Color stereoscopic images requiring only one color image

Yael Termin
Gal A. Kaminka
Bar Ilan University
Computer Science Department
Ramat Gan, 52900
Israel
E-mail: terminy@cs.biu.ac.il

Sarit Semo
Raanana, Israel

Ari Z. Zivotofsky
Bar Ilan University
Brain Sciences Center
Ramat Gan, 52900
Israel

Abstract. Utilizing remote color stereoscopic scenes typically requires the acquisition, transmission, and processing of two color images. However, the amount of information transmitted and processed is large, compared to either monocular images or monochrome stereo images. Existing approaches to this challenge focus on compression and optimization. This paper introduces an innovative complementary approach to the presentation of a color stereoscopic scene, specialized for human perception. It relies on the hypothesis that a stereo pair consisting of one monochromatic image and one color image (a MIX stereo pair) will be perceived by a human observer as a 3-D color scene. Taking advantage of color redundancy, this presentation of a monochromatic-color pair allows for a drastic reduction in the required bandwidth, even before any compression method is employed. Herein we describe controlled psychophysical experiments on up to 15 subjects. These experiments tested both color and depth perception using various combinations of color and monochromatic images. The results show that subjects perceived 3-D color images even when they were presented with only one color image in a stereoscopic pair, with no depth perception degradation and only limited color degradation. This confirms the hypothesis and validates the new approach. © *2007 Society of Photo-Optical Instrumentation Engineers.*
[DOI: 10.1117/1.2772235]

Subject terms: stereoscopic image compression; video bandwidth reduction; binocular vision; stereopsis; color perception; depth perception.

Paper 060586R received Jul. 27, 2006; revised manuscript received Feb. 1, 2007; accepted for publication Feb. 9, 2007; published online Aug. 21, 2007.

1 Introduction

Color and depth perception are important capabilities in everyday human activities and in many applications requiring human supervision or control. Both properties, color and stereoscopic perception, are of significance. Binocular (stereoscopic) vision provides us with the ability to determine the 3-D structure and the relative and absolute distance of objects present in our field of view (FOV). Color perception allows segmentation and evaluation of the perceived objects.

Stereopsis, which is perhaps the most important mechanism of depth perception, depends on the use of both eyes. Two disparate images (via the two eyes) are presented to the brain and combined into a single image that yields stereoscopic depth perception. Although many of the cues used to judge depth involve only one eye (viz., relative size, parallax, texture, interposition, relative position, linear perspective, shadows), these monocular depth cues result in less depth-estimation accuracy than disparity, which can only be obtained from binocular vision. Thus, by using binocular disparity information, humans are able to make fine depth judgments that they cannot make when using just one eye.[1]

Color is additionally important in many applications of 3-D imaging.[2,3] Color is used in segmenting the images, allowing the viewer to distinguish between objects in the image and between objects and their background. It is also important in distinguishing types of objects or their status (e.g., in medical applications).

Unfortunately, color stereo imaging requires large capacity of the transmission channel. Some approaches to this challenge focus on compression and optimization of both images by various means.[4–7] Other approaches exploit the fact that the left and right images differ only in small areas, and thus improved compression can be achieved by taking advantage of the redundancies in the stereo pair. These techniques often rely on disparity compensation,[8,9] a process that can be computationally complex. Nearly all of the techniques used for stereo compression assume at least two cameras that yield similar chromatic information. Section 2 includes a review of related work.

This paper evaluates a novel complementary approach in which a single color image is combined with a single monochromatic (grayscale) image. It relies on the hypothesis that the sophisticated human visual perception system will fuse the two images that differ in chromatic content into one perceived color stereoscopic image. Thus, it is specifically targeted towards human perception. Using this approach, bandwidth requirements may be reduced even before any compression technique is applied.

We experimentally validate the hypothesis by conducting a set of experiments with human subjects, evaluating their perception of depth and color in stereoscopic images generated from color-color pairs, color-gray pairs (MIX),

and gray-gray pairs. The image pairs were presented via either a stereoscope or a head-mounted display.

This paper is organized as follows. Section 2 discusses related work and motivation for the approach. The experimental procedure designed to test the proposed approach is described in Sec. 3. The experimental results are presented in Sec. 4. We discuss the implications of the results in Sec. 5. Section 6 concludes and outlines future work.

2 Background and Motivation

Color and depth are both important in real-world applications. Batavia and Singh[2] describe an obstacle detection methodology for robotics, which combines two complementary vision-based methods: color segmentation and stereo-based color homography. Siegal and Akiya[3] present a methodology of obstacle detection in autonomous robot-incorporated systems, based on color and stereo segmentations. Depth and color are tightly coupled also in computer vision applications such as tracking multiple 3-D moving objects (e.g., cars, humans).[10,11] For instance, Darrell et al.[12] combined depth from stereo with skin color and face detection in order to track moving people.

Drascic and Grodski[13] show that the addition of a third dimension to displays in human remote-controlled teleoperation improves the decision-making capabilities of the operator, while adding color improves the operator's recognition and obstacle detection performance. An integration of two detection methodologies, one using color segmentation and the other depth from stereo, enhances the efficiency of both.

A stereo pair can be produced from nearly any kind of camera (still or video, digital or analog, color or monochrome) and viewed in a variety of displays. In order to acquire the stereo pair to be viewed in a stereoscopic display system, the system needs to mimic the views seen by the two eyes, thus requiring that two horizontally displaced cameras be used. The principles are the same whether video cameras or still images are being used. The stereoscopic display system presents the left and right perspective separately to each eye. These images can be viewed in a variety of techniques, from a simple computer monitor with active shutter glasses that sequentially blocks one eye from viewing the monitor while the other eye's image is displayed, to more sophisticated head-mounted displays (HMDs), where each eye has its own screen. These recent developments allow the addition of 3-D realism to 2-D displays.

However, color stereoscopic imaging comes at the cost of increasing the raw data to be transmitted in the channel between the two color cameras and the displays. An additional cost factor is the need for placing two full-resolution color cameras at the sensing end of the device.

Two of the most important parameters in the design of a stereoscopic device are the memory size and transmission channel capacity necessary for the storage and transfer of the stereoscopic images. Thus, much research effort today is dedicated to the optimization and compression of a stereoscopic image, and this has led to the proposal of various compression schemes for stereo images and stereo video streams.

Various monoscopic image-compression approaches have been developed, and are continually refined and improved (e.g., the JPEG family of compression standards[14]).

However, while these methods can successfully reduce the bandwidth requirements of each channel, they fail to take advantage of the large redundancy in stereoscopic imaging due to the fact that both channels contain information about the same objects, albeit taken from slightly different perspectives.

Information about the disparity between the channels—the amount of shift one needs to perform on the pixels within one image (target) to locate the corresponding pixels in the other image (reference)—can be used to further reduce the combined bandwidth requirements of the two channels. This redundancy of information between the two images was already noticed by Yaroslavsky more than 20 years ago.[15] Today, many stereo compression schemes utilize the correlation between the two stereo images, the *inter frame redundancy*.[7,16] These methods exploit the redundancy between the two images by using disparity-compensated prediction. Such predictions depend on estimating the magnitude and direction of the disparity.

Several stereo compression methods have been developed based on the concept of disparity estimation. Lukacs[17] takes advantage of the binocular redundancy by using a pixel-based stereo disparity. However, most of the compression techniques use fixed block-size based disparity estimation schemes.

Woo and Ortegs[7,9] use a block-based disparity estimation method rather than estimation based on pixel or feature, and claim it to be simpler and more effective to implement.[18] Whereas Woo and Ortega[7] estimate disparity based on a small block, Sethuraman et al.[4–6] present methods for stereo compression based on the concept of multi-resolution disparity estimation. A single image and a disparity map are used to synthesize a second image for the stereoscopic pair. Other methods for the compression of a stereoscopic image apply different compression strategies to each of the images separately, while using information about the disparity between the channels.

In fact, any of the techniques used in motion compensation (for example, MPEG,[19,20] which works at a block level) are applicable to disparity estimation.[8] The disparity-estimation/disparity-compensation approach resembles motion-estimation/motion-compensation methods, which are popular for video coding. Intraframe redundancy can be found as well between any two sequential image frames or between left and right 3-D stereoscopic image pairs and can be used in order to achieve significant data compression.[21]

Kim et al.[22] suggest a scheme for high-resolution 3-D stereoscopy via monoscopic HDTV bandwidth, using three cameras. They use a monoscopic color camera to construct the *main stream*, and two monochromatic low-resolution cameras to create a disparity map. The compressed disparity map is transmitted as a low-bandwidth *auxiliary stream*, and a synthesized color pair is computed on the receiver end using the main and auxiliary streams. However, this method requires three synchronized cameras.

While the preceding methods take advantage of the *spatial* similarity between the stereo images, other methods[23] use frequency as the basis for compression, without relying on any complex calculation of disparity compensation.

Other investigators have suggested relying on the versatile capabilities of the human visual system to further reduce the bandwidth requirements of the stereoscopic color

image, before compression. The motivation for such approaches is that there is much research that deals with the interaction impact of binocular vision and color. For instance, it is well known[24] that chromatic information helps address the correspondence problem in stereo vision. According to Shevell and Miller,[25] only a small difference in chromatic adaptation is caused by introducing a 3-D representation of the stimuli. This implies that color and depth perception are only loosely coupled and therefore might be processed separately. There is evidence that there is a reduced contrast threshold for binocular detection as compared with monocular detection.[26] This means that there is a sort of interocular facilitation that enhances binocular performance in detection tasks.[27]

Thus, investigators have examined approaches in which two different images are presented to the two eyes. Such a procedure may result either in a fused stereoscopic image (as in stereograms), or in *binocular rivalry*. Binocular rivalry is a form of multistable perception that occurs when the two eyes are presented with visual stimuli that are different from each other and cannot be fused into a single coherent percept. Under these conditions, the percept typically alternates between two states corresponding to the left eye's stimulus and the right eye's stimulus, or between two gestalts that are formed by combining parts of the monocular stimuli.[28] When large stimuli are used, the alternations are often piecemeal, while for small stimuli the whole image alternates in unison. It has been shown that it is not necessary to have images that the subject can interpret: random dot stereograms in which high spatial frequencies have been removed from one image and low spatial frequencies from the other are no longer fusible and produce binocular rivalry[29] Reviews and a summary of work on binocular rivalry can be found in Refs. 30–32

Investigators have sought techniques in which fusion is likely to occur. One set of approaches focuses on maintaining chromatic content and reducing the resolution in one of the images, relying on the fusion process for reconstructing the image in the original high resolution. Dinstein et al.[33] take advantage of these capabilities of human visual perception. They use a stereo image pair in which one of the images is subsampled and compressed, yet the human observer produces from those two images an image that is perceived stereoscopically as a sharp 3-D image.

The underlying hypothesis for our investigation calls for presenting two images that differ in their *chromatic* content to the two eyes. Previous work has examined rivalry between competing colors, not between color and lack of color. Andrews and Lotto report that when two different-colored panels are presented to a human viewer, the perception will often alternate between the colors.[34] Alternatively, Julesz[35] reported that two images of different colors can create the impression of a third, mixed color. This percept is known as binocular color mixture.[36] Makous and Pulos[37] reported that during binocular rivalry between a red-and-black grating and a perpendicular green-and-black grating, the colors mixed, so that the observer reported seeing alternation between perpendicular yellow and black gratings (not even a hint of a plaid). In other words, there is some form of integration of the colors seen by the two eyes. To the best of our knowledge, there is no research on rivalry between a color and a monochrome image such as presented here.

Our hypothesis would extend the preceding results by reducing the amount of information coming from 3-D color space by reducing one of the images to a monochrome image, thus minimizing the amount of data transmitted before any compression method is applied. Moreover, this approach requires only one of the cameras to be a color camera.

We were encouraged in exploring this direction (given the question of whether fusion or rivalry will occur in our case) by Andrews and Lotto,[34] who have shown that when physically different monocular stimuli are likely to represent the same object at the same location in space, fusion is likely to result. Moreover, there is evidence that even in the case of rivalry, information from the suppressed eye is still processed in the brain, i.e., the eye's sensitivity is reduced but not shut off. Abruptly increasing the luminance of or moving an object seen by the suppressed eye will cause it to be detected.[38,39]

To summarize, the literature makes several important points, which we build on in our approach: (i) if the disparate-appearing objects are likely to be the same object at the same location in space, fusion may occur; (ii) integration of the colors between the two eyes may occur; and (iii) information from the nonobserving eye is still getting through the visual processing areas, possibly enabling a 3-D percept. By using one achromatic image (gray) instead of a competing color, we anticipate reconstructing an image with depth. If the image from one eye were fully suppressed, there could be no fusion and no depth perception.

3 Methods

Controlled psychophysical experiments were conducted in order to validate the hypothesis that a stereo pair consisting of one monochrome image and one color image produce a 3-D color perception. This section presents the experiments that were conducted in order to confirm this fundamental principle of the proposed technique.

Various combinations of color and monochromatic images were presented to the subjects, using either a stereoscope (Fig. 1) or a HMD (Fig. 2). The first experiment made use of a 1905 stereoscope and two sets of images, selected to have different color contents. One set, dominated by yellow colors, included only three objects, positioned at various distances from the camera, against a neutral homogeneous background. The second set of images contained color-saturated objects presented against a cluttered background of colorful content, and were mainly green and magenta in color.

The second set of experiments tested the hypothesis with a HMD. The importance of testing with an HMD is that recent technological developments allow the addition of 3-D realism through sophisticated new display techniques. These technologies introduce elements such as refresh rates that might cause the visual system to perform differently than under conditions of a stereoscope. One popular option is a sophisticated HMD where each eye has its own miniscreen. In light of that, the second experiment, whose intent was to establish the phenomenon under more application-like conditions, made use of an nVisor-SX HMD controlled by a PC. This setting enabled a controlled procedure of
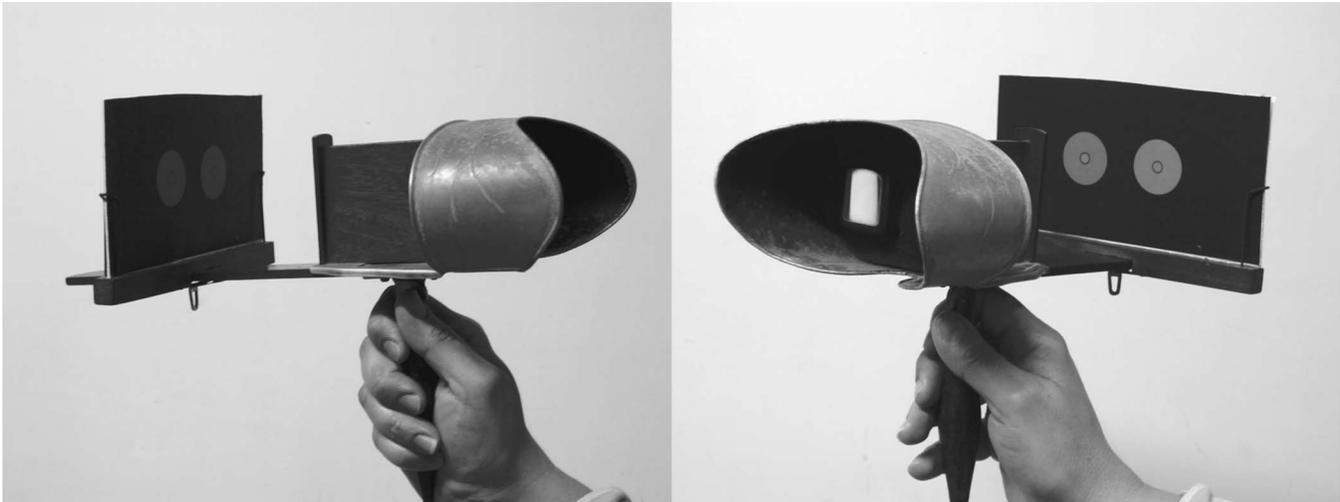
Fig. 1 The 1905 stereoscope used in the experiments.

automatic display of images in random order, with multiple repetitions. The images displayed in this experiment contained objects of saturated colors (red, yellow, green, blue).

3.1 Experiment 1

The 1905 stereoscope used in the experiments (shown in Fig. 1) is designed to display two photographs to the viewer's eyes, using a partition that enables each eye to view only one of the two pictures. The two pictures, taken from slightly different viewpoints, are positioned on a stiff mounting, side by side, and viewed through a set of prisms and lenses. Since the invention of the stereoscope by Wheatstone in 1838, it has been utilized in many experiments involving fusion, color mixing, and rivalry.[40]

These experiments were performed on 11 subjects (9 males and 2 females) between the ages of 16 and 45. Except for two subjects, all were naive as to the purpose of the experiments. All subjects were clinically tested for normal color and depth perception and equal visual acuity in both eyes, as determined by optometric examinations. Stereoacuity was tested via the Titmus stereo fly test. Color vision was tested using Ishiara plates.

Fig. 2 The nVisor-SX head-mounted display used in the experiments.

Two sets of stereo image pairs were used. Each set contained different objects at different distances and colors, with the two static images taken from slightly different viewpoints in order to generate depth perception. One set of stereo images included objects that were mostly yellow, with low average saturation (Fig. 3). The other pair consisted of objects in saturated colors, such as green and magenta. Some of the images were taken by us (Fig. 3), and some (Fig. 4) were modified from Refs. 41 and 42.

The two static images presented to the observers were at times different in their color content, to evaluate our hypothesis. The original color stereo images were modified to fit the experiment (e.g., cropped and partially or fully converted to grayscale). In each set, four different image pairs were used, creating four color combinations each:

(a) Left eye: color image; right eye: color image (color-color)
(b) Left eye: grayscale image; right eye: grayscale image (gray-gray)
(c) Left eye: color image; right eye: grayscale image (MIX)
(d) Left eye: grayscale image; right eye: color image (MIX).

Observers were asked to adjust the distance between the displayed stereo pairs and their eyes until the pair of images fused into one 3-D image, if possible. The image pairs were presented in random order; each set was presented twice. For each image pair, subjects were asked to rate their color perception on a scale of 1 to 10, with 1 denoting a grayscale image and 10 an image in vivid colors. They were also asked to rate the depth perception from 1 to 10, with 1 corresponding to a flat image with no depth at all, and 10 to a full 3-D image. Subjects were unaware of the color content of the images they were viewing.

3.2 Experiment 2

In the second experiment the stereoscope was replaced by a HMD. This nVisor-SX display is usually used for advanced virtual reality applications. It incorporates high-resolution
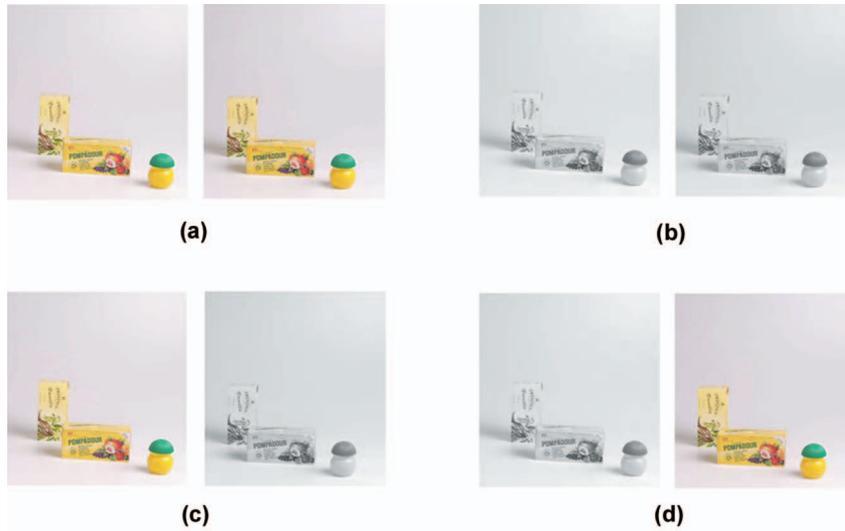
Fig. 3 Four combinations of low-saturation stereo pairs, used in the experiments: (a) color-color; (b) gray-gray; (c) MIX (color on left); (d) MIX (color on right).
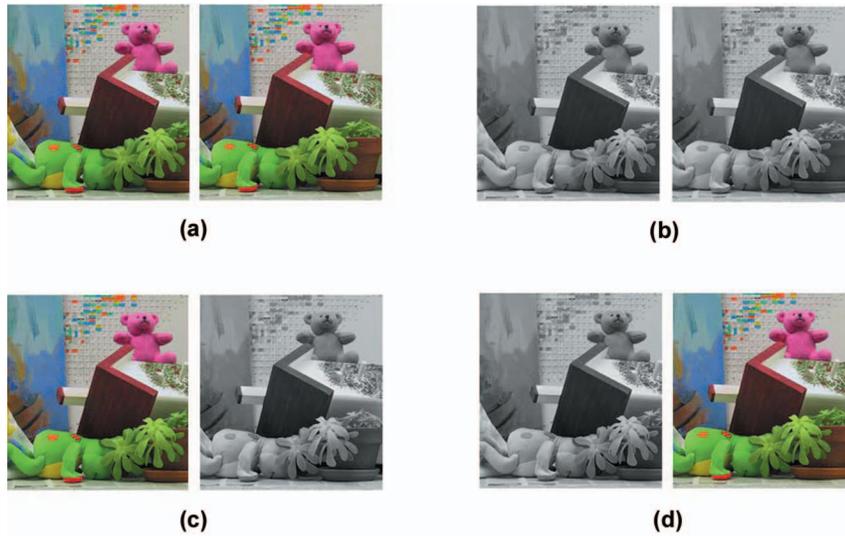

Fig. 4 Four combinations of high-saturation stereo pairs used in the experiments: (a) color-color; (b) gray-gray; (c) MIX (color on left); (d) MIX (color on right).
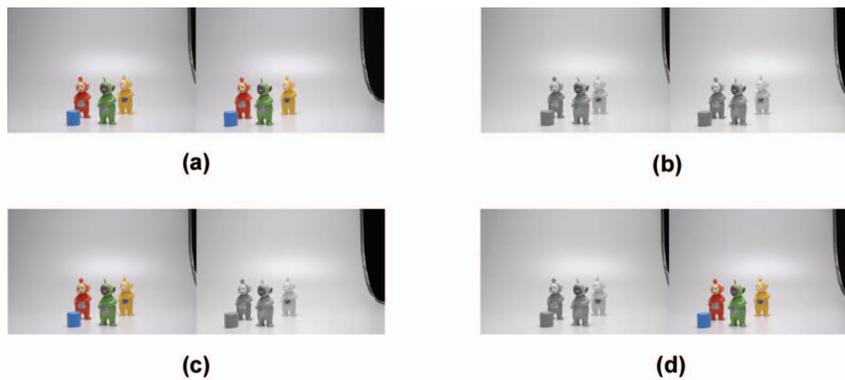

Fig. 5 First HMD set: four combinations of stereo pairs used in the HMD experiments: (a) color-color; (b) gray-gray; (c) MIX (color on left); (d) MIX (color on right).
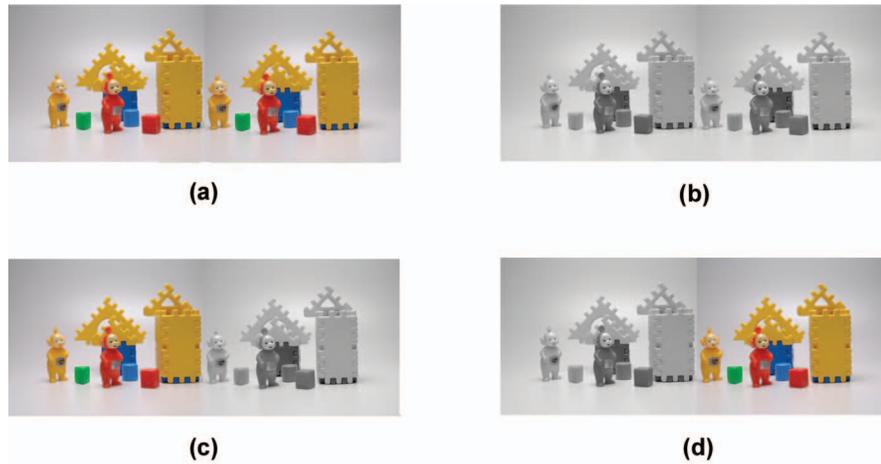
Fig. 6 Second HMD set: Four combinations of stereo pairs used in the HMD experiments: (a) color-color; (b) gray-gray; (c) MIX (color on left); (d) MIX (color on right).

color microdisplays with custom-engineered optics designed to deliver high visual acuity in a wide FOV. The nVisor-SX has access to user adjustments, including interocular distance, and an eye relief to accommodate users with eyeglasses. The HMD specifications are: dual channel support (stereo), monocular FOV (diagonal) 60° overlap 100%.

These experiments were performed on 15 subjects (10 males and 5 females) between the ages of 21 and 43. All subjects were clinically tested for normal color and depth perception and equal visual acuity in both eyes, as determined by optometric examinations (using standard noninvasive optical tests). All subjects were naive as to the purpose of the experiments.

Three sets of stereo image pairs were used. Each set contained different objects at different distances and of dif-

ferent colors. Within each pair the two images were taken from slightly different viewpoints in order to generate depth perception. One set of stereo images (Fig. 5) included four colored objects (red, green, yellow, and blue) at different distances positioned in front of a homogeneous monochromatic background (uncluttered). In the second set of stereo images (Fig. 6), the same objects were used, with more objects added in the back, but keeping the background homogeneous and monochromatic. The third set of the images (Fig. 7), contained two greenish objects at two different distances from the camera. Overall, the three image sets covered a range of color contents, in terms of hues and saturation. The variety was intended to cover a range of colors, distances, interpositions, and saturations. Each image set comprised four chromatic combinations of the basic

Fig. 7 Third HMD set: Four combinations of stereo pairs used in the HMD experiments: (a) color-color; (b) gray-gray; (c) MIX (color on left); (d) MIX (color on right).

Table 1 Depth perception results (stereoscope).

	Subjective depth perception		
	Color pair	MIX pair	Monochromatic pair
Mean	9.62	9.35	9.20
Std. dev.	0.92	1.38	0.96

Table 2 Color perception results (stereoscope).

	Subjective color perception		
	Color pair	MIX pair	Monochromatic pair
Mean	9.56	7.37	1
Std. dev.	0.92	1.55	0

stereo pair: (a) color-color, (b) gray-gray, (c) MIX (color on left), and (d) MIX (color on right), as presented in Figs. 5–7.

4 Results

The results show that as a rule, subjects successfully fused the images in all sets. Irrespective of the color differences in a given image pair, the subjects perceived a strong 3-D image in depth. Color perception was successful in all mixed cases, though some degradation of color quality can be seen in the results. We provide below detailed results of the depth and color perception grades for each of the two experiments.

4.1 Experiment 1

4.1.1 Depth perception

Depth was perceived in all four combinations of stereo images and in both image sets that were presented to the subjects. Color rivalry, if it occurred, seemed to have no influence on depth perception. The results across the 11 subjects are summarized in Table 1, and presented graphically in Fig. 8. Figure 8(a) shows the results averaged over all subjects. Figure 8(b) shows the average for each subject individually.

In Table 1 the row marked "Mean" includes the mean results (across 11 subjects, multiple viewings per subject, both sets of images). The row marked "Std. dev." includes

the standard deviation for these mean values. Columns refer to the content of the presented stereo pairs. The column labeled "MIX pair" includes both the color-gray and gray-color presentations.

Figure 8(a) shows the same results graphically. The leftmost bar indicates the mean depth perception for the color-color images; the middle bar, for the mixed images; and the rightmost bar, for the gray-gray images. The offset indicators on each bar show the extent of the standard deviation.

Figure 8(b) shows the individual average results for each subject. Each of the 11 triplets of bars corresponds to one subject. The order of the bars within each triplet is the same as in Fig. 8(a). The Y axis denotes the score on the scale of 1 to 10 described earlier. The X axis separates the different subjects. Here, for each subject, the result is their average across both image sets (Figs. 3 and 4).

In summary, depth perception seems unaffected by the use of mixed pairs. Average depth scores for all three presentations (color-color, MIX, and gray-gray) show insignificant differences (the maximum difference obtained between color pair and monochromatic pair has a paired, two-tailed t-test value $p=0.21$).

4.1.2 Color perception

All subjects reported that they perceived a color image when they were presented with one color image and one monochromatic image (MIX pair), as well as when viewing
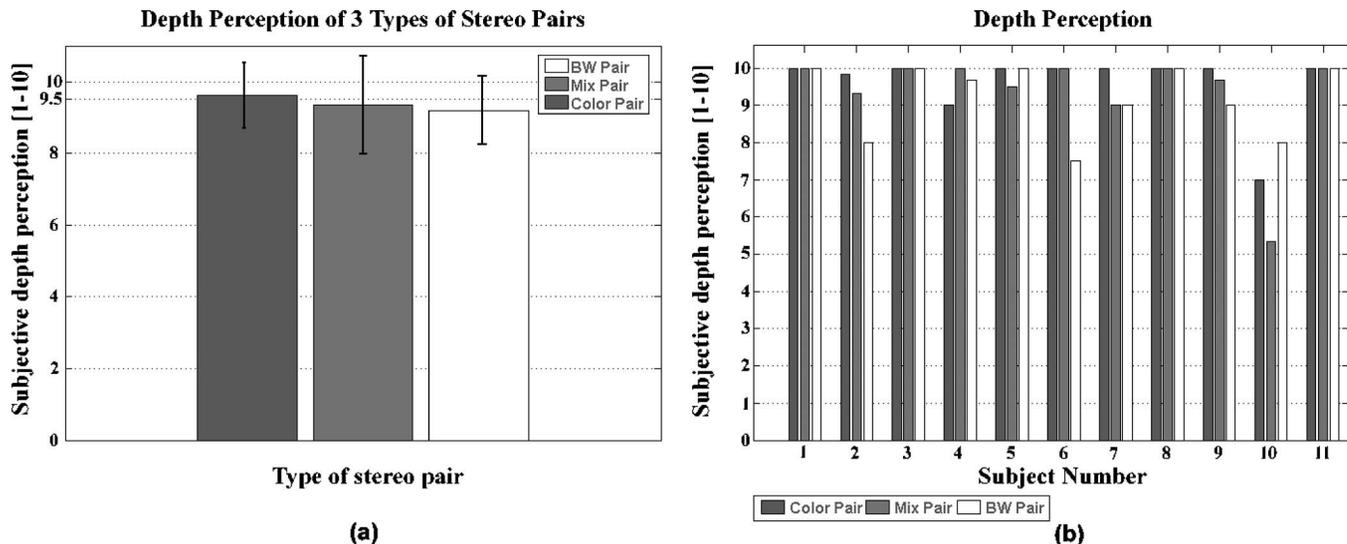

Fig. 8 (a) Depth perception of three types of stereo pairs (stereoscope). (b) Subjective depth perception (stereoscope).

Table 3 Color difference distance results (stereoscope).

	Difference	
	Color–MIX	MIX–monochromatic
Mean	2.2	6.37
Std. dev.	1.33	1.55

Table 4 Depth perception results (HMD).

	Subjective depth perception		
	Color pair	MIX pair	Monochromatic pair
Mean	10.00	10.00	10.00
Std. dev.	0.0	0.0	0.0

two color images. The mean and standard deviation results are shown in Table 2. The average color score is 9.56 (std. dev. 0.92) in the color-color pairs, and 7.37 (std. dev. 1.55) in the mixed pairs. Indeed, some subjects reported a color fading effect when viewing some images. Also, some subjects reported that the color of some of the objects within an image faded in and out during viewing. These reports were not consistent for the same subject across different viewings, or for the same image pair across subjects.

The results presented in Table 2 are plotted in Fig. 9(a). The average results for each of the 11 subjects are shown in Fig. 9(b). The axes are the same as in Fig. 8.

The results show that subjects clearly perceived color in the mixed presentations. There was, however, a certain amount of degradation in the quality of the perceived color. In order to ascertain whether the degraded color perception in the mixed condition more closely resembled the color pair or the monochrome pair, average color perception differences were calculated in pairs (i.e., for each subject separately) and for the group as a whole. These are presented in Table 3. The table shows that the difference in results between the color-color images and the mixed images is smaller than the difference between the mixed images and the monochromatic (gray-gray) images. This difference is statistically significant (paired one-tailed t-test, $p = 0.0005$). Given that the measures are most likely on an ordinal scale rather than an interval one, from a purist ob-

jective perspective we cannot completely reject the null hypothesis that the results from the mixed presentations are not closer to the color presentations than to the monochrome presentation. However, from a subjective impression we think that the comparison indeed holds and that the scale may be interval.

4.2 Experiment 2

4.2.1 Depth perception

Depth was perceived in all four combinations of stereo images and in all three image sets that were presented to the subjects. The results here were even more impressive than in the stereoscope experiments, as can be seen in Table 4 and graphically in Fig. 10. All subjects in all viewings rated the depth as a 10. Figure 10(a) shows the average results for all subjects. Figure 10(b), shows the average for each individual subject. Table 4 is arranged the same as Table 1 and Fig. 10 parallels Fig. 8. In summary, when using an HMD, depth perception is unaffected by the use of mixed pairs of images.

4.2.2 Color perception

With the HMD, all subjects reported that they perceived a color image when they were presented with one color image and one monochromatic image, as well as when viewing two color images. The mean and standard deviation are
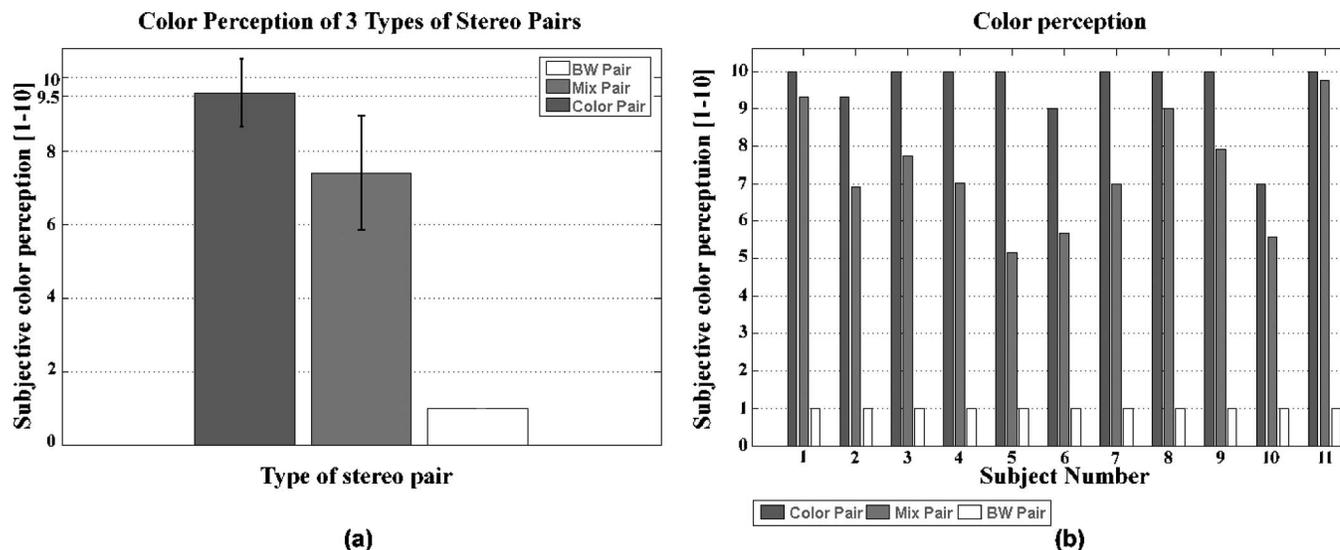

Fig. 9 (a) Color perception of three types of stereo pairs (stereoscope). (b) Subjective color perception (stereoscope).

Table 5 Color perception results (HMD).

	Subjective color perception		
	Color pair	MIX pair	Monochromatic pair
Mean	9.95	7.20	1.01
Std. dev.	0.18	0.92	0.05

Table 6 Color difference distance results (HMD).

	Difference	
	Color−MIX	MIX−monochromatic
Mean	2.7	6.18
Std. dev.	0.83	0.9

shown in Table 5. The average color score is 9.95 (std. dev. 0.18) in the color-color pairs, and 7.20 (std. dev. 0.92) in the mixed pairs. Here too, some subjects reported a color fading effect when viewing some mixed pairs or reported that the color of some of the objects within an image faded in and out during viewing. These reports were not consistent for the same subject across different viewings, or for the same image pair across subjects. The averages of the color perception results for all subjects are presented in Table 5 and Fig. 11(a). The average results for each of the 15 subjects are shown in Fig. 11(b). The axes are the same as in the respective depth perception graphs [Fig. 9(a) and 9(b)].

As with the stereoscope, the results clearly show that in a mixed presentation the subjects perceived a 3-D color image, but with color that was of a slightly degraded quality. Here too, average color perception differences were calculated in pairs (i.e., for each subject separately) and for the group as a whole. The group results are presented in Table 6. The table shows that the difference in results between the color-color images and the mixed images is smaller than the difference between the mixed images and the monochromatic (gray-gray) images. This difference is statistically highly significant (paired one-tailed t-test, $p=1e-17$).

5 Discussion

The approach presented and tested in this paper is innovative, and may prove useful for imaging systems combining color and depth. It allows for the use of two cameras, only one of which is color. This technique utilizes the human visual system's innate abilities and the manner in which the image is perceived in the brain to reduce the redundant color information. The experimental results confirm our hypothesis that one monochrome and one color image are sufficient for the visual system to generate full color and depth perception.

By using one monochromatic image instead of a competing color, our subjects perceived an image with depth, and one in which the "colors" blend to produce a color 3-D image. No degradation in depth perception was measured. However, some degradation in the perception of color is evident, though the subjects still rank their color perception of mixed images significantly closer to the full-color images (color-color) than to the monochromatic (gray-gray) images.

The proposed approach for reducing bandwidth and using only one color camera appears to be a viable complementary approach to current techniques used in transmitting color 3-D images.
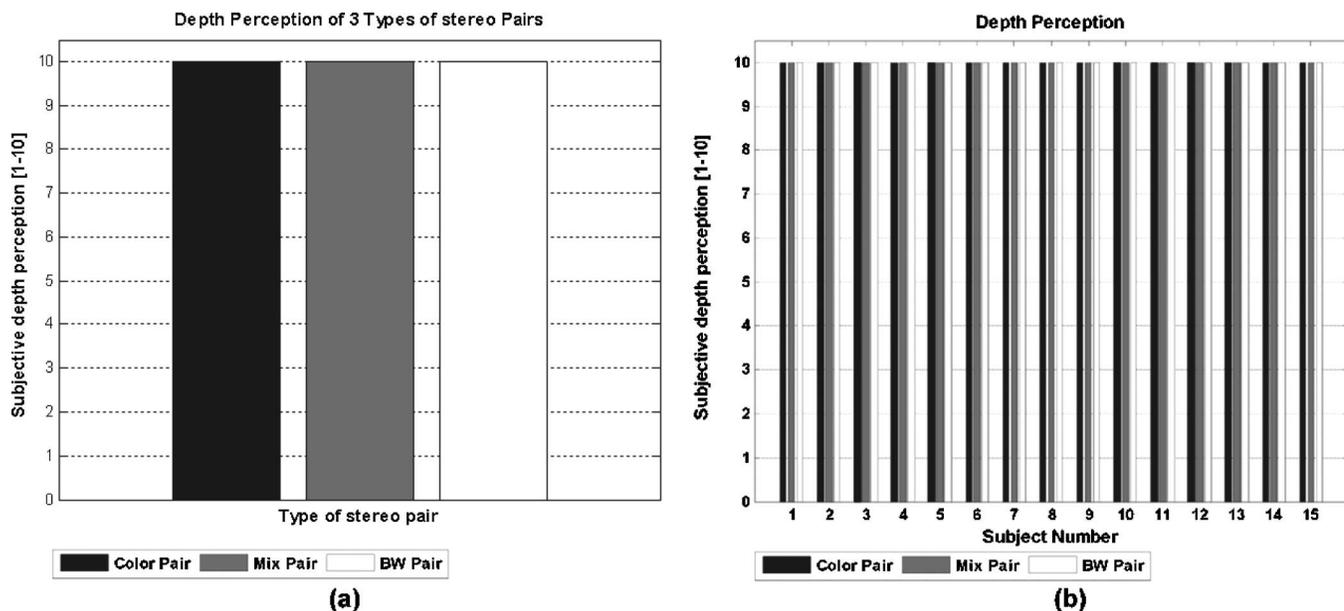

Fig. 10 (a) Depth perception of three types of stereo pairs (HMD). (b) Subjective depth perception (HMD).
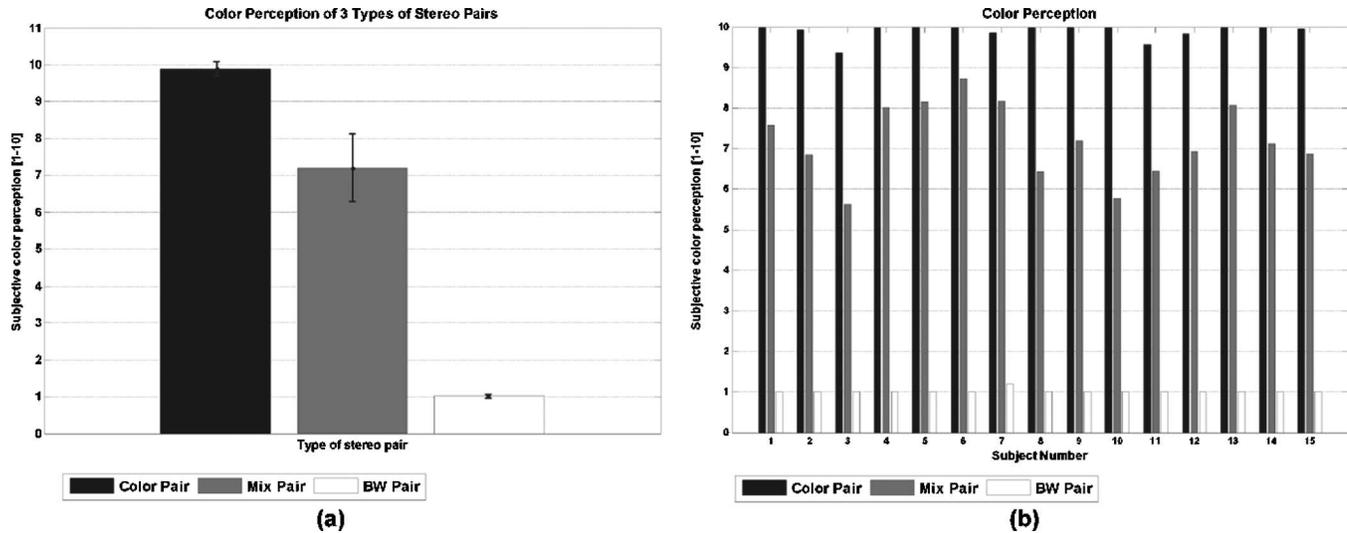
Fig. 11 (a) Color perception of three types of stereo pairs (HMD). (b) Subjective color perception (HMD).

There are several potential shortcomings to using this approach. The most obvious, based on the presented data, is that although the hypothesis that a mixed pair will be seen in color has been validated, it is clear that the color in that scene is less vivid than in a color pair. Furthermore, as can be seen from Fig. 9(a), there is a large variation in color perception between subjects. Thus, if a system were designed based on this approach and were intended for a broad audience, it might be necessary to first ascertain that indeed the intended user is among that section of the population that perceives the color more vividly rather than less. Investigation of the color degradation, variation between subjects, and means for improving the color perception score will be explored in future work.

6 Conclusions

We continue to further explore this phenomenon with the goal of defining the minimal requirements for which a mixed stereoscopic image contains enough information for the human visual system to perceive both color and depth. The results from the experiments described in this paper validate the basic principle. Future experiments will be made to further quantify the efficacy of this approach and its robustness. In addition, attempts will be made to use this technique in actual applications, such as teleoperation of robots.

References

1. R. Skuler and R. Blake, *Perception*, 3rd ed., McGraw-Hill (1994).
2. P. H. Batavia and S. Singh, "Obstacle detection using adaptive color segmentation and color stereo homography," in *Proc. IEEE Int. Conf. on Robotics and Automation* (2001).
3. Y. T. M. W. Siegal and T. Akiya, "Kinder gentler stereo," in *Stereoscopic Displays and Virtual Reality Systems VI*, *Proc. SPIE* **3639A**, 18–27 (Jan. 1999).
4. A. J. S. Sethuraman and M. Siegel, "Multiresolution based hierarchical disparity estimation for stereo image pair compression," in *Proc. Symp. on Application of Subbands and Wavelets*, A. Akansu, Ed., IEEE (1994).
5. M. S. S. Sethuraman and A. Jordan, "A multiresolutional region based segmentation scheme for stereoscopic image compression," in *Digital Video Compression: Algorithms and Technologies*, *Proc. SPIE* **2419**, 265–274 (1995).
6. M. W. S. S. Sethuraman and A. Jordan, "A multiresolution framework for stereoscopic image sequence compression," in *Proc. ICIP-94*, Vol. **II**, pp. 361–365, IEEE (1994).
7. W. Woo and A. Ortega, "Stereo image compression based on disparity field segmentation," in *VCIP '97*, *Proc. SPIE* **3024**, 391–402 (Feb. 1997).
8. W. Woo and A. Ortega, "Dependent quantization for stereo image coding," in *PW-EI-VCIP'98*, *Proc. SPIE* **3309**, 902–913 (1998).
9. W. Woo and A. Ortega, "Overlapped block disparity compensation with adaptive windows for stereo image coding," *IEEE Trans. Circuits Syst. Video Technol.* **10**(2), 194–200 (2000).
10. J. Rehg, M. Loughlin, and K. Waters, Vision for a smart kiosk, in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 690–696 (1997).
11. C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**(7), 780–785 (1997).
12. T. Darrell, G. Gordon, M. Harville, and J. Woodfill, "Integrated person tracking using stereo, color, and pattern detection," in *Proc. Conf. on Computer Vision and Pattern Recognition*, pp. 601–609, IEEE Computer Soc. (1998).
13. D. Drascic and J. J. Grodski, "Defence teleoperation and stereoscopic video," in *Stereoscopic Displays and Applications IV*, *Proc. SPIE* **1915**, 58–69 (Feb. 1993).
14. JPEG Committee, http://www.jpeg.org/committee.html.
15. L. P. Yaroslavsky, "On redundancy of stereoscopic pictures," in *Image Science '85*, *Proc., Acta Polytechnica Scandinavica*, Vol. **1**, pp. 82–85.
16. M. G. Perkins, "Data compression of stereo pairs," *IEEE Trans. Commun.* **40**(4), 684–696 (1992).
17. M. E. Lukacs, "Predictive coding of multi-viewpoint image sets," in *Proc. Int. Conf. on Acoustics Speech and Signal Processing*, pp. 521–524, IEEE (1986).
18. W. Woo and A. Ortega, "Optimal blockwise dependent quantization for stereo image coding," *IEEE Trans. Circuits Syst. Video Technol.* **9**(6), 861–867 (1999).

19. L. Chiariglione, "Short MPEG-1 description," http://www.chiariglione.org/mpeg/standards/mpeg-1/mpeg-1.htm (1996).
20. R. Koenen, "Overview of the MPEG-4 standard," http://www.chiariglione.org/mpeg/standards/mpeg-4/mpeg-4.htm (2002).
21. M. Siegel, P. Gunatilake, S. Sethuraman, and A. Jordan, "Compression of stereo image pairs and streams," in *Stereoscopic Displays and Virtual Reality Systems*, *Proc. SPIE* **2177**, 258–268 (Feb. 1994).
22. K. T. Kim, M. Siegal, and J. Son, "Synthesis of a high-resolution 3d-stereoscopic image pair from a high-resolution monoscopic image and low-resolution depth map," in *Stereoscopic Displays and Virtual Reality Systems*, *Proc. SPIE* **3295**, 76–86 (1998).
23. W. D. Reynolds and R. V. Kenyon, "The wavelet transform and the suppression theory of binocular vision for stereo image compression," in *IEEE Int. Conf. on Image Processing*, Vol. **2**, pp. 557–560 (1996).
24. D. R. Simmons and F. A. Kingdom, "Interactions between chromatic- and luminance-contrast-sensitive stereopsis mechanisms," *Vision Res.* **42**(12), 1535–1545 (2002).
25. S. K. Shevell and P. R. Miller, "Color perception with test and adapting lights perceived in different depth planes," *Vision Res.* **36**(7), 949–954 (1996).
26. D. R. Simmons and F. A. Kingdom, "On the binocular summation of chromatic contrast," *Vision Res.* **38**(8), 1063–1071 (1998).
27. A. Anzai, J. M. A. Bearse, R. D. Freeman, and D. Cai, "Contrast coding by cells in the cat's striate cortex: monocular vs. binocular detection," *Visual Neurosci.* **12**(1), 77–93 (1995).
28. D. Alais, R. P. O'Shea, C. Mesana-Alais, and I. G. Wilson, "On binocular alternation—a translation of Diaz-Caneja (1928)," Dept. of Psychology, Univ. of Otago, Dunedin, New Zealand (2000). Retrieved December 6, 2006 from http://psy.otago.ac.nz/r_oshea/BR_bibliography/br_DJtrans.html.
29. J. E. Mayhew and J. P. Frisby, "Rivalrous texture stereograms," *Nature (London)* **264**(5581), 53–56 (1976).
30. R. Blake, "Primer on binocular rivalry, including controversial issues," *Brain Mind* **2**, 5–38 (2001).
31. R. Fox, *Binocular Vision and Psychophysics*, pp. 93–110, Macmillan Press, London (1991).
32. I. P. Howard and B. J. Rogers, *Binocular Vision and Stereopsis*, Oxford Univ. Press, New York (1995).
33. I. Dinstein, M. G. Kim, A. Henik, and J. Tzelgov, "Compression of stereo images using subsampling and transform coding," *J. Soc. Photo-Opt. Instrum. Eng.* **30**(9), 1359–1364 (1991).
34. T. J. Andrews and R. B. Lotto, "Fusion and rivalry are dependent on the perceptual meaning of visual stimuli," *Ecologist* **14**(5), 418–423 (2004).
35. B. Julesz, *Foundation of Cyclopean Perception*, Univ. of Chicago Press, Chicago (1971).
36. C. Erkelens and R. van Ee, "Multi-coloured stereograms unveil two binocular colour mechanisms in human vision," *Vision Res.* **42**(9), 1103–1112 (2002).
37. W. Makous and E. Pulos, "Grating colors mix while their contours rival," *Invest. Ophthalmol. Visual Sci.* **20**(Suppl.), 225 (1981).
38. R. Fox and R. Check, "Detection of motion during binocular suppression," *J. Exp. Psychol.* **78**(3), 388–395 (1968).
39. H. Wiesenfelder and R. Blake, "Apparent motion can survive binocular rivalry suppression," *Vision Res.* **31**, 1589–1599 (1991).
40. W. L. Gulick and R. B. Lawson, *Human Stereopsis. A Psychophysical Analysis*, Oxford Univ. Press, New York (1976).
41. D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.* **1–3**, 7–42 (Apr.-June 2002).
42. D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Proc. IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition* (2003).

Yael Termin received her BSc in physics and MSc in computer science from Bar-Ilan University and has recently submitted her PhD dissertation on perception of a 3-D colored stereo image from one colored and one grayscale images. During her PhD studies she has received the dean's award and a doctoral fellowship of excellence. She is beginning her postdoc in the the Neural Imaging Laboratory at the Gonda Multidisciplinary Brain Research Center at Bar Ilan University, Israel. This paper is based on her PhD thesis.

Gal A. Kaminka is a senior lecturer in the Computer Science Department at Bar Ilan University, Israel. His research expertise includes teamwork and coordination, robotics, behavior and plan recognition, multiagent systems, and modeling social behavior. He received his PhD from the University of Southern California, and spent two years as a post doctoral fellow at Carnegie Mellon University. Today, Dr. Kaminka leads the MAVERICK group at Bar Ilan, supervising close to 20 MSc and PhD students—the largest computer science group in Israel. He has been awarded an IBM faculty award and top places at international robotics competitions.

Sarit Semo received her BA in physics and BSc in electrical engineering with honors from the Technion Institute of Technology in Haifa, Israel. She is currently completing her MSc in biomedical engineering at Tel-Aviv University on color adaptation and color constancy in vision research.

Ari Z. Zivotofsky is a senior lecturer in the neuroscience program at Bar Ilan University, Israel. He did his undergraduate degree in electrical engineering at The Cooper Union and received his PhD from Case Western Reserve University in biomedical engineering. Following that, he spent four years as a postdoc in the Laboratory of Sensorimotor Research at the NIH. His research interests are in the areas of ocular motility, visual perception, and cognitive functioning.