# Optical Image Processing Using Light Modulation Displays

Gordon Wetzstein[1], Wolfgang Heidrich[1] and David Luebke[2]

1) The University of British Columbia    2) NVIDIA Research

**Abstract**

*We propose to enhance the capabilities of the human visual system by performing optical image processing directly on an observed scene. Unlike previous work which additively superimposes imagery on a scene, or completely replaces scene imagery with a manipulated version, we perform all manipulation through the use of a light modulation display to spatially filter incoming light. We demonstrate a number of perceptually-motivated algorithms including contrast enhancement and reduction, object highlighting for preattentive emphasis, color saturation, de-saturation, and de-metamerization, as well as visual enhancement for the color blind. A camera observing the scene guides the algorithms for on-the-fly processing, enabling dynamic application scenarios such as monocular scopes, eyeglasses, and windshields.*

Categories and Subject Descriptors (according to ACM CCS):  I.4.3 [IMAGE PROCESSING AND COMPUTER VISION]: Enhancement - Filtering

## 1. Introduction

The human visual system (HVS) is a remarkable optical device possessing tremendous resolving ability, dynamic range, and adaptivity. The HVS also performs an impressive amount of processing in early (preattentive) stages to identify salient features and objects. However, the HVS also has some properties that limit its performance under certain conditions. For example, veiling glare due to extremely high contrast can dangerously limit object detection in situations such as driving at night or driving into direct sunlight. On the other hand, conditions such as fog or haze can reduce contrast to a point that significantly limits visibility. The tristimulus nature of human color perception also limits our ability to resolve spectral distributions, so that quite different spectra may be perceived as the same color (*metamers*). Any form of color blindness exacerbates the problem.

We propose to enhance the power of the human visual system by applying on-the-fly optical image processing using a spatial light modulation display. To this end, we introduce the concept of see-through optical processing for image enhancement (SOPhIE) by means of a transparent display that modulates the color and intensity of a real-world observation. The modulation patterns are determined dynamically by processing a video stream from a camera observing the same scene. Our approach resembles and builds on work in computational photography and computer vision (see Section 2), but we target a human observer rather than a camera
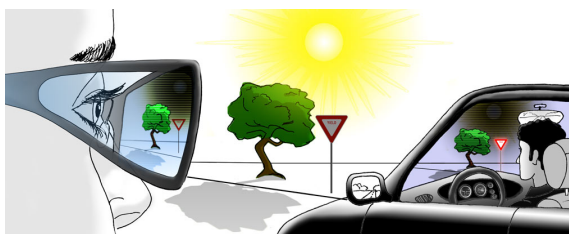


**Figure 1:** *A conceptual illustration of our approach. A light modulation display locally filters a real-world scene to enhance the visual performance of a human observer, in this case by reducing the contrast of the sun and boosting the saturation of the traffic sign for a driver. Our approach would apply to see-through scenarios such as car windshields and eye glasses as depicted, as well as to binoculars, visors, and similar devices.*

sensor, and our goals are thus perceptual rather than photographic. Our work also resembles traditional 'see-through' augmented reality (AR) displays, which present the observer with an image that shows synthetic imagery overlaid (i.e. added to) the real world scene. Unlike optical AR displays, we use the display to spatially *filter* the incoming light at the observer's position, allowing us to perform image processing operations such as contrast reduction, contrast enhancement, or color manipulation without the latency introduced by video-based AR systems.

Potential applications for such an approach, once perfected,

range from tone-mapping sunglasses and ski goggles to 'smart' automobile windshields that reduce dangerous contrast, enhance low contrast, or subtly draw attention to important features such as street signs or an erratic vehicle. We take the first steps toward such ambitious applications by demonstrating several examples of optical processing in a prototype setup consisting of a monocular scope-like device (see Section 3.1). Our primary prototype enables the display and the observed scene to be in focus, and enables the camera and observer to share an optical axis. We also envision and analyze different scenarios such as the aforementioned glasses or car windshields that place the display out of focus and impose a parallax between the camera and observer.

The goal of this paper is to introduce a novel display technology and to show its potential using a variety of intuitive applications. We make several specific contributions:

- We show several perceptually-motivated applications of optical contrast manipulation, including the application of gamma curves, tone mapping-style contrast reduction, and contrast enhancement using counter-shading.

- We show the use of spatially modulated color filters to perform de-metamerization and color saturation enhancement or reduction, as well as providing visual aids for the color blind.

- We discuss object highlighting using manipulations of color, intensity, or contrast.

- We analyze and demonstrate both in- and out-of-focus geometries for applications such as the above.

- We analyze and demonstrate both in-line and parallax geometries for the camera and observer.

- We evaluate the effect of our contrast reduction approach on the perception of low-contrast details in high-contrast scenes with a user study.

## 2. Related Work

**Optical Image Processing.** A vast literature exists on optical image processing methods using Fourier optics [BW97], including effects such as edge enhancement [YWK*06] and image sharpening [SSK01]. Unfortunately, Fourier optics require coherent light and are thus ill-suited for natural environments.

**Night Vision.** Image amplification for low-light vision enhancement has also been well studied, including wearable solutions e.g. for military personnel [Hra02]. Night-vision solutions using optical image amplification approaches typically employ cascading opto-electrical effects. Like SOPhIE, these approaches have a multiplicative effect on the incoming imagery, but they perform uniform amplification rather than spatially selective filtering.

**Programmable Imaging.** Our work echoes the programmable imaging concept introduced by Nayar et al. [NBB04, NBB06]. Their work introduces a reflective dis-

play – a digital micromirror array (DMD) – into the optical path of a camera in such a way that it is in focus and modulates the pictured image. The modulation is dynamically determined in a feedback loop, with the camera observing the already modulated light reflected off the DMD. The authors demonstrate applications to high dynamic range imaging and computer vision algorithms such as feature detection and object recognition. Our 'window' prototype (Section 3.2) also resembles earlier work by Nayar and Branzoi [NB03], who used a liquid crystal display (LCD) as an out-of-focus spatial light modulator for adaptive dynamic range imaging, selectively attenuating incoming light to resolve dark regions without saturating the image sensor in bright regions.

Whereas programmable imaging manipulates the radiometric characteristics of the incoming light to aid photography and computer vision, our manipulations are intended for direct viewing. Our target is the human visual system rather than a photographic or computational camera. This opens up a wide range of perceptually-motivated applications as outlined above, and also impacts the algorithms we use. For example, adaptive dynamic range imaging seeks to bring the image sensor close to saturation everywhere for maximal sensitivity. By comparison, a contrast-reducing application in our system more closely resembles tone mapping: it must produce a single image that is immediately comprehensible to the human viewer and that best conveys the relative contrasts of the scene. Finally, rather than relying on a camera-display feedback loop, we use a camera image sensor with an unobstructed view of the object. This configuration also eliminates moiré patterns caused by the interaction of the pixel grids from the display and the image sensor.

**Augmented Reality.** Researchers in augmented reality (AR), also known as mediated reality, have been working on live manipulation of observed scenes for decades [Sut68, FMS93, ABB*01]. AR displays can be classified as *optical see-through* or *video see-through* [RHF94], with optical see-through AR displays further categorized as either head-worn [CR06] or spatial [BR05].

SOPhIE uses an optical design resembling optical see-through displays long used in AR, but differs crucially in the way the display content is combined with the real world. Traditional optical see-through AR uses combining optics, such as beam-splitters, to overlay synthetic content from a display, such as an OLED or LCD, onto real-world scenery in an *additive* fashion. In SOPhIE, however, we use a partially transparent display as a spatial light modulator (SLM) to filter, in real-time, the light arriving from the real world environment in a *multiplicative* fashion. Rather than using the display to show artificial content, we program the SLM transparency in order to perform a variety of image processing operations to aid the user in understanding the real-world environment. To our knowledge, SLMs in the optical paths of AR displays have so far only been employed to achieve mutual occlusion [KKO01, CHR04]. Since the optical fil-

tering acts immediately on the incident light, the processed real-world scene can be observed without latency in the SOPhIE approach. While filter updates are required in regular intervals to account for object or observer motion, we show that relatively infrequent updates are sufficient for many applications, due to the low spatial frequencies required for light modulation in most of our image processing tasks.

Video see-through AR captures the incoming light with a camera, processes the resulting image to merge it with the synthetic elements, and then displays the result. This design simplifies image processing and registration between synthetic and real imagery, and allows arbitrary manipulation of the real-world imagery, but at the cost of introducing the full system latency to the user's perception of the real world. Video see-through AR systems are also limited by the resolution and dynamic range of the display and camera. SOPhIE, just like traditional optical see-through AR, thus possesses a crucial advantage over video-based AR: the real world is viewed directly at the full resolution of the human eye, and without latency. As a result, SOPhIE avoids motion sickness, and with appropriate safeguards could be used for safety-critical applications.

On the other hand, the hardware components required in SOPhIE are very similar to the ones used in head-mounted AR applications. Recent advances in miniaturizing system components for head-worn displays could be applied directly to SOPhIE systems. It would thus be possible to seamlessly integrate our approach into everyday clothing, a concept thoroughly explored by Mann and colleagues [MFA*05, Man97] as well as accessories such as multipurpose contact lenses [LP08]. We have not attempted such miniaturization in our prototypes, but consider previous work a strong indication of its feasibility.

### 3. Prototype Designs

To demonstrate the feasibility of the SOPhIE approach, we have implemented two physical prototypes that let us experiment with different algorithms. The first one is a scope, and the second one a small see-through window-like configuration that can simulate, for example, one side of a pair of glasses. We do not deal with binocular stereo parallax that occurs when both eyes of the user look through the same display. For windshields or visors, this issue will eventually have to be addressed, but we believe this will be possible given HVS characteristics such as the dominant eye effect. We leave the investigation of such solutions for future work.

### 3.1. Monocular Scope Prototype

Our first prototype setup (Figure 2, left) is a scope-like system in which a lens assembly brings both the scene and a see-through LCD panel into focus simultaneously. A beam splitter ensures that the camera shares the optical axis of the system. Both the camera and the LCD panel are greyscale-only in this setup.
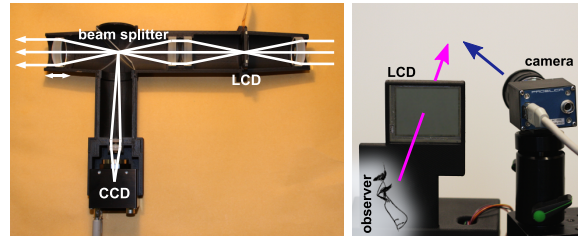
**Figure 2:** *Two prototypes. Left: the opened scope system with optical paths indicated. Right: the window-style setup is simply an LCD panel stripped of its backlight in a custom housing with the camera located next to it.*

**LCD Display.** The LCD panel was taken from a Sharp PG-D100U video projector, which uses three such panels, each with a resolution of 800x600, to project a color image. The control electronics were left in the projector housing, which was connected to the scope via a ribbon cable. Focusing optics for the display were assembled from four achromatic doublet lenses.

**Camera.** The camera, a 1.5 megapixel C-mount camera from Prosilica, observes the scene through a beam-splitter. We successfully used both a standard half-silvered mirror and a reflective polarizer for this purpose. The reflective polarizer has the advantage of minimizing light absorption through the LCD panel, as it pre-polarizes the light transmitted through it. The camera image is formed by two additional achromatic doublets that re-image the image projected by the front-most, i.e. object-side, lens of the assembly.

**Assembly.** We used a rapid prototyping machine to create a custom housing for the components of the scope. The white ABS plastic was subsequently spray-painted black to minimize scatter. The housing allows the front-most lens to be moved along the optical axis for focusing the scope at different depths.

**Calibration.** The response curves of both the camera and LCD panel are measured and compensated for using standard techniques [RWPD05]. Geometric calibration and alignment procedures are performed by replacing the human observer with a second camera, and manually aligning the captured images using a model of affine transformation plus radial distortion.

### 3.2. Window-style Prototype

Our second prototype allows a user to directly observe a scene through a color LCD panel without refocusing optics. The camera, a color version of the same Prosilica model used in the monocular scope, is located off-axis in this setup and the system corrects for the resulting parallax. This second prototype allows us to analyze the issues associated with target applications such as car windshields, sunglasses, or helmet visors where the relative positions of eyes and display are fixed.

Such setups can suffer from color fringing and other diffraction artifacts caused by small pixel structures. This problem can be avoided by using displays with large pixels. The resulting lower resolution can be tolerated for our application, since the display will generally be much closer to the human observer than the scene under investigation. This causes the display to be strongly blurred due to defocus, so that the display resolution is not a primary concern.

For our prototype, we chose a 2.5" active matrix TFT panel with a resolution of $480 \times 234$ pixels from Marshall Electronics (V-LCD2.5-P). In order to turn this display into a see-through spatial light modulator, we removed the backlight. Like most LCD displays, the Marshall panel is coated on the front surface with a film that acts as a polarizer, but also shapes the directional distribution of emitted light. Since this film prevents see-through applications by blurring the scene, we replaced it with a generic polarizer.

Geometric and radiometric calibration of this second setup proceed as for the first setup. A radial distortion term is not necessary due to the lack of refocusing lenses.

## 4. Applications

The SOPhIE framework enables a variety of applications. Some scenarios require certain optical setups, such as an in-focus display or a color display, while other applications also work with out-of-focus or black & white displays. We introduce several contrast manipulation methods with our scope prototype in Section 4.1. Section 4.2 includes a discussion on color manipulation techniques with our window-style prototype, then we show how optical object highlighting can be performed with both configurations in Section 4.3 and in Section 4.4 that most of the in-focus applications also work for out-of-focus displays.

### 4.1. Contrast Manipulations

**Direct View Gamma Curves.** A straightforward use of SOPhIE is the application of a non-linear 'response' function to the view. For example, a gamma curve with $\gamma < 1$ can be used to reduce contrast in environments with harsh lighting, while $\gamma > 1$ could boost contrast, for example on an overcast day. Gamma adjustments are frequently used to boost contrast on conventional displays. For example, the inverse gamma curve used by video standards deliberately differs from the gamma of physical output devices, in order to improve contrast [ITU90]. SOPhIE enables the same method for use in direct observations by human viewers.

To apply a response curve such as a gamma curve, we record a normalized image $I$ of the scene using the SOPhIE camera. The ratio image $I^\gamma/I$ is the desired transmission of the LCD panel. For contrast enhancement ($\gamma \geq 1$), all pixel values in the ratio image are between 0 and 1, and are thus directly realizable. For contrast reduction ($\gamma < 1$), however, the values are generally larger than 1. We address this issue by dividing the ratio image by the global constant $k = \max_x x^\gamma/x$,
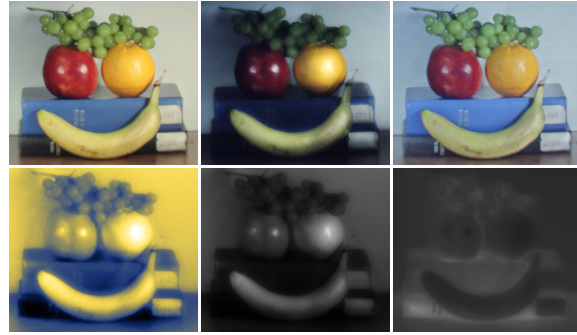


**Figure 3:** *Optical gamma modulation. The top row shows photographs taken from the point of view of a human observer using SOPhIE. Left: unmodified image. Center and right: a gamma of 2.5 and 0.625, respectively. Bottom left: false-color rendition of the image observed by SOPhIE's camera. Bottom center/right: modulation images shown on the SLM to produce the gamma curve applications above.*

where $x$ is every possible camera pixel intensity value. The constant only changes if $\gamma$ changes. This division reduces the peak brightness in the image but maintains the desired gamma curve characteristics. Figure 3 shows an example of this approach, photographed through the optical system of our scope, *as seen from the point of view of a human user*. The intensity adaptation of the HVS was simulated by adjusting the exposure times of the photographs.

**Contrast Reduction.** A gamma curve results in a relatively subtle compression or expansion of contrast. For more dramatic contrast reduction, we can apply strategies similar to tone mapping (see, for example [RWPD05]). One simple approach is to display the inverse of the image seen by the camera on the LCD panel, so that brighter regions in the camera image become darker (i.e. more heavily attenuating) regions on the see-through display. We choose a suitable scaling factor to effect the desired contrast reduction, and perform a spatial blur to avoid artifacts in the form of hard edges. Alternatively, we can leave most of the scene untouched, but dim the regions with very bright light sources or specular reflections. To determine such regions, the SOPhIE camera can be operated in an under-exposed setting, such that only very bright regions have non-zero pixel values. We again blur the resulting mask to avoid distracting artifacts.

Figure 4 shows examples of both approaches. Note how both methods darken the lamp enough to make out the light bulb, and remove lens glare from the photographs taken through the SOPhIE system. The effect is even stronger for a human observer, due to the stronger scattering in the human eye compared to camera lenses. This scattering causes *veiling glare*. In Section 5 we show with a user study that the proposed contrast reduction method can significantly improve detection rates under such glare conditions.

**Contrast Enhancement with Countershading.** In this section, we investigate an alternative strategy to contrast enhancement. Since we cannot amplify light with our setup,

**Figure 4:** *Contrast reduction. Three photographs taken from the point of view of a human observer. From left to right: no correction; correction with an inverted and blurred camera image; and darkening a blurred region around saturated pixels. The insets show the corresponding modulation patterns.*

we exploit the special characteristics of the HVS to achieve images that appear to have increased contrast even though the actual dynamic range is unchanged. Two examples of such methods are the familiar *unsharp-mask* filter, and the more sophisticated *countershading* approach by Krawczyk et al. [KMS07]. Both methods rely on the *Cornsweet effect* [KM88], in which the perceived difference in intensity between adjacent image regions can be amplified by exaggerating the difference in the boundary region only. The unsharp-mask filter provided by many image manipulation packages is an ad-hoc realization of this effect, while adaptive countershading as introduced by Krawczyk et al. [KMS07] is based on a principled analysis of the spatial frequencies in the image, and how the human visual system perceives them.

In the SOPhIE system, we experimented with both unsharp-masking and countershading. As seen in the intensity profiles in Figure 5, unsharp masking amplifies the contrast by modifying intensities in one of the higher frequency bands, while countershading alters multiple spatial frequency bands at the same time. The unsharp-mask implementation is similar to the method discussed before; we compute the ratio image of the processed and unprocessed camera image, and show the result on the LCD display. In order to ensure that the scanlines can be compared with the same exposure settings, the unmodified case was captured with the LCD transmission set to a constant value of 0.8.

Our countershading approach borrows from the work by Krawczyk et al. We process different spatial frequencies using an image pyramid [BA83]. For each frequency band, we compute the countershading profile, and add a multiple of it into the image. In the work by Krawczyk et al., the multiplier is set such that the contrast from a reference HDR image is reproduced unless doing so would result in visible halo effects. They use a simple perceptual model based on luminance and contrast masking, as well as contrast sensitivity, to determine the largest multiplier that would not result in such artifacts. Since we do not have reference images in our applications, we only use the perceptual model to determine the multiplier. This modified approach boosts the local contrast as much as possible without introducing banding.
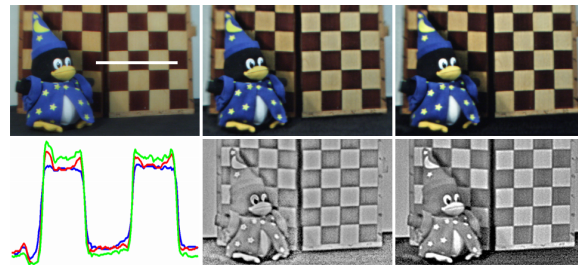


**Figure 5:** *Contrast enhancement. Top row: photographs taken from the viewpoint of a user with (from left to right) no contrast enhancement, unsharp-masking, and countershading. Bottom left: a cross-section of the intensity profile for the scanline marked in white (blue: original image, red: unsharp-masking, green: countershading). Bottom center/right: modulation images shown on the SOPhIE display.*

### 4.2. Color Manipulations.

The applications discussed thus far only require modulating intensity. Next, we discuss a number of applications that become possible with the addition of color displays and cameras.

**Manipulations of Color Saturation.** An intuitive application of a color system is the spatial manipulation of color saturation. For example, we can reduce the color saturation in the following way. For each pixel, the camera observes an RGB color. We set the transmission of the color channel with the smallest value to one, but reduce the transmission of the other two channels such that they match the smallest channel value. Likewise, we can boost saturation by choosing a transmission value of one for the dominant, i.e. largest, channel, but reducing the transmission for the other channels by a certain factor.

Note that this approach enables us to drastically reduce color saturation, but not, in general, to completely eliminate all color in the scene. The SOPhIE camera senses color with RGB sensor arrays and filters it with RGB displays. Like all tri-stimulus systems (including the HVS) such an approach permits metamers, i.e. different spectral distributions that result in an identical sensor response. The RGB filters in the SOPhIE system have different spectral distributions from the S, M, L photoreceptor types in the human eye, so differ-

ent metamers with a given sensor response would require slightly different RGB filter settings to completely remove all perceived color from the image.

Since our monocular scope prototype does not have a color display, we demonstrate the method with the second prototype, which has a parallax between camera and observer. Results are shown in Figure 9 (top row). They demonstrate that it is possible to drastically reduce color saturation with a colored SOPhIE system, although some residual color remains due to the effects discussed above.

**Color De-metamerization.** *De-metamerization* addresses the related problem of making metamers of a single perceived color visually distinct for a human observer. We can achieve this in a straightforward manner by using the color display as a spatially uniform, but programmable color filter. Figure 6 shows a result of this approach, again using the second prototype with a parallax between camera and observer. The exposure times of the right photograph was adjusted, simulating intensity adaptations performed by the HVS.
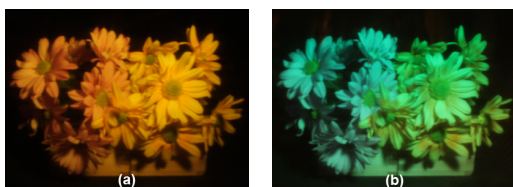


**Figure 6:** *De-metamerization: Colored flowers that appear similar under orange illumination (a). Modifying the color transmission of the LCD reveals the visual differences (b).*

**Color Deficiency.** A significant portion of the population suffers from some kind of color vision deficiency. The most widespread such deficiency is deuteranopia (red-green color blindness), which results from a lack of medium-wavelength sensitive cones. Brettel et al. [BVM97] showed how to simulate images as perceived by color-deficient viewers. For a given RGB image (Figure 7 (a)), the view of a deuteranopic dichromat is simulated in Figure 7 (b). Note how the colors of various objects in the scene become indistinguishable.

We can optically modulate the colors of a real-world scene so that visible differences are preserved for people suffering from color blindness. To this end, we employ Rasche's [RGW05] algorithm for re-colorization to calculate a desired color-modified image from a captured photograph. Dividing the former by the latter and accounting for the camera and display response curves enables us to compute a compensation image that is displayed on the SOPhIE LCD as seen in Figure 7 (d). A deuteranopic person would not be able to perceive the original colors, but could discriminate individual objects which would appear similar otherwise (Figure 7 (c)). Again, the exposure time of (c) was increased.

In recent work, Mohan et al. [MRT08] achieved a similar effect by illuminating a scene with spectrally controllable light



**Figure 7:** *Using SOPhIE for aiding color-deficient viewers. (a) shows the original scene and (b) the scene as perceived by a deuteranopic color-deficient viewer. Modulating the colors of the SOPhIE display as shown in (d) preserves visual differences for color blind persons (c). All images except (d) are photographed through the SOPhIE prototype assembly and images (b) and (c) are post-processed to simulate the perception of a deuteranopic viewer.*

sources to improve visible details for color-deficient viewers. However, it is often not feasible to control the illumination spectrum for a real-world scene. Due to its passive nature, the SOPhIE approach is more practical, because it could be integrated into glasses or other private devices of an individual user. As such, our approach can improve the vision of a color deficient viewer according to the type and severity of the deficiency, without affecting other observers.

### 4.3. Object Highlighting Using Preattentive Cues

The above examples apply the same image processing operation uniformly across the image. These methods can be made spatially varying by making use of any external object recognition or tracking method. For example, de-metamerization trivially extends to spatially selective de-metamerization, that is if one wished to de-metamerize camouflage from foliage but leave the color of the sky perceptually unmodified. More generally, spatially varying processing enables us to highlight objects of interest and to shift the user's attention. The effect ranges from subtle to quite dramatic.

For example, by increasing contrast in regions of interest we obtain an effect much like depth of field-driven attention in cinematography. On the other hand, reducing brightness or color saturation outside the regions of interest is a very strong way of directing attention (see Figure 8). Significant brightness [BPR83, HE99] and color [NS90, HE99] differences trigger preattentive processing mechanisms in the human visual system, so that in effect visual search tasks are performed in parallel rather than sequentially [BPR83, NS90]. As a result, the regions of interest stand out strongly even over very cluttered surroundings.

In our demonstration system, we highlight regions determined by either manual selection or simple background subtraction to identify regions of change. However, one could easily use dedicated tracking or recognition systems to highlight, for example, people, faces, or street signs. In multi-use collaborative systems, a user could use an interactive interface to draw the attention of other users to specific features.

### 4.4. Defocused Light Modulation

Conceptually, response function manipulations are operations that require a one-to-one correspondence between display pixels and scene points, and thus an in-focus display. However, in practice most algorithms also work very well with blurred ratio images, and thus out-of-focus displays. One exception is the countershading approach, which requires good alignment between display and world, and thus does not work with out-of-focus settings. All other algorithms presented here, however, can be used with window-like or near-eye setups without refocusing optics, as required for glasses, helmets, or windshields.

The ability to use defocused light modulation also makes the SOPhIE approach quite robust under misalignment between the see-through display and the real world. Consequently, we can handle setups in which a parallax exists between camera and observer, a situation unlike most results described so far, in which the camera shared the optical axis with the human observer. This feature is again important for systems such as glasses or windshields. Figure 9 shows several examples of out-of-focus photographs with a 5° parallax between viewer and camera.

Finally, and crucially, the robustness under misalignment produces a resulting robustness under motion. Even though our configuration has a system latency comparable to optical see-through augmented and mixed reality systems – or worse, since we have invested little effort in reducing latency – this does not translate into a noticeable misalignment if the modulation image is defocused or synthetically blurred. As a
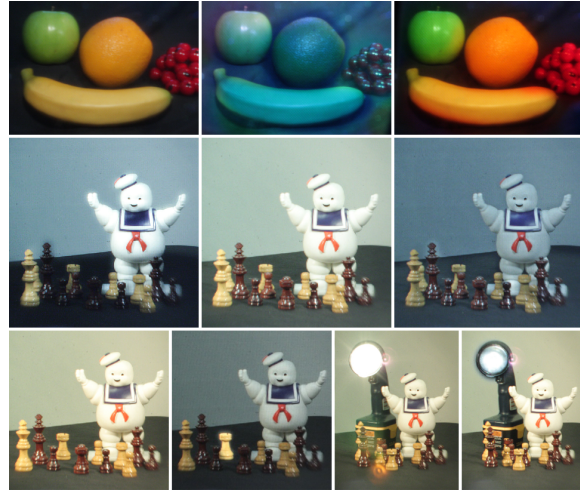


**Figure 9:** *Results from parallax experiments: color saturation changes using our second prototype (top), synthetic gamma (center row), object highlighting (lower left), and contrast reduction (lower right).*

result, SOPhIE successfully processes dynamic scenes with moderately fast-moving objects without apparent latency, a key advantage over watching a processed video stream on a standard display.

## 5. User Study

To validate the effectiveness of the SOPhIE framework, we picked one of the proposed applications for evaluation with a user study. The application we chose is contrast reduction according to Section 4.1. We selected this application due to its wide applicability in all possible hardware embodiments of SOPhIE, including both binoculars and scopes, as well as sunglasses or windshields. We leave a full evaluation of the other application scenarios for future work.

For our study, we set up a bright light source and an LCD screen that subjects had to observe through our scope prototype (Figure 10). The visual path was blocked such that the only means to see the screen was through the scope. One hundred random characters, that were closely located to the light source, were presented to each of 12 subjects with normal or corrected-to-normal vision. Each character flashed for a fixed duration of 0.5 seconds. The lamp had a size of 2 visual degrees, the characters 0.7 degrees, and the distance from the center of the light source to the center of the stimuli was 2 degrees.

For every character, we randomly enabled or disabled our contrast reduction. The characters were shown on a black background in five different intensity levels, randomly selected for each character, with Weber contrasts of 5.25, 8.55, 14.2, 45.6, and 102 respectively. We measured the Weber contrast of the light source with respect to the background as 14900. Weber contrast is given as $(L - L_b)/L_b$, where $L$ is the luminance of the stimulus and $L_b$ that of the back-



**Figure 8:** *Example of object highlighting. Top row: object highlighting by darkening all but the region of interest. Middle row: more subtle direction of attention (right) by selectively sharpening some regions (left) of the image (center). Bottom row: highlighting by brightness (center) and color (right) manipulations using our second setup.*

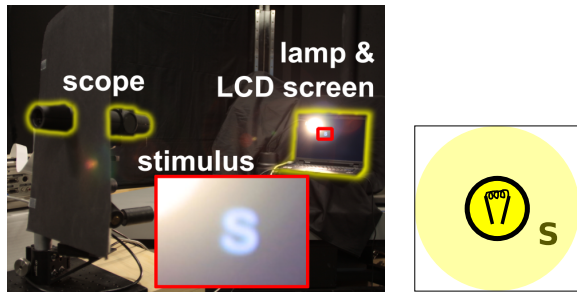ground. Weber contrast is the preferred unit for non-grating stimuli presented on uniform backgrounds.



**Figure 10:** *A photograph of the physical setup of our user study with its different components highlighted (left). The schematic on the right shows what a human observer sees through the scope.*

Results of our study are shown in Figure 11. The detection probability is the number of correctly recognized characters divided by the number of displayed characters for each of the 10 different settings. The plot shows average detection probability, and according variances across the subjects (left). Psychometric functions (as seen on the right) were fitted using the *psignifit* toolbox for Matlab (see bootstrap-software.org/psignifit/) which implements the maximum-likelihood method described by [WH01]. The expected detection probability of 0.5 could be shifted from a Weber contrast of 42 without modification to 5 using our contrast modulation technique, which is a significant improvement.

As expected, the light source creates glare that is caused by light scattering in the human eye. Effectively, this increases the perceived intensity around the lamp and masks details in the veiling glare. Instead of uniformly darkening the entire scene like regular sunglasses, our contrast reduction decreases the light intensity for only the brightest parts. This unveils imperceivable details that would otherwise be hidden in the halo of the light source created by scattering in the human visual system.

The acquired data supports our arguments; stimuli of lower contrast can be veiled by objects of higher contrast. In situations such as driving, the sun or headlights of other cars can limit a driver's visual performance. Traffic signs or other crucial information may not be perceived as intended. By selectively reducing the contrast of a scene as proposed in this paper, we can significantly reduce glare created in the eye, and enhance the visual performance of a human observer.

## 6. Discussion and Conclusions

### 6.1. Feasibility and Limitations

The sunglasses concept presented in Figure 1 may seem futuristic, but all the necessary ingredients are already present in a ubiquitous computing platform. Modern mobile phones contain small liquid crystal displays, very small cameras,
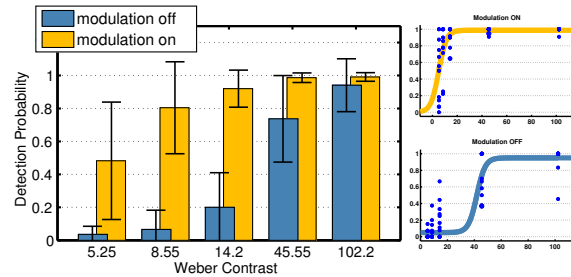


**Figure 11:** *The results of our user study confirm that details in high contrast scenes that are imperceivable due to veiling glare can be unveiled with our framework. These plots shows average stimulus detection probabilities and variances (left) as well as psychometric functions fitted to the data (right).*

and low-power processors suited for image processing calculations. As more processing power becomes available, sunglasses seem quite feasible as a SOPhIE platform in the near future. Work in augmented reality has already demonstrated that these components can be integrated into wearable systems [Man97]. Other scenarios, such as automobile windshields, still face some limitations of the underlying technology that must be addressed. Here we discuss the challenges in display, tracking, and processing technology that face a practical implementation of the SOPhIE approach.

**Display technology.** Liquid crystal displays are a mature, mass-produced technology, but have several disadvantages for our purposes. First, they rely on polarization of the incoming light and thus immediately cut the incoming intensity by half. This may prove acceptable for sunglasses, or related scenarios like ski goggles, but not for windshields. Our prototypes have a light throughput of about 10%. For context, modern greyscale LCD panels have maximum transparencies of up 45%, considerably more transmissive than most sunglasses (5%-20%).

Mass-produced LCDs have increasingly high resolution. However, for strongly defocused settings such as sunglasses this resolution is unnecessary, as fine details are blurred away, and even undesirable, since a small dot pitch causes color banding due to diffraction. We viewed scenes through pixel grids printed onto transparencies to empirically determine that a minimum pixel size of about 0.02 mm avoids visible diffraction artifacts in such out-of-focus settings. LCD panels designed for SOPhIE applications would require fast response, low resolutions, high contrast, and for some applications greyscale rather than color—an unusual combination not currently mass-produced. Finally, the flat nature of current LCDs could hamper the ergonomic and fashion design of products like sunglasses; however, companies are beginning to offer curved and flexible LCD panels [Cra05].

Other technologies exist but possess their own limitations; for example, commercial *electrochromic* products offer an excellent attenuation range but are slow and power hungry. The reflective micromirror arrays used by Nayar et

al. [NBB06] would be ergonomically awkward and suffer diffraction problems in defocused situations.

**Real-time image processing**. Clearly any image processing performed by the SOPhIE system must be done in real time. This limits the algorithms we can consider employing. However, modern mobile processors are extremely powerful, particularly for algorithms with dedicated fixed-function support. For example, the NVIDIA Tegra processors perform H.264 720p video decode at 30 frames per second, using well under 200 milliwatts full-chip [NVI08]. Furthermore, many of our algorithms operate at relatively low resolutions because of the eventual defocus blurring of the already low resolution display. In our prototypes almost all of the presented algorithms run at several hundred frames per second on current GPUs and thus have no measurable impact on latency. The exceptions are countershading and Rasche's recolorization algorithm [RGW05], for which we currently use unoptimized CPU-based implementations.

The latency of our prototypes is fairly high due to the loose coupling of cameras and displays with a normal desktop PC. Specifically, the scope and the window-style prototype have system latencies of 72 ms and 61 ms respectively. In a commercial handheld system these latencies would be drastically reduced, because the display and camera would be custom designed and much more tightly integrated with the processor. Nevertheless, even our current prototypes proved sufficiently fast to handle speeds up to 1 $m/s$ at a distance of 3 m without noticeable visual artifacts.

**Head and feature tracking.** We assume in all of our applications that the relative positions of the camera and display are fixed and calibrated, and that the eye position relative to the display is either fixed, or known approximately through other means. These assumptions hold in any setup in which the optical system is either held to the head, such as binoculars and monocular scopes, or attached to the head in a fixed fashion as with eye glasses or helmet visors. Larger systems like windows or car windshields could use head tracking, although multi-user scenarios require further research.

Our object highlighting applications may require more sophisticated real-time recognition and tracking of important features, which imply more daunting computational requirements. However, advances in graphics hardware and computer vision algorithms appear promising. For example, real-time face tracking is available commercially [See08] and robust detection and tracking of pedestrians is a 'hot topic' in computer vision [SM07]. As processors continue to evolve, it seems reasonable to imagine such applications running in real-time on mobile hardware.

### 6.2. Conclusions and Future Work

In summary, we have presented a novel approach for improving the performance of the human visual system by spatially modulating the light incident at a human observer. Applications such as contrast enhancement and reduction, color ma-

nipulation, and object highlighting could help humans process visual information more effectively, as demonstrated by our user study. This approach could someday help reduce risk in safety critical applications.

We have evaluated the feasibility of both a scope-like setup, and a thin device, such as eye glasses. The basic technologies for lightweight, mobile SOPhIE systems exist today; no significant technical hurdles prevent implementations of such devices. Advances in display technology and processing power will increase the reach and application of the SOPhIE approach. We plan to investigate the feasibility of more complex optical setups and algorithms including spatial convolution and non-linear operations.

In the future we will further customize specific algorithms and applications for SOPhIE, and extensively test and evaluate these with more user studies. A detailed analysis of binocular parallax effects is also left for future work. We believe that the dominant eye effect and other properties of the human visual system can be effectively exploited to produce high quality results in setups where both eyes share the same modulation display.

We also plan to investigate applications of the SOPhIE approach to future AR display systems. Combining traditional additive AR features with our multiplicative attenuation technique to achieve effects such as consistent illumination [BGWK03] appears natural; many such tasks, while difficult to realize with either additive or multiplicative display in isolation, become feasible when using both at once.

### 7. Acknowledgments

### References

[ABB*01] AZUMA R., BAILLOT Y., BEHRINGER R., FEINER S., JULIER S., MACINTYRE B.: Recent advances in augmented reality. *IEEE Comput. Graph. Appl. 21*, 6 (2001), 34–47.

[BA83] BURT P., ADELSON E.: The Laplacian Pyramid as a Compact Image Code. *IEEE Trans. Comm. 31*, 4 (1983), 532–540.

[BGWK03] BIMBER O., GRUNDHÖFER A., WETZSTEIN G., KNÖDEL S.: Consistent Illumination within Optical See-Through Augmented Environments. In *Proc. ISMAR* (2003), pp. 198–207.

[BPR83] BECK J., PRAZDNY K., ROSENFELD A.: *Human and Machine Vision*. Academic Press, 1983, ch. A theory of textural segmentation.

[BR05] BIMBER O., RASKAR R.: *Spatial Augmented Re-*

*ality: Merging Real and Virtual Worlds*. A K Peters, Ltd., July 2005.

[BVM97] BRETTEL H., VIÉNOT F., MOLLON J. D.: Computerized simulation of color appearance for dichromats. *JOSA A 14*, 10 (1997), 2647–2655.

[BW97] BORN M., WOLF E.: *Principles of Optics*, 6th ed. Cambridge University Press, 1997.

[CHR04] CAKMAKCI O., HA Y., ROLLAND J. P.: A Compact Optical See-Through Head-Worn Display with Occlusion Support. In *Proc. ISMAR* (2004), pp. 16–25.

[CR06] CAKMAKCI O., ROLLAND J.: Head-Worn Displays: A Review. *IEEE Journal of Display Technology 2*, 3 (2006), 199–216.

[Cra05] CRAWFORD G.: *Flexible Flat Panel Displays*. John Wiley and Sons, July 2005.

[FMS93] FEINER S., MACINTYRE B., SELIGMANN D.: Knowledge-Based Augmented Reality. *Comm. of the ACM 36*, 7 (1993), 53–62.

[HE99] HEALEY C., ENNS J.: Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE TVCG 5*, 2 (1999), 145–167.

[Hra02] HRADAYNATH R.: *An Introduction to Night Vision Technology*. Defence Research & Development Organization (India), 2002.

[ITU90] ITU: *ITU-R BT.709, Basic Parameter Values for the HDTV Standard for the Studio and for International Programme Exchange*. Standard Recommendation 709, International Telecommunication Union, 1990.

[KKO01] KIYOKAWA K., KURATA Y., OHNO H.: ELMO: An Enhanced Optical See-Through Display Using an LCD Panel for Mutual Occlusion. In *Proc. ISMAR* (2001), pp. 186–187.

[KM88] KINGDOM F., MOULDEN B.: Border effects on brightness: a review of findings, models and issues. *Spatial Vision 3*, 4 (1988), 255–262.

[KMS07] KRAWCZYK G., MYSZKOWSKI K., SEIDEL H.-P.: Contrast restoration by adaptive countershading. *Computer Graphics Forum (Proc. Eurographics) 26* (2007), 581–590.

[LP08] LINGLEY A., PARVIZ B.: Multipurpose Integrated Active Contact Lenses. *The Neuromorphic Engineer* (2008).

[Man97] MANN S.: Wearable Computing: A First Step Toward Personal Imaging. *IEEE Computer 30*, 2 (1997), 25–32.

[MFA*05] MANN S., FUNG J., AIMONE C., SEHGAL A., CHEN D.: Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences. In *CHI extended abstracts* (2005).

[MRT08] MOHAN A., RASKAR R., TUMBLIN J.: Agile Spectrum Imaging: Programmable Wavelength Modulation for Cameras and Projectors. *Computer Graphics Forum (Proc. Eurographics) 27*, 2 (2008), 709–717.

[NB03] NAYAR S., BRANZOI V.: Adaptive Dynamic Range Imaging: Optical Control of Pixel Exposures over Space and Time. In *Proc. ICCV* (2003), pp. 1168–1175.

[NBB04] NAYAR S., BRANZOI V., BOULT T.: Programmable Imaging using a Digital Micromirror Array. In *Proc. CVPR* (Jun 2004), pp. 436–443.

[NBB06] NAYAR S. K., BRANZOI V., BOULT T. E.: Programmable Imaging: Towards a Flexible Camera. *IJCV 70*, 1 (2006), 7–22.

[NS90] NAGY A., SANCHEZ R.: Critical color differences determined with a visual search task. *JOSA A 7*, 7 (1990), 1209–1217.

[NVI08] NVIDIA TEGRA:. www.nvidia.com/page/handheld, 2008.

[RGW05] RASCHE K., GEIST R., WESTALL J.: Recoloring Images for Gamuts of Lower Dimension. *Computer Graphics Forum (Proc. Eurographics) 24*, 3 (2005), 423–432.

[RHF94] ROLLAND J., HOLLOWAY R. L., FUCHS H.: A comparison of optical and video see-through head-mounted displays. In *SPIE 2351* (1994), pp. 293–307.

[RWPD05] REINHARD E., WARD G., PATTANAIK S., DEBEVEC P.: *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers, December 2005.

[See08] SEEING MACHINES:. www.seeingmachines.com, 2008.

[SM07] SABZMEYDANI P., MORI G.: Detecting Pedestrians by Learning Shapelet Features. In *Proc. CVPR* (2007), pp. 1–8.

[SSK01] SHIH M., SHISHIDO A., KHOO I.: All-optical image processing by means of a photosensitive nonlinear liquid-crystal film: edge enhancement and image addition&subtraction. *Optics Letters 26* (2001), 1140–1142.

[Sut68] SUTHERLAND I. E.: A head-mounted three dimensional display. *Fall Joint Computer Conference* (1968), 757–764.

[WH01] WICHMANN F. A., HILL N. J.: The Psychometric Function: Fitting, Sampling and Goodness-of-Fit. *Perception and Psychophysics 63*, 8 (2001), 1293–1313.

[YWK*06] YELLESWARAPU C., WU P., KOTHAPALLI S., RAO D., KIMBALL B., SAI S., GOWRISHANKAR R., SIVARAMAKRISHNAN S.: All-optical spatial filtering with power limiting materials. *Optics Express 14* (2006), 1451–1457.