Title
Color Influences Fast Scene Categorization

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 18(0)

Authors
Oliva, Aude
Schyns, Phillipe G.

Publication Date
1996

Peer reviewed

Color Influences Fast Scene Categorization

Aude Oliva and **Philippe G. Schyns**
Department of Psychology
University of Glasgow, Glasgow
Glasgow, G12 8QQ, UK
{aude,philippe}@psy.gla.ac.uk

Abstract

A critical aspect of early visual processes is to extract shape data for matching against memory representations for recognition. Many theories of recognition assume that this is being done on luminance information. However, studies in psychophysics have revealed that color is being used by many low-level visual modules such as motion, stereopsis, texture, and 2D shapes. Should color really be discarded from theories of recognition? In this paper, we present two studies which seek to understand the role of chromatic information for the recognition of real scene pictures. We used three versions of scene pictures (gray-levels, normally colored and abnormally colored) coming from two broad classes of categories. In the first category, color was diagnostic of the category (e.g., *beach, forest* and *valley*). In the second category color was not diagnostic (e.g., *city, road* and *room*). Results revealed that chromatic information is being registered and facilitates recognition even after a 30 ms exposure to the scene stimuli. Similar results were recorded with exposures of 120 ms. However, influences of color on speeded categorizations were only observed with the color-diagnostic categories. No influence of color was observed with the other categories.

Introduction

There is little doubts that shape is a critical aspect of recognition. In fact, most recognition theories assume that shape serves as the basis of object and scene representations (e.g., Biederman, 1987; Marr & Nishihara, 1978; Tarr & Pinker, 1989). Consequently, a critical aspect of early visual processes is to extract shape data for matching against memory representations. This is accomplished by early visual modules such as edge detection, depth perception, motion and stereo which are often assumed to operate on achromatic information--the luminance, or gray-levels of the image. However, there is a wealth of psychophysical evidence suggesting that early vision operates simultaneously in luminance and chromatic pathways for motion (e.g., Cavanagh & Ramachandran, 1988), simple shapes (Damasio, Yamada, Damasio, Corbett, McKee, 1980), texture (McIlhaga, Hine, Cole & Schneider, 1990) and stereo (Logothetis, Schiller, Charles & Hurlbert, 1990). In light of these evidence, should chromatic information really be discarded from recognition theories?

The influence of chromatic information on object and scene recognition has been little studied, and results are often contradictory. For example, Biederman and Ju (1988) have argued that color had no impact on the identification of an object. Their experiments compared the naming of colored object photographs and line drawings of the same objects. No naming advantage was reported for colored objects (at presentations times of 50 ms and 100 ms), even when the color was characteristic (diagnostic) of the objects. This contrasts with the earlier work of Ostergaard and Davidoff (1985) who reported that objects like fruits and vegetables were named faster when the stimuli were colored slides then when they were gray-level slides (see also Wurm, Legge, Isenberg, & Luebker, 1993). Recently, Tanaka & Bunosky (1996) have shown that color influences objects that have a characteristic color.

Two factors could account for these conflicting data. This first factor is the diagnosticity of chromatic information for a categorization (see Tanaka & Bunoski, 1996). For example, in the *lemon* category, the distribution of color is clustered around only two modes (yellow and green). In contrast, the distribution of color is multi-modal in the *car* category. Thus, an effect of chromatic information on recognition might only be observed when color predicts the categorization, insisting for the requirement to control for color diagnosticity. The second factor is the control of luminance across conditions of stimulation. Too often (e.g., Biederman & Ju, 1988; Ostergaard & Davidoff, 1985), comparisons are made between stimuli whose luminance differs across conditions, making it difficult to know whether the absence of effect of chromaticity in one condition results from an enhanced luminance in the other condition.

While it is probably true that color is not diagnostic of many real world object categories (with the exception of fruits and vegetables), little is known about the possible effects of color on the recognition of real world scenes. Real-world scenes are arguably more diagnosed by their colors than are objects (think for example of beaches, seas, mountains, fields, and so forth. Unfortunately the little research existing on the categorization of real world scene has mostly focused on luminance information (e.g., Oliva & Schyns, 1995; Schyns & Oliva, 1994). However, a recent work of Gegenfurtner, Sharpe, & Wichmann (1995) showed that color entered the way scenes were represented in memory. Their study asked subjects to learn colored and black and white scene pictures to be later tested on their familiarity with these scenes. It was found that subjects performed significantly better for memorized colored images, suggesting that their memory of scenes contained chromatic information.

This paper presents two studies which seek to understand the respective role of luminance and chromatic information in speeded scene categorizations. One could claim that the addition of color has a relatively low-level contribution which facilitates segmentation processes. In this case a beach scene whose sky was painted red, the beach blue, and the sea pink should be categorized faster than its gray-level counterpart. An alternative claim is that color exerts a high-level influence on processing. For example, recognition could operate independently in luminance and color pathways, both of which would provide sources of cues for recognition. In this case, the beach would be recognized faster when the luminance and chromatic cues both represent a beach. However, this should only hold when chromatic information is diagnostic of a category. That is, it should hold only for categories with few dominant colors (e.g., *beach* or *forest*), but not for categories with many variable colors (e.g., *room* or *city*).

Our experiments were designed to get a better understanding of these issues. In a categorization task, all subjects were exposed to three versions of 120 different scene stimuli: normally colored, abnormally colored and gray levels. Precautions were taken to insure that the three versions of a scene had an identical luminance. In Experiment 1, the stimuli were presented for 120 ms and we measured subjects' reaction times of denomination. In Experiment 2, scenes were presented for only 30 ms. Both experiments used the same scene categories. To control for diagnosticity, two classes of categories were used: natural (*beach, forest* and *valley*) and artifact (*room, road* and *city*). Natural categories have fewer color modes than artifact categories, and color is therefore more diagnostic of the categories. We expected chromatic information to facilitate recognition only when it was diagnostic of a category. Conversely, we expected chromatic information to impair recognition if the colors were abnormal of the category. The difference in presentation times between the two experiments was there to inform about the time course of luminance and chromatic processing.

Experiment 1

Method

Subjects. Twenty Glasgow University students with normal or corrected vision were paid to participate to the experiment.

Stimuli. We chose 120 high quality color scene pictures from the Corel CD Photo Library. These images formed two categories: natural and artifact scenes. Natural and artifact scenes were themselves subdivided into 3 subcategories (*beach, forest, valley* and *city, road, room*) with an equal number of exemplars per category. Our experiments involve transformations of chromatic information. A colored images has two main components: luminance (the gray levels of the image) and chromaticity (the color itself). A standard 3 dimensional encoding of color in Red Green Blue (RGB) space does not separate luminance from chromaticity. Thus, changes of color affect luminance. To control this potential confound we transformed RGB data in a space (L*a*b*) which represents luminance (L*) and color (a*b*) independently. We then obtained three different versions of the same scene. The first version, the Normal Color (NC) scene simply corresponded to the original image. The Incongruent Color (IC) was computed by changing the a* and b* axis in L*a*b*. There are three ways to change color, because the a* and b* axis can be swapped and inverted. To illustrate, if a* represents the red to green spectrum and b* represents the yellow to blue spectrum, a swap of a* and b* (L*b*a*) changes the color of a banana from yellow to red. A swap of the axis plus an inversion of their values would create a blue banana. For each of the 120 scenes, 3 IC images were computed as just explained. A panel of three independent judges was asked to rate how atypical these images appeared of the category. We kept the most atypical color transformation for the IC condition. The final version of a scene (LU) simply discarded all color information, only keeping the L* component of L*a*b* (the LUminance, or gray levels). Stimuli were presented on a computer monitor; images subtended 6.4 x 4.4 degrees of visual angle. As there were 120 original pictures, together, NC, IC and LU composed a set of 360 experimental stimuli.

Procedure. The experiment was composed of 360 trials: 120 NC, 120 IC and 120 LU. Each trial used a different scene picture. A trial consisted of the presentation of a fixation square, followed by a mid-gray image presented for 500 ms, immediately followed by a target image (NC, IC or LU) presented for 120 ms. Subjects were instructed to name the scene as quickly and as accurately as they possibly could. Reaction Times (RTs) were measured with a vocal key. Prior to the experiment, subjects were told that all scene stimuli belonged to one of six possible categories (*beach, city, forest, room, road* and *valley*) and were asked to categorize the pictures with these names. Six practice trials were given to familiarize subjects with the experimental apparatus, and to calibrate the sensitivity of the vocal key. Trials were randomly assigned into 10 blocks of 36 trials. Subjects were allowed a one minute pause between each block.

Results and Discussion

The analysis was only performed on correct categorizations when naming latencies were within 2.5 standard deviations of the means. Figure 1 summarizes the results. A two-way, within-subjects ANOVA with type of scenes (natural vs. artifact) and stimulus conditions (NC, IC and LU) revealed a main effect of scene type, $F(1,19) = 131.71$, $p <$

.0001, a main effect of stimulus condition, $F(2,38) = 24.62$, $p < .0001$, and a significant interaction $F(2,38) = 24.76$, $p < .0001$. Our results demonstrate that chromatic manipulations affect recognition performance. However, the interaction reveals that this influence was different for different categories. We hypothesized that chromatic information would only be useful for recognition when it was diagnostic of a category. That is, categorization times should be significantly faster for a chromatic NC scene than for its gray-level LU equivalent. Indeed, mean RTs were significantly higher when chromatic information was removed from a natural scene, $F(1,38) = 54.755$, $p < .001$, (respectively, LU = 816 ms and NC = 786 ms) but the LU and NC versions of artifact scenes elicited almost identical RTs (742 ms and 738 ms, respectively). Another consequence of processing chromatic information for recognition is that a change of color could in principle impair recognition. However, this should only apply when color is diagnostic of a categorization. To test this hypothesis, we compared categorization latencies of IC and RC stimuli. We found that IC scenes took 46 ms longer to be categorized than NC scenes, $F(1,38) = 124.56$, $p < .0001$, and also that IC were slower to name that LU, $F(1,38) = 14.14$, $p < .001$. No such difference was observed for artifact scenes (see Figure 1).

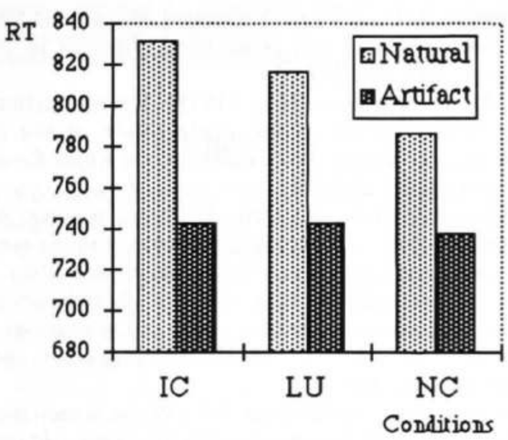

Figure 1: Mean reaction times as a function of type of category (Natural vs. Artefact) and color condition (IC = Incongruent Color, LU = Luminance, NC = Normal Color) for a stimulus presentation of 120 ms.

In sum, in opposition to the general conclusions of Biederman and Ju (1988), color is being used for recognition--at least for the fast identification of scenes. However, in agreement with Ostergaard and Davidoff (1985) this influence is limited to categories which are distinguishable on the basis of their color. Our data suggest that color and luminance are two independent cues available for fast scene recognition. However, we also found that color has no effect when it is not diagnostic of a category.

Experiment 2

The results of Experiment 1 were gathered in conditions of 120 ms presentations of scenes. Other studies have shown that complex scenes may be recognized with very fast accuracy even after shorter presentation times (Schyns & Oliva, 1994). However, these studies did not distinguish between the relative contribution of luminance and chromatic cues for recognition. An argument could be made that the time course of luminance and chromatic pathways are different (Billock, 1995) and that one (e.g., chromaticity) is systematically dependent on processing being initialized in the other (e.g., luminance). Thus, a 120 ms presentation of the stimuli could in fact mask the fact that one source of cues is being registered much later that the other cues. Experiment 2 replicated Experiment 1, using identical stimuli, but we now presented them for 30 ms. Although changes of presentation times do not test time course of processing per se, this experiment could reveal whether a very short presentation time was sufficient to register chromatic information sufficiently well to later speed up recognition.

Method

Subjects. Twenty Glasgow University students with normal or corrected vision were paid to participate to the experiment.

Stimuli. The same 120 color scenes as in Experiment 1, subdivided into natural and artifact scenes, were used in Experiment 2. The experiment was composed of the same 360 trials (120 NC, 120 IC and 120 LU).

Procedure. The procedure was identical to Experiment 1, except that the presentation time of the scenes was now 30 ms. Subject's task was to name the input as fast as they possibly could, using one of the possible category names (*beach, city, forest, room, road and valley*).

Results and Discussion

Subjects' mean RTs were collected for correct categorizations. A two-way within-subjects ANOVA with type of scene (natural vs. artifact) and stimulus condition (NC, IC and LU) revealed a significant of the main effect of scene type, $F(1, 19) = 24.10$, $p < .0001$, stimulus condition $F(2, 38) = 19.30$, $p < .0001$, and a significant interaction, $F(2, 38) = 4.87$, $p < .02$. Figure 2 summarizes the results. Globally, the patterns of results after a 30 ms presentation are similar to those reported for Experiment 1 with a 120 ms presentation: When it is diagnostic, chromatic information exerts a strong influence on scene recognition. This is observed for natural, NC categories which were recognized significantly faster then their LU versions, $F(1,38) = 21.96$, $p < .0001$, but no such effect was observed for artifact scenes (LU = 772 ms and NC = 767 ms). Locally, the categorization times of natural scenes revealed an interesting difference with the results observed in Experiment 1: IC scenes were categorized 36 ms slower than NC scenes, $F(1,38) = 22.21$, $p < .0001$, but not slower than their gray-level counterparts. No such difference existed for artifact scenes.

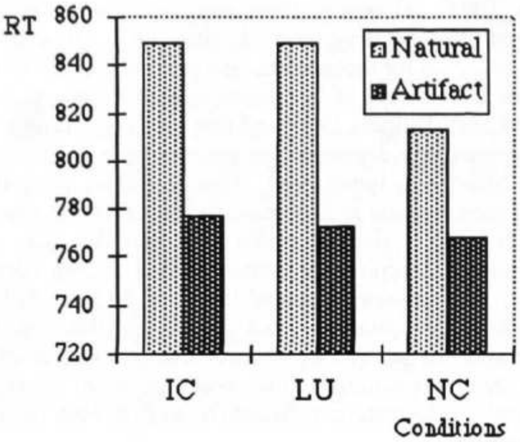

Figure 2: Mean RT as a function of type of category and color condition for a stimulus presentation of 30 ms.

These results confirm that chromatic information is used for speeded scene recognition. Why we did not obtain an interference of incongruent colors when the stimuli were presented very briefly deserves further studies. Clearly, it is not that color registration is impoverished by of a brief presentation, as color facilitates recognition in the NC condition. It could be that very brief presentations did no leave enough time for such an interference to develop.

General Discussion

Experiment 1 and 2 showed that color influences the speeded recognition of scene categories, but only when color is diagnostic. This effect has not been reported before in the scene recognition literature (but see Tanaka & Bunoski, 1996, for object recognition), and doubts have been raised about the control of diagnosticity and luminance in object recognition studies. Our results also suggest that color attributes are stored in memory along with luminance attributes (see Gegenfurtner et al., 1995). Of course, the absence of an effect for artifact scenes does not mean that color is not represented, but that it is less effective in this case as a categorization cue for recognition at a glance.

These results should have a strong impact on computational models of scene recognition which are more traditionally based on luminance (though see Liu & Yang, 1994; Uchiyama & Arbib, 1994). Speeded recognition of natural scenes requires an efficient encoding of different scene aspects. The results suggest that color should now be used as a supplementary categorical cue for complex image recognition. The integration of color and luminance information over time will be the topic of further studies.

References

Biederman, I. (1987). Recognition-by-components : A theory of human image understanding. *Psychological Review*, 94, 115-147.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, 20, 38-64.

Billock, V.A. (1995). Cortical simple cells can extract achromatic information from the multiplexed chromatic and achromatic signals in the parvocellular pathways. *Vision Research*, 35, 2359-2369.

Cavanagh, P. & Ramachandran, V. S. (1988). Structure from motion with equiluminous stimuli. *Annual Meeting CPA, Montreal.*

Damasio, A., Yamada, T., Damasio, H., Corbett, J, & McKee, J. (1980). Central achromatopsia: behavioral, anatomic, and physiological aspects. *Neurology*, 30, 1064-1071.

Gegenfurtner, K.R., Sharpe, L.T., & Wichmann, F.A. (1995). The contribution of color to recognition memory in normal and color-deficient observers. *Perception*, supp. 24, 12.

Liu, J. & Yang, Y. (1995). Multiresolution color image segmentation. *IEEE:PAMI*, 16, 689-700.

Logothetis, N. K., Schiller, P. H., Charles, E. R., & Hurlbert, A. C. (1989). Perceptual deficits and the role of color-opponent and broad-band channels in vision. *Science*, 247, 214-217.

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of the three-dimensional shapes. *Proceedings of the Royal Society of London*, B200, 269-294.

McIlhaga, W., Hine, T., Cole, G. R., & Snyder, A. W. (1990). Texture segregation with luminance and chromatic contrast. *Vision Research*, 30, 489-498.

Oliva, A., & Schyns, P.G. (1995). Mandatory scale perception promotes flexible scene categorizations. *17th Annual Conference of the Cognitive Science Society*, Pennsylvania, 159-163.

Ostergaard, A.L., & Davidoff, J.B. (1985). Some effects of color on naming and recognition of objects. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 11, 579-587.

Schyns, P.G., & Oliva, A. (1994). From blobs to boundary edges : evidence for time and spatial scale dependant scene recognition. *Psychological Science*, 5, 195-200.

Tanaka, J.W., & Bunoski, L.M. (1996). Is the object recognition system really colorblind? Submitted for publication.

Uchiyama, T., & Arbib, M. A. (1994). Color image segmentation using competitive learning. *IEEE:PAMI*, 16, 1197-1206.

Wurm, L.H., Legge, G.E., Isenberg, L.M., & Luebker, A. (1993). Color improves object recognition in normal and low vision. *Journal of Experimental Psychology : Human Perception and Performance*, 19, 899-911.