073–3080.
- Clifford, C.W. (in press). Dynamics of visual feature binding. In *Space & Time in Perception & Action*, R. Nijhawan and B. Khurana, eds. (Cambridge: Cambridge University Press).
- Haynes, J.D., and Rees, G. (2006). Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534.
- Kamitani, Y., and Tong, F. (2006). Decoding seen and attended motion directions from activity in the human visual cortex. *Curr. Biol.* 16, 1096–1102.
- Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–685.
- Haynes, J.D., and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nat. Neurosci.* 8, 686–691.
- Sumner, P., Anderson, E.J., Sylvester, R., Haynes, J.D., and Rees, G. (2008). Combined orientation and colour information in human V1 for both L-M and S-cone chromatic axes. *Neuroimage* 39, 814–824.
- Andersen, R.A. (1997). Neural mechanisms of visual motion perception in primates. *Neuron* 18, 865–872.
- Bartels, A., Logothetis, N.K., and Moutoussis, K. (2008). fMRI and its interpretations: An illustration on directional selectivity in area V5/MT. *Trends Neurosci.* 31, 444–453.
- Haynes, J.D., and Rees, G. (2005). Predicting the stream of consciousness from activity in human visual cortex. *Curr. Biol.* 15, 1301–1307.
- Xiao, Y., Wang, Y., and Felleman, D.J. (2003). A spatially organized representation of colour in macaque cortical area V2. *Nature* 421, 535–539.
- McKeefry, D.J., Laviers, E.G., and McGraw, P.V. (2006). The segregation and integration of colour in motion processing revealed by motion after-effects. *Proc. Biol. Sci.* 273, 91–99.
- Mayhew, J.E.W., and Anstis, S.M. (1972). Movement aftereffects contingent on colour intensity and pattern. *Percept. Psychophys.* 12, 77–85.
- Engel, A.K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* 5, 16–25.
- Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature* 453, 869–878.
- Lamme, V.A., and Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579.
- Kaiser, P.K. (1991). Flicker as a function of wavelength and heterochromatic flicker photometry. In *Limits of Vision*, J.J. Kulikowski, V. Walsh, and I.J. Murray, eds. (Basingstoke: MacMillan), pp. 171–190.
- Dumoulin, S.O., Bittar, R.G., Kabani, N.J., Baker, C.L.J., Le Goualher, G., Bruce Pike, G., and Evans, A.C. (2000). A new anatomical landmark for reliable identification of human area V5/MT: A quantitative analysis of sulcal patterning. *Cereb. Cortex* 10, 454–463.
- Vapnik, V.N. (1998). *Statistical Learning Theory* (New York: Wiley).