Perception, 2002, volume 31, pages 995 – 1003

DOI:10.1068/p3376

Contribution of color to face recognition

Andrew W Yip, Pawan Sinha¶
Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology,
45 Carleton Street, Cambridge, MA 02142, USA; e-mail: sinha@ai.mit.edu
Received 11 January 2002, in revised form 30 March 2002; published online 17 July 2002

Abstract. One of the key challenges in face perception lies in determining how different facial attributes contribute to judgments of identity. 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 gray-scale 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.

1 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 on 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; Wurm et al 1993; Humphrey et al 1994; 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 dealt with 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, Kemp et al (1996) and Bruce and Young (1998) have suggested that the lack of a contribution of color cues to face recognition is because they do

¶ Author to whom all correspondence and requests for reprints should be addressed.

not affect shape-from-shading processes, which are believed to be largely 'color-blind' (Cavanagh and Leclerc 1989).

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 focus specifically on degradation by blurring because of its high ecological significance. This degradation mimics the reduction in image information that accompanies increasing viewing distances or common refractive errors in the eye.

2 Methods

2.1 *Subjects*

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

2.2 *Stimuli*

Frontal images of twenty-four celebrity faces (twelve male, twelve 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. Gray-scale 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.

2.3 *Procedure*

Subjects were seated in front of a computer monitor without restrictions on their head movements. On an average, each face subtended 2 deg of visual angle at a distance of 0.6 m. Subjects were randomly divided into the 'color' and 'gray-scale' groups. Nineteen subjects were assigned to the color condition, eighteen to the gray-scale 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 between the eyes) 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.

3 Results

Figure 1 summarizes our results. The graph shows how subjects' face-identification performance for color and gray-scale images changes as a function of available image resolution. Like Kemp et al (1996), we found that recognition performance with gray-scale 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 gray-scale images. Pairwise t-tests revealed that the

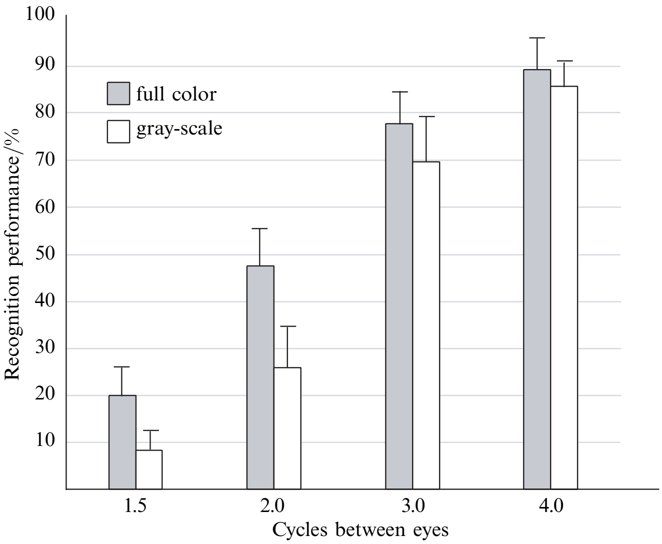


Figure 1. Face-recognition performance with full-color and gray-scale images as a function of image resolution. Consistent with the work of Kemp et al (1996), we found no significant difference in performance between the color and gray-scale conditions at high resolutions. However, as shape cues were 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.

two conditions were significantly different when the facial image resolution was 1.5 cycles between the eyes ($p = 0.0053$) or 2 cycles between the eyes ($p < 0.001$). These results support the hypothesis that color cues do enhance face-recognition performance and their contribution becomes very evident when shape information is degraded.

This evidence of the contribution of color to face recognition brings up an interesting question regarding the specific role it plays. One possibility is that color provides diagnostic information. The expression 'diagnostic information' refers to color cues that are specific to an individual, for instance the particular hue of hair or skin that may allow us to identify the individual. 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 and Shapiro 1985; Felzenszwalb and Huttenlocher 1998), this task is notoriously difficult and becomes even more intractable as images are degraded. Color may facilitate this

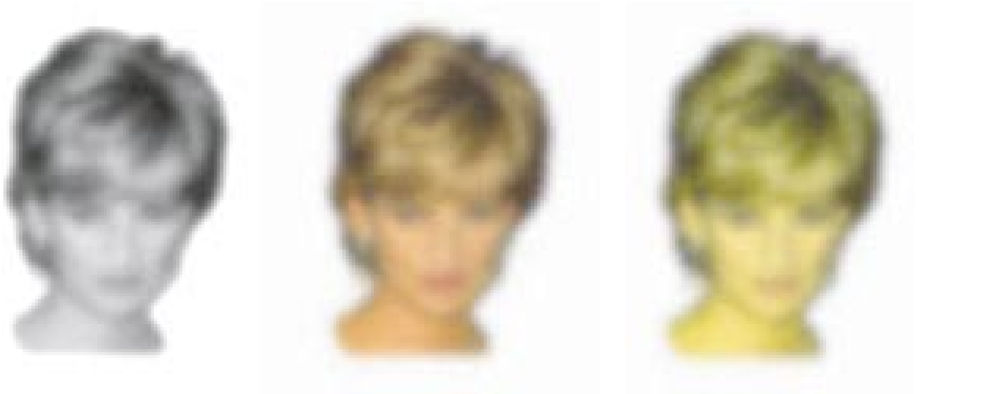


Figure 2. A sample face in gray-scale, full-color, and pseudo-color conditions.

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 gray-scale, 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, gray-scale, 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 gray-scale stimuli.

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 earlier 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-scale condition (or possibly worse, owing 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 ($p \leqslant 0.39$ on all tests). A t-test comparison of the gray-scale and pseudo-color conditions exhibited the same pattern as that seen in the gray-scale

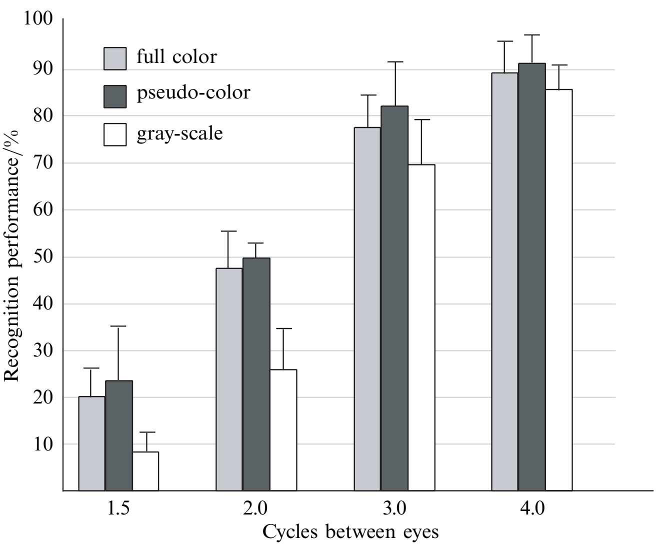


Figure 3. Face-recognition performance with pseudo-color images relative to the gray-scale condition. At all resolutions tested, performance with pseudo-color images was statistically the same as with full-color images and was significantly better than with gray-scale images at low resolutions. This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

to full-color comparison; performance differences were significant when the image resolution was 1.5 cycles between the eyes ($p < 0.002$) or 2 cycles between the eyes ($p < 0.001$). In summary, performance with pseudo-color images is statistically the same as that with full-color images, and significantly better than with gray-scale images at low resolutions.

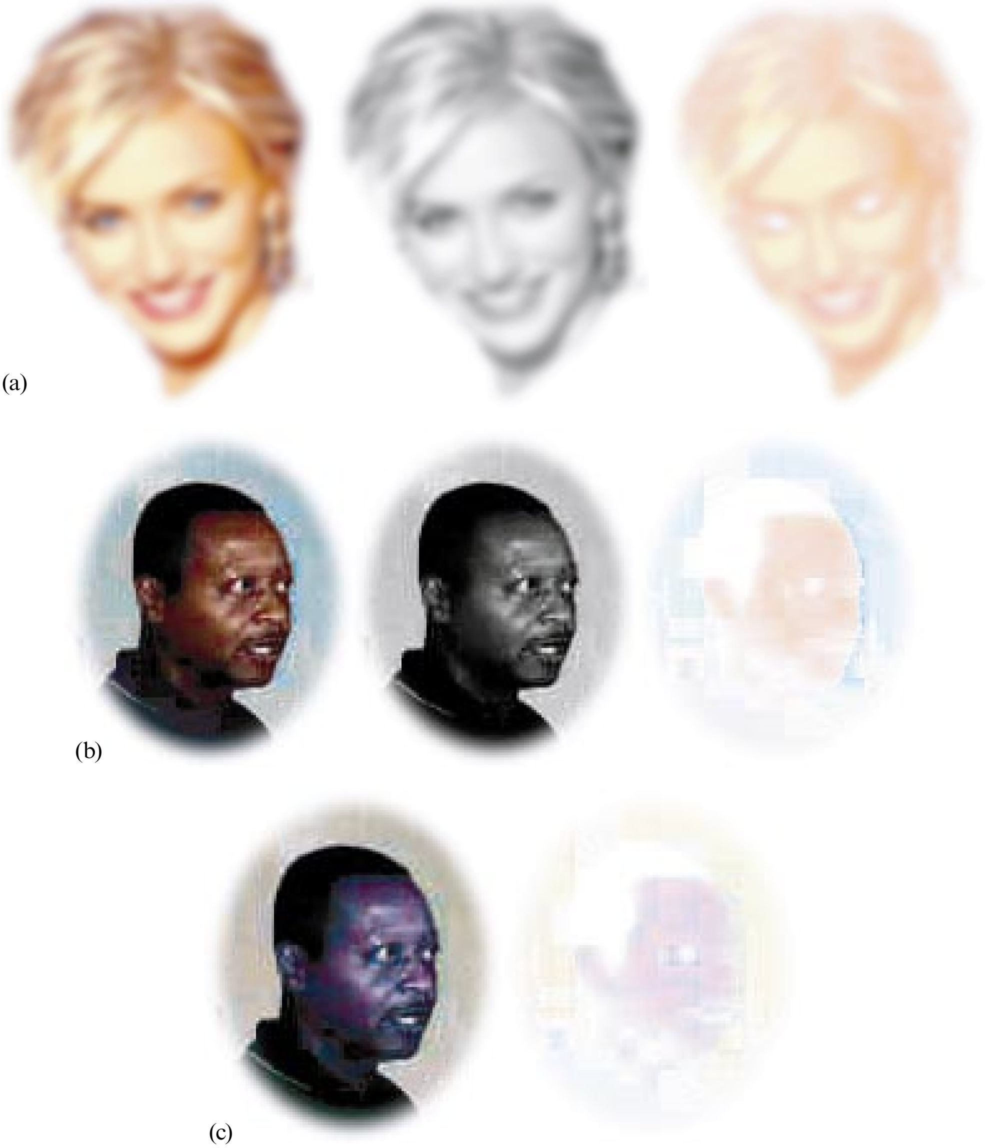
A two-factor ANOVA on the combined data from all three experimental conditions showed highly significant main effects of blur level ($F_{4,220} = 346.9$, $p < 0.001$) and color condition ($F_{2,220} = 17.03$, $p < 0.001$), as well as a significant blur-level color-condition interaction ($F_{8,220} =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

4 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. They suggest that 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 hairlines (see figure caption for details). A note about how the hue maps were generated. The images were represented in CIE Lab format, which dissociates the luminance and color signals. The Lab color model is a refinement of the Commission Internationale d'Eclairage (CIE) model originally proposed in 1931. The Lab image format consists of three channels. The first channel is luminance (L) while the other two channels, a and b, represent color ranges. The a channel contains colors ranging from green to red, and the b channel contains colors ranging from blue to yellow. The hue maps shown in figure 4 were derived by discarding the luminance channel.

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.

Given the kind of stimuli we have used in our experiments, a natural question to ask is whether an alternative way to examine the use of color for face recognition is to remove the luminance variations in the image. While this is an interesting image manipulation, it takes us in a direction that is a little different from what we want to stress in the paper. The point the paper makes is not that color information on its own is sufficient for recognition, but rather that the presence of color extends performance beyond the luminance-only condition, possibly by facilitating tasks like segmentation and edge localization. It is not our intention to suggest that the two sources of information—luminance on the one hand and color on the other—are used in fundamentally different ways: luminance for shape and color for segmentation. The emphasis we place on the role of color in segmentation and edge localization is because color information, on its own (as embodied in hue maps of faces), is



(a) (b) (c)

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 hairline 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 hairline.

largely devoid of shading information and thus not very useful for assessing 3-D structure information. However, given that many facial landmarks, such as eyes, mouth, and hairline, are often defined both by luminance as well as hue changes, color information can aid in the recovery of 2-D shape information. This is indeed what we mean when we talk about color providing segmentation information: segmentation implies not only a separation into regions, but also a determination of where the region boundaries are and the shapes of the regions. This idea also gains support from the work of Mullen and colleagues on the contribution of color cues to form perception (Mullen and Kingdom 1991; McIlhagga and Mullen 1996, 1997). A possible explanation why the contribution of color becomes more evident with blur may lie in the different contrast-sensitivity functions for luminance and color contrast (Webster et al 1990; Losada and Mullen 1995). Given the lower acuity for color, blurring has a greater perceptual effect on luminance rather than color. The latter may, therefore, provide a more stable input under different blur manipulations.

Given the manipulations we have used in our experiments, a question we need to address is whether studying recognition with such highly degraded images is too contrived a setting to be worthwhile. There are two ways to address this concern. First, several studies such as those by Harmon (1973), Harmon and Julesz (1973), and Costen et al (1994) have demonstrated the high level of performance obtainable even with very low spatial frequencies. Thus, such stimuli are very likely to be tapping into important face-recognition mechanisms. Second, understanding how we are able to interpret and recognize such impoverished images may be of great ecological significance, since it touches upon recognition under sub-optimal viewing conditions so common in the real world. We often have to recognize people even when their images are degraded through refractive errors in our eyes, haze, or sensor noise in electronic cameras. Studying such stimuli is, therefore, inherently interesting and important. The results also touch upon the issue of recognition at a distance. However, in this context, it is important to note that while the increasing amount of blur used in our study does capture the loss of information associated with increasing viewing distance, the two manipulations are not entirely equivalent. Specifically, a distant face has lower resolution, but does not look blurry. An interesting question for further investigation is whether color facilitates recognition at large distances in the same way that it does for the highly degraded images we used here.

The results of our experiments open up an interesting debate from the perspective of visual neurophysiology. It is generally believed that the parvocellular channel is the dominant stream responsible for the processing of color information (De Yoe and Van Essen 1988; Livingstone and Hubel 1988). However, this channel is also thought to be specialized for high spatial frequencies. And yet, the experimental data reported above suggest that low spatial frequencies show greater benefit with the addition of color in terms of face-recognition performance. How might one account for this seeming puzzle? There are at least two possibilities. First, even though the parvo and magno streams are relatively segregated in the early stages of the visual system, there is little evidence to suggest that this segregation is strictly maintained in later stages as well. There is now significant evidence of crosstalk between the dorsal and ventral streams (the later, and approximate, continuations of the magno and parvo streams). Thus, color information can conceivably contribute to processing in both these pathways. Second, and more importantly, luminance and color information may encode spatial structure somewhat redundantly. The analysis of color information, therefore, may facilitate recognition even when the luminance information is highly degraded. In other words, it is not the low-spatial-frequency channels themselves that are benefiting from the inclusion of color, but recognition at low spatial frequencies.

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. Indeed, an alternative way of accounting for the small increase in recognizability that Lee and Perrett observed with 'color caricatures' is that exaggeration of the hue signal also strengthens color-based low-level cues.

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 individuals' state of health. Recently, Tarr and colleagues have suggested that facial color cues may be involved in the task of gender classification (Tarr et al 2001). 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 (Sun et al 1998; Terrillon 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.

Can we extrapolate from the results reported here and make hypotheses about the role of color in object recognition in general? It is likely that for object-discrimination tasks where color is a diagnostic cue, say when distinguishing between different varieties of flowers, color may play a role beyond just assisting with low-level image analysis. The results may generalize to the recognition of other object classes that are, like faces, relatively similar in their chromatic structure. However, in order to be scientifically conservative, we believe that the results most directly apply to the domain of faces.

Acknowledgments. 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. The authors wish to thank Jim Dannemiller, Dan Kersten, and Keith Kluender 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** 115 – 147

Biederman I, Ju G, 1988 "Surface versus edge-based determinants of visual recognition" *Cognitive Psychology* **20** 38 – 64

Brainard D H, Wandell B A, 1992 "Asymmetric color matching: How color appearance depends on the illuminant" *Journal of the Optical Society of America A* **9** 1433 – 1448

Bruce V, Young A, 1998 *In the Eye of the Beholder. The Science of Face Perception* (Oxford: Oxford University Press)

Cavanagh P, Leclerc Y G, 1989 "Shape from shadows" *Journal of Experimental Psychology: Human Perception and Performance* **15** 3 – 27

Costen N P, Parker D M, Craw I, 1994 "Spatial content and spatial quantisation effects in face recognition" *Perception* **23** 129 – 146

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

De Yoe E A, Van Essen D C, 1988 "Concurrent processing streams in monkey visual cortex" *Trends in Neurosciences* **11** 219 – 226

Felzenszwalb P, Huttenlocher D, 1998 "Image segmentation using local variation", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Los Alamitos, CA: IEEE Computer Society Press) pp 98 – 104

Grimson W E L, 1990 "The combinatorics of object recognition in cluttered environments using constrained search" *Artificial Intelligence* **44** 121 – 166

Haralick R, Shapiro L G, 1985 "Survey: image segmentation techniques," *Computer Vision, Graphics, and Image Processing* **29** 100 – 135

Harmon L D, 1973 "The recognition of faces" *Scientific American* **229**(5) 70 – 83

Harmon L D, Julesz B, 1973 "Masking in visual recognition: Effects of two-dimensional noise" *Science* **180** 1194 – 1197

Humphrey G K, Goodale M A, Jakobson L S, 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, 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, Poggio T, 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* (Los Alamitos, CA: IEEE Computer Society Press) pp 91 – 96

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

Livingstone M S, Hubel D, 1988 "Segregation of form, color, movement and depth: Anatomy, physiology and perception" *Science* **240** 740 – 749

Losada M A, Mullen K T, 1995 "Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking" *Journal of the Optical Society of America A* **12** 250 – 260

McIlhagga W H, Mullen K T, 1996 "Contour integration with color and luminance contrast" *Vision Research* **36** 1265 – 1279

McIlhagga W H, Mullen K T, 1997 "The contribution of color to contour detection", in *Colour Vision Research: Proceedings of the John Dalton Conference* Eds C M Dickenson, I Murray, D Carden (London: Taylor and Francis) pp 187 – 196

Mullen K T, Kingdom F A, 1991 "Color contrast in form perception", in *The Perception of Colour* Ed. P Gouras (Basingstoke, Hants: Macmillan) pp 198 – 217

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

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

Sun Q B, Huang W M, Wu J K, 1998 "Face detection based on color and local symmetry information", in *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan* (Los Alamitos, CA: IEEE Computer Society Press) pp 130 – 135

Swain M J, Ballard D H, 1991 "Color indexing" *International Journal of Computer Vision* **7** 11 – 32

Tanaka J W, Presnell L M, 1999 "Color diagnosticity in object recognition" *Perception & Psychophysics* **61** 1140 – 1153

Tarr M J, Kersten D, Cheng Y, Rossion B, 2001 "It's Pat! Sexing faces using only red and green" *Journal of Vision* **1**(3) 337a (abstract) http://journalofvision.org/1/3/337, DOI:10.1167/13337

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

Webster M A, De Valois K K, Switkes E, 1990 "Orientation and spatial frequency discrimination for luminance and chromatic gratings" *Journal of the Optical Society of America A* **7** 1034 – 1049

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