massachusetts institute of technology — artificial intelligence laboratory

# Role of Color in Face Recognition

Andrew Yip and Pawan Sinha

AI Memo 2001-035 December 2001
CBCL Memo 212

## Abstract

*One of the key challenges in face perception lies in determining the contribution of different cues to face identification. In this study, we focus on the role of color cues. Although color appears to be a salient attribute of faces, past research has suggested that it confers little recognition advantage for identifying people. Here we report experimental results suggesting that color cues do play a role in face recognition and their contribution becomes evident when shape cues are degraded. Under such conditions, recognition performance with color images is significantly better than that with grayscale images. Our experimental results also indicate that the contribution of color may lie not so much in providing diagnostic cues to identity as in aiding low-level image-analysis processes such as segmentation.*

This research was supported in part by funds from the Defense Advance Research Projects Agency and by the Alfred P. Sloan Foundation Fellowship in Neuroscience to PS.

## Introduction

Our vivid perception of colors in the environment leads us to expect that color must be important for allowing us to interpret complex scenes and, specifically, recognizing the objects therein. However, the role played by color information in object recognition has been the subject of much debate. Much of the theoretical modeling and experimental work in recognition has focused on shape cues. A theory that has enjoyed enduring popularity is that objects may be encoded primarily in terms of their luminance-defined bounding edge structure [Biederman, 1987; Grimson, 1990] with surface properties such as color and texture being of little consequence. Researchers on the other side of this debate have proposed models of recognition that rely strongly on color cues [Swain and Ballard, 1991].

Mirroring the debate in the theoretical domain, experimental studies investigating the effects of color on recognition of non-face objects have yielded mixed results. Ostergaard and Davidoff (1985) found that achromatic and incorrect-color images were recognized with as much accuracy as their veridical counterparts. When Biederman and Ju (1988) compared recognition performance for color images and line drawings of the same objects, they found no advantage for the color stimuli. Other researchers (Price and Humphreys, 1989; Humphrey et al., 1994; Wurm et al. 1993; Tanaka and Presnell, 1999) however, showed that objects were named faster in color photographs than in uncolored outline drawings or gray-scale images. De Valois and Switkes (1983) suggested that color constancy may aid image segmentation by allowing for improved discrimination between reflectance and illumination contours.

A relatively small body of research has examined the contribution of color in face-recognition. A notable study in this regard was conducted by Kemp and his colleagues [Kemp et al., 1996]. The researchers found that observers were able to process quite normally even those faces that had been subjected to hue-reversals (tasks included recognizing familiar faces or spotting differences between faces). Color appeared to confer no significant recognition advantage beyond the luminance information. In explaining these data, Bruce and Young (1998) have suggested that the lack of a contribution of color cues to face recognition is because they do not affect shape from shading processes, which are believed to be largely 'color-blind' (Cavanagh and Leclerc, 1987).

However, another possible reason for the observed lack of a contribution of color to face recognition in these studies is that in situations where strong shape cues are available (as in high-resolution face images), performance may already be at ceiling and the contribution of color may not be evident. The interesting condition to examine would be one wherein shape cues are progressively degraded. If color does play a role in face recognition, its contribution would be more evident under such conditions. Our experiments are designed to test this hypothesis. We shall focus specifically on degradation by blurring because of its high ecological significance and validity. This degradation mimics the reduction in image information that accompanies increasing viewing distances or common refractive errors in the eye.

## Methods

*Subjects*

A total of 37 subjects participated in the study. The subjects' ages ranged from 18-40, with a mean of 22. Subjects were recruited from the MIT Brain and Cognitive Sciences Department subject pool. All subjects had normal visual acuity or wore appropriate corrective lenses.

*Stimuli*

Frontal images of 24 celebrity faces (12 male, 12 female) were selected as the experimental stimuli. The faces were rotoscoped from their respective backgrounds and the distance between the eyes was normalized to be uniform (60 pixels) across the entire set (henceforth referred to as the 'high resolution' set).

The high resolution images were then subjected to four different levels of Gaussian blur to create images with 1.5, 2, 3, and 4 cycles between the eyes. These four blurred image sets and the original high resolution image set constituted the stimuli for the 'full-color' condition. Grayscale stimuli were created by removing the hue information from these images. MATLAB software was used for all data analysis and for the Gaussian blur image transformations.

*Procedure*

Subjects were seated in front of a computer monitor without restrictions on their head movements. On an average, each face subtended 2° of visual angle at a distance of 0.6 meters. Subjects were randomly divided into the 'color' and 'grayscale' groups. 19 subjects were assigned to the color condition, 18 to the grayscale condition. The two groups were mutually exclusive and each subject saw only one of the two sets of stimuli (color or gray-scale). The stimuli were presented in order, proceeding from the lowest resolution images (1.5 cycles/eye-to-eye) to the high-resolution reference images. No restrictions were placed on viewing time. Subjects were asked to identify the faces and to record the celebrity's name on a response sheet. A separate response sheet was provided for each resolution level. Subjects were free to change their answers from one resolution level to another.

## Results

Figure 1 summarizes our results. The graph shows how subjects' face identification performance for color and grayscale images changes as a function of available image resolution. Like Kemp et al (1996), we found that recognition performance with grayscale images was not significantly different from that with color images at high resolutions. However, performance for the two groups diverged as image resolution was progressively decreased. At low resolutions, performance with color images was significantly better than that with grayscale images. Pair-wise t-tests revealed that the two conditions were significantly different when the facial image resolution was 1.5 cycles/eye-to-eye ($p = 0.0053$) or 2 cycles/eye-to-eye ($p < 0.001$). These results support the hypothesis that color cues do enhance recognition performance and their contribution becomes very evident when shape information is degraded.

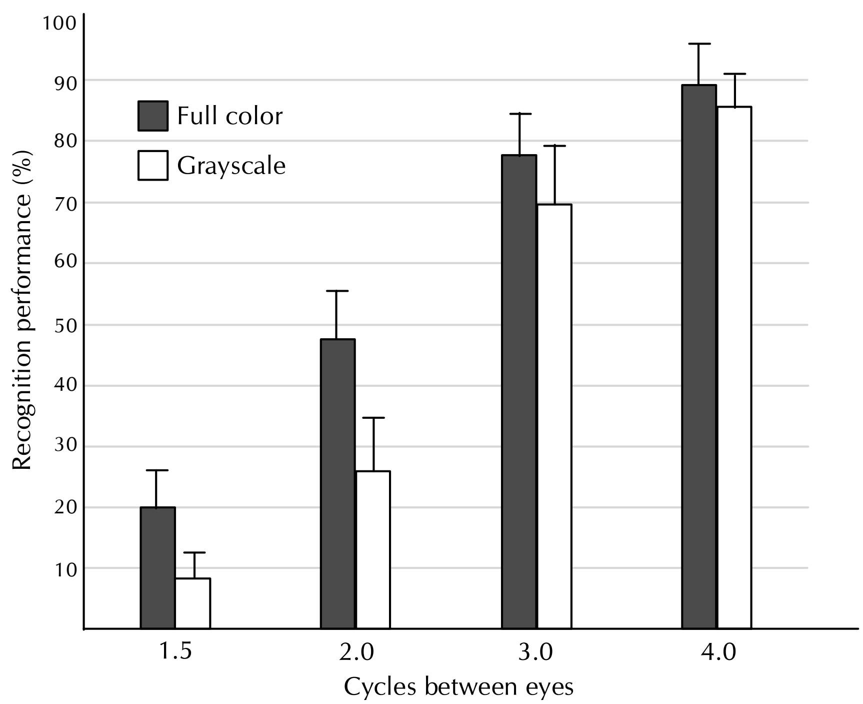


**Figure 1.** Face recognition performance with full-color and grayscale images as a function of image resolution. Consistent with the work of Kemp et al [1996], we found no significant difference in performance between the color and grayscale conditions at high-resolutions. However, as shape cues were degraded by reducing image resolution, performance with color images significantly exceeded that with gray-scale ones. Since subjects were not pre-cued about which celebrities they were going to be tested on, chance-level performance is close to zero.

This evidence of color's contribution to face recognition brings up an interesting question regarding the specific role it plays. One possibility is that color provides diagnostic information – the precise hue of a person's hair or skin may allow us to identify them. On the other hand, color might facilitate low-level image analysis, and thus indirectly aid face recognition. An example of such a low-level task is image segmentation – determining where one region ends and the other starts. As many years of work in computer vision has shown [Haralick, 1985; Felzenszwalb and Huttenlocher, 1998], this task is notoriously difficult and becomes even more intractable as images are degraded. Color may facilitate this task by supplementing the luminance-based cues and thereby lead to a better parsing of a degraded face image in terms of its constituent regions. An interesting open question, therefore, is: which of these two possibilities (providing diagnostic information for identification versus facilitating low-level analysis) is the primary way in which color contributes to face recognition?

To address this question, we conducted an additional experiment with variants of the face stimuli described earlier. We created 'pseudo-color' versions of these faces wherein they were assigned unnatural colors. Such stimuli dispense with the diagnostic aspects of color cues but preserve the low-level cues that may aid in tasks like segmentation. To create the stimuli for the pseudo-color condition, each high-resolution full-color face was subjected to a hue transformation involving a partial rotation around the color wheel. Figure 2 shows a sample face in grayscale, full-color and pseudo-color conditions. The luminance channel of the resulting pseudo-color image was maintained to be identical to the luminance channel of the original full-color image. Luminance was therefore controlled across all three (full-color, grayscale and pseudo-color) experimental conditions. Finally, the high-resolution pseudo-color images were prepared at the same four levels of Gaussian blur that were used with the full-color and grayscale stimuli.

**Figure 2.** A sample face in grayscale, full-color and pseudo-color conditions.

Our subjects comprised ten individuals who had not participated in any of the earlier experiments. The instructions to these subjects and their task was the same as for the prior ones. We reasoned that if color-cues were used to provide diagnostic information, then performance with pseudo-color stimuli would be quite poor, similar to the gray-level condition (or possibly worse, due to interference from unnatural color cues). If, however, color primarily assists in low-level tasks such as segmentation, then we would expect performance with pseudo-color images to be on par with performance with full-color images.

Our results, shown in figure 3, support the latter possibility. Pairwise t-tests indicated that the full-color and pseudo-color data were not significantly different for any of the blur conditions ($p > 0.39$ on all four tests). A t-test comparison of the grayscale and pseudo-color conditions exhibited the same pattern seen in the grayscale to full-color comparison; performance differences were significant when the image resolution was 1.5 cycles/eye-to-eye ($p < 0.002$) or 2 cycles/eye-to-eye ($p < 0.001$). In summary, performance with pseudo-color images is statistically the same as that with full color images, and significantly better compared to gray-scale images at low-resolutions.

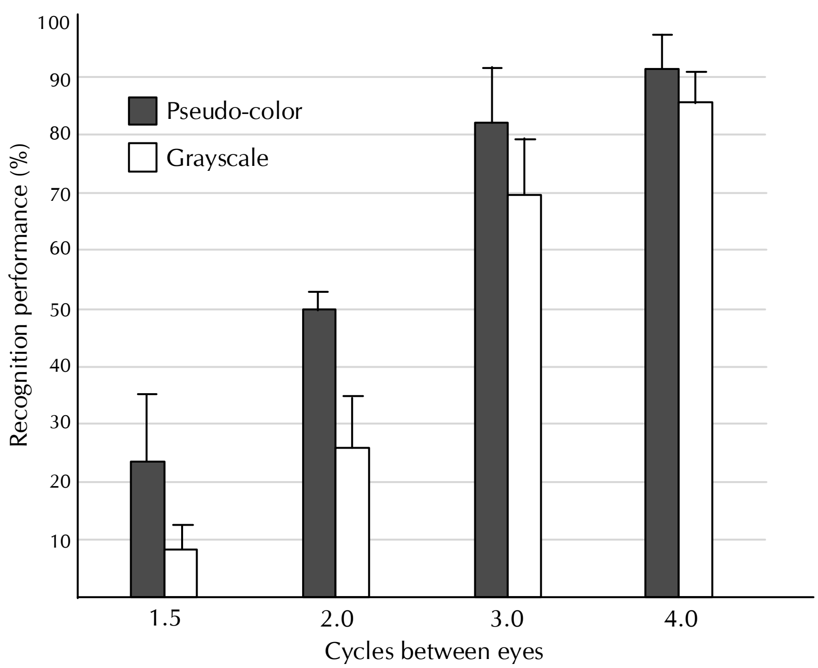


**Figure 3.** Face recognition performance with pseudo-color images relative to full-color and grayscale conditions. At all resolutions tested, performance with pseudo-color images is statistically the same as with full-color images and is significantly better than with gray-scale images at low resolutions. This pattern of results suggests that color cues may facilitate low-level image analysis rather than providing precise hue-related diagnostic cues to identity.

A two-factor ANOVA on the combined data from all three experimental conditions showed highly significant main effects of blur level ($F_{4, 220} = 346.9$, $p < 0.001$) and color condition ($F_{2, 220} = 17.03$, $p < 0.001$), as well as a significant blur level × color condition interaction ($F_{8, 220} = 2.517$, $p = 0.0123$).

## Discussion

The luminance structure of face images is undoubtedly of great significance for recognition. Past research has suggested that the use of these cues may adequately account for face-identification performance with little remaining need to posit a role of color information. However, our experimental results suggest that color does in fact play an important role in face recognition and that its contribution becomes evident under conditions that degrade available shape information. Furthermore, our results with pseudo-color images indicate that color cues may contribute to recognition primarily by facilitating low-level image analysis tasks such as segmentation rather than by providing specific diagnostic information. Figure 4 illustrates this idea. The images show the luminance and color components of sample face inputs. The color distributions can supplement luminance information to allow for a better estimation of the boundaries, shapes and sizes of facial attributes such as eyes and hair-lines (see figure caption for details). It is worth emphasizing that the segmentation task referred to here is not segmentation of the face from the background, since the background was uniform white for all faces, but rather segmentation of features within a face (such as where the forehead ends and the hairline begins). Also, it is difficult to explain the pseudo-color results by invoking the idea of color constancy rather than facilitation of a low-level operation like segmentation. Not only is constancy imperfect for isolated objects (such as our face stimuli which sit on featureless white backgrounds) [Brainard and Wandell, 1992], but the hue transformation used cannot be characterized as an overall illuminant change.

Our results do not rule out a diagnostic contribution when the color attributes for an individual are distinctive, as Lee and Perrett (1997) found. However, in general, since faces have similar spectral properties and since the observed colors can change significantly under different illumination conditions, it makes sense for the visual system not to rely on color as a diagnostic cue, but rather for tasks such as segmentation where differences instead of absolute values are sufficient.

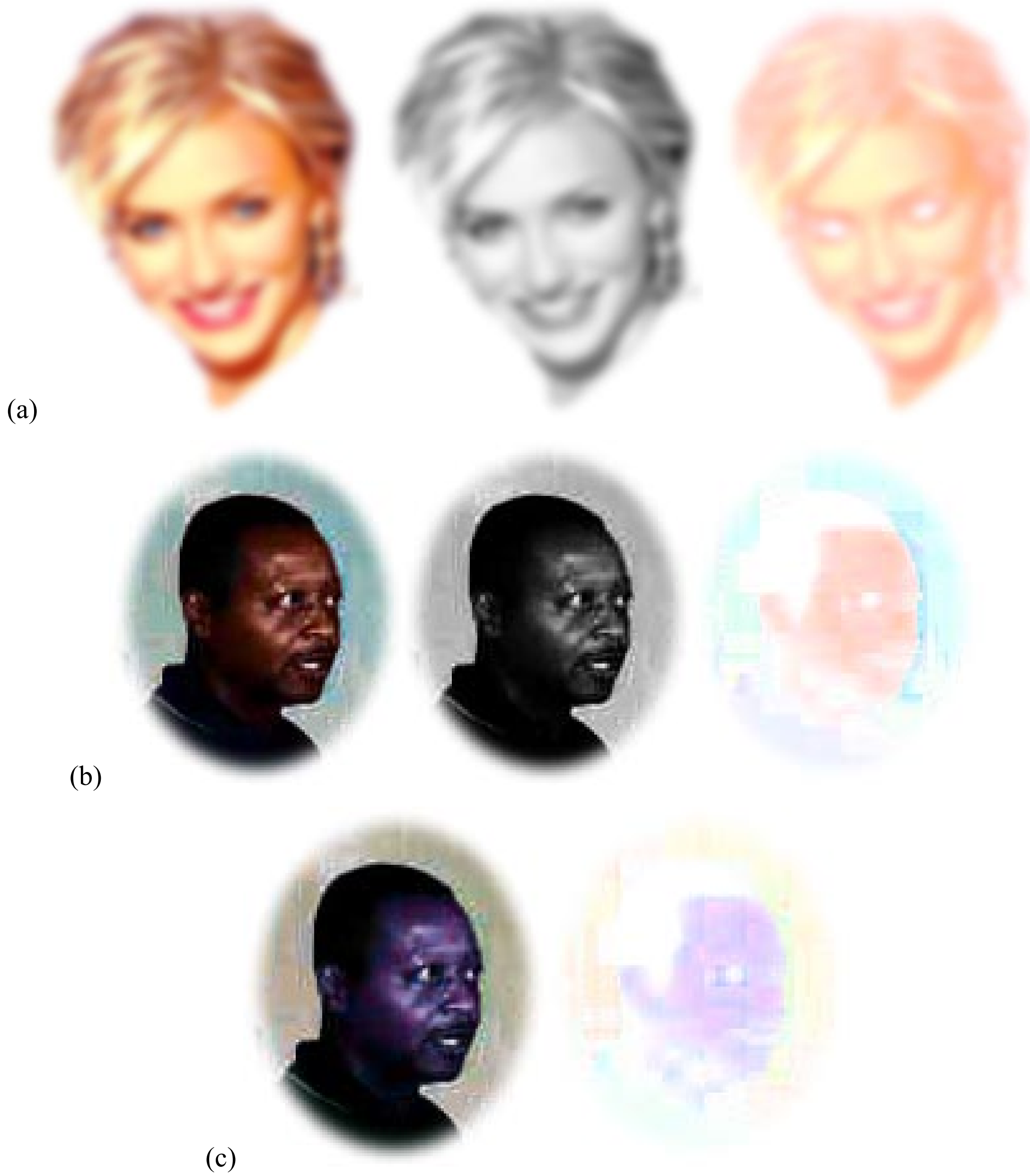


**Figure 4.** Examples that illustrate how color information may facilitate some important low-level image analysis tasks such as segmentation. In (a), the hue distribution (right panel) allows for a better estimation of the shape and size of the eyes than the luminance information alone (middle panel). Left panel shows the original image. Similarly, in (b), hue information (right panel) allows for a better segmentation and estimation of the location and shape of the hair-line than just the luminance information (middle panel). This facilitation of low-level analysis happens with other choices of colors as well, such as in the pseudo-color image shown on the left in (c). The hue distribution here, as in (b) aids in estimating the position of facial attributes such as the hair-line.

It is to be noted that although precise estimates of hue do not appear critical for face identification, they may be important for some other face perception tasks. Fine differences between hues may, for instance, allow an observer to assess an individual's emotional state ('flushed with anger' or 'pale with fear') and also to judge the state of their health. Availability of hue information can also benefit the task of face-detection (searching for a face in a large scene) by reducing the problem from a complex pattern matching one to a simpler color search task. Several computational systems do, in fact, make use of this intuition [Terrillon et al., 1998; Sun et al., 1998; Kumar and Poggio, 2000].

To the extent that our results suggest that color does not provide diagnostic information for face-recognition, they are consistent with earlier reports [Kemp et al., 1996]. However, they go beyond those findings by suggesting that color cues are not entirely disregarded by the face-recognition processes. Rather, they contribute significantly under degraded conditions by facilitating low-level facial image analysis.

## Acknowledgments

The authors wish to thank Jim Dannemiller, Antonio Torralba and Richard Russell for helpful discussions about the ideas presented in this paper.

## References


Biederman, I. (1987). Recognition by components: a theory of human image understanding. *Psychological Review*, **94**(2), 115-147.

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

Brainard, D. H. and Wandell, B. A. (1992). Asymmetric color matching: How color appearance depends on the illuminant. *J. Opt. Soc. Am.*, **9**: 1433-1448.

Bruce, V. and Young, A. (1998). *In the eye of the beholder. The science of face perception.* Oxford University Press.

Cavanagh, P. and Leclerc, Y. G. (1989). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance,* **15**, 3-27.

De Valois, K. K. and Switkes, E. (1983). Simultaneous masking interactions between chromatic and luminance gratings. *Journal of the Optical Society of America*, **73**, 11-18.

Felzenszwalb, P., and Huttenlocher, D. (1998). Image Segmentation Using Local Variation. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 98-104.

Grimson, W. E. L. (1990). The combinatorics of object recognition in cluttered environments using constrained search, *Artificial Intelliegence*, **44**.

Haralick, R. (1985). Survey, image segmentation techniques, *Computer Vision, Graphics, and Image Processing,* **29**, 100-135.

Humphrey, G. K., Goodale, M. A., Jakobson, L. S. and Servos, P. (1994). The role of surface information in object recognition: studies of a visual form agnosic and normal subjects. *Perception*, **23**, 1457-1481.

Kemp, R., Pike, G., White, P. and Musselman, A. (1996). Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception*, **25**, 37-52.

Kumar, V. and T. Poggio. (2000). Learning based approach to real-time tracking and analysis of faces. In: *Proceedings of the Fourth International Conference on Face and Gesture Recognition*, Grenoble, France, 91-96.

Lee, K. J. and Perrett, D. (1997). Presentation-time measures of the effects of manipulations in colour space on discrimination of famous faces. *Perception*, **26**, 733-752.

Ostergaard, A. L. and 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.

Price, C. J. and Humphreys, G. W. (1989). The effects of surface detail on object categorization and naming. *The Quarterly Journal of Experimental Psychology*, **41A**, 797-828.

Sun, Q. B., Huang, W. M., and Wu, J. K. (1998). Face detection based on color and local symmetry information. *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, pp. 130-135.

Swain, M. J. and Ballard, D. H. (1991). *Color indexing*. International Journal of Computer Vision, **7**:11--32.

Tanaka, J. W. and Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics*, **61**, 1140-1153.

Terrillon, J., C., David, M. and Akamatsu, S. (1998). Automatic detection of human faces in natural scene images by use of a skin-color model and of invariant moments. *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, Japan, pp. 112-117.

Wurm, L. H., Legge, G. E., Isenberg, L. M. and Luebker, A. (1993). Color improves object recognition in normal and low vision. *J. Exp. Psych: Hum. Percep. and Perf.*, **19**, 899-911.