Colorimetric and Photometric Compensation for Optical See-Through Displays

Christian Weiland[1], Anne-Kathrin Braun[2], and Wolfgang Heiden[1]

[1] University of Applied Sciences,
Bonn-Rhein-Sieg, Germany
Christian.Weiland@iais.fraunhofer.de,
wolfgang.heiden@fh-bonn-rhein-sieg.de
[2] Fraunhofer Institute for Applied Information Technology,
Sankt Augustin, Germany
Anne-Kathrin.Braun@fit.fraunhofer.de

Abstract. Optical see-through displays are an established technology within augmented reality. Wearing such a display the users eyes automatically adapt to the luminance of the real world environment, while the virtual part is displayed using a steadybrightness. This often results in clear differences between real and virtual elements. This paper shows a technique for colorimetric compensation which avoids this effect. Furthermore algorithms for photometric compensation will be demonstrated. The appearance of background shapes and colours arise from the combination of the luminance of the background and the projected luminance of the object. These "ghosts" are photometrically compensated.

Keywords: augmented reality, see-through display, colorimetry, photometry, compensation.

1 Introduction

The human eye can perceive a huge range of luminance, with absolute levels ranging $1 : 10^{14}$ orders from nocturnal starlight to sunlight on a very bright day and the dynamic range of light can exceed $1 : 10^4$ orders. Despite this wide range the human visual system is able to generate valuable information for us. It becomes even more astonishing as one single photo receptor is only capable of a dynamic range of about 10^3 orders.

The human eye uses four mechanisms to handle different levels of luminance and adapt to the specific situation: the pupil, the rod and cone system, bleaching and regeneration of photo pigments and neural processes. The eye's receptor cells called rods and cones respond differently to luminance levels: While the achromatic rods work at scotopic illumination levels of 10^{-6} to 10 cd/m^2, the chromatic cones react at the photopic range of 0.01 to 10^8 cd/m^2. In the mesopic range between 0:01 and 10 cd/m^2 both receptors contribute to vision. All these mechanisms effect our perception of luminance and color. They usually show non-linear behavior and their functionality is still not understood completely.

A person wearing a see-through display will move around and encounter various scenes with various illumination levels, but an optical see-through display, does not consider the current adaptation of the eye. This causes a visual difference between the real and the virtual part of the viewed scene especially when the eye adapts to the real world while the virtual part remains uncorrected. In this paper a technique for photometric compensation is developed which measures scene-luminance with a camera and changes the virtual part of the scene to be adapted to the same luminance level, so that perception is uniform across the scene. The colorimetric compensation algorithm control the color of the virtual image per pixel. In this paper a technique is developed which uses a camera-recorded image to compensate radiance from reality in the virtual image.

For realization, we attached a HDR camera to a head-mounted display. Despite this camera location shifts the center of gravity to the front the wearing comfort is still given and the see-through display does not slip. The user's field of view is not limited by the camera as it is mounted high enough. The compensation algorithm was integrated into the VR/AR framework MORGAN [6].

2   Related Work

There are several publications containing valuable information for this work. Nevertheless none of these covers the exact topic of colorimetric or photometric compensation for see-through displays, so this work is considered to be a novelty. Most information for the colorimetric part arises from the area of tone reproduction and realistic image synthesis. Photometric compensation for see-through displays was not developed yet, but there is literature covering compensation for video projectors, which cannot be applied directly.

Ferwerda et al. [2] present an advanced model based on psychophysical data from experiments, which produced a *threshold-versus-intensity* function. This function describes the threshold luminance for which a flashed disk is visible on a background of lower luminance. Another experiment analyzed the changing sensitivity or the rods and cones under different levels of luminance intensity. The results of both experiments are important parts in the tone reproduction system.

The model includes four eye adaptation mechanisms, which are receptor system behaviour, pigment bleaching and fast and slow neural processes. Ferwerda thankfully provided images from his work which are used for box filter calibration in colorimetric compensation.
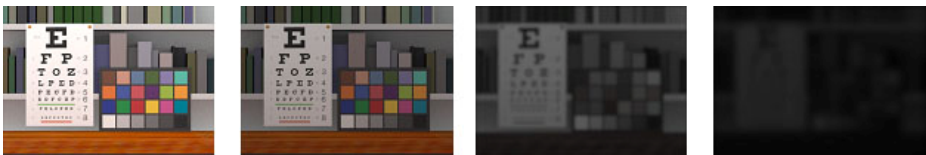

Fig. 1. Tone reproduction images from Ferwerda et al. [1]

Irawan et al. [3] base their method on Ferwerda's work and extend the model. They set their focus on processing HDR image streams and introduce an operator capable of computing adaptation from a single starting point of time. The main part of the operator is an exponential decay function which mimics biological behaviour.

There is work on photometric compensation for video projectors done by Bimber et al. [1]. They compensate the virtual image emitted by the projector by considering the environmental light and the surface material. Bimber implemented a version for smart beamers with shaders which is real-time capable. The main difference to see-through displays is the direct interaction of emitted light and the surface which allows manipulation of the result. For see-through displays the interaction is independent from the material term.

Another interesting solution in hardware is introduced by Kiyokawa et al. [5]. They mount a LCD display directly in front of the combiner and by opening and shutting its pixels can block the background light.
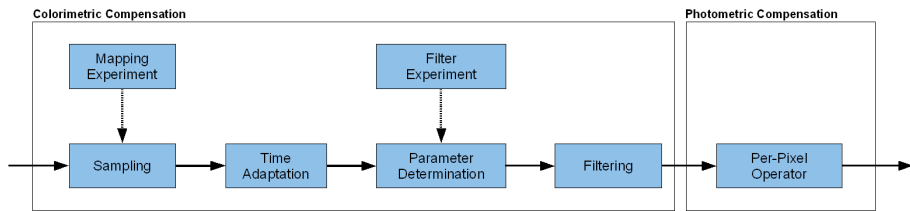
3 Compensation

3.1 Colorimetric Compensation

The colorimetric compensation system is based on the model of Ferwerda et al. [2] for tone reproduction to compute an image under certain levels of eye adaptation. Since Ferwerda's model is too complex to be computed in real-time we developed an approach to estimate the effects by image filters, e.g. Gaussian or Box filters. The virtual image is read from the current framebuffer and the optimized filter is applied on it to achieve effects close to those Ferwerda's model would have produced. The result is written back to the buffer and displayed by buffer swapping.

Fig. 2. Colorimetric compensation applied on the Stanford Dragon. With increasing luminance sharpness, brightness and chrominance are modulated accordingly. (The negative pattern arises from photometric compensation.)

The approach consists of six steps shown in figure 3. The mapping experiment, the sampling and the time adaptation determine the luminance level the user's eye is adapted to. Essentially is a mapping from RGB pixel value to a luminance value, which is created in an offline experiment. Using the results from the mapping the environmental image is sampled. Therefore, a number of sampling points are updated each frame and their corresponding luminance values are added. The sum is divided by the number of samples to get the mean luminance in the image.

Fig. 3. Overview of compensation algorithm

Sampling is implemented for rods and cones independently. The coordinates of the sampling points are determined randomly at start-up time. The cone points are distributed over a small region at the center of the captured image, while the rod points are uniformly distributed over the entire image.

The number of sampling points is a critical parameter. Using too many points, sampling the pixels would lead to considerable time costs. Yet if chosen too low, deviations and per-frame jitter would occur. To solve this problem a sample buffer is used as suggested in [8]. This buffer is divided into several slices, of which one per frame is updated. The sums of all samples of a buffer are held and only the sum for the updated slice is recomputed. This technique allows usage of a sufficient high number of sample points while at the same time causing minimal computation costs and reducing jitter signicantly. In the final version of this work, 80 sample points for the rod and cone system each are used, divided into four slices. Higher numbers up to 1000 samples are tested which lead to more stable mean values but higher computational costs.

After the sampling step, the environment luminance is known and can be used to compute the eye's adaptation level including time effects. Originating from the time adaptation model presented in [3] and the time courses in [2] the form of the time adaptation function is known. Time course is modeled by an exponential decay function and for the intensity of adaptation a scaling factor from the difference between adaptation and measured luminance is included [3].

Another important feature is the blindness resulting from rapid changes in the environment luminance. The eye adaption function above compute the alignment of the current adaptation state to the measured luminance level, but do not consider the effects a rapid luminance change causes in the human eye. These effects result in a temporary shift to the monochromatic domain for the whole image. For simulating this effect a combination of two parameters is developed in this work and integrated in the colorimetric compensation system.

The first parameter will recover data by a reverse lookup of the mapping experiment, which allows a simple transformation from the measured luminance to a RGB value of which the arithmetic mean is taken. By this lookup a luminance value can be extracted from the measured luminance. As this parameter depends not on the eye adaptation level but on the measured luminance it gives a value that is independent on whether the measured luminance is larger or smaller than the adaptation level.

The second parameter describes the power how the first parameter influences the current pixel color.

The further steps described in figure 3, the filter experiment, the parameter determination and the filtering, modify the virtual image to adjust sharpness, brightness and chrominance. In an offline filter experiment the parameters for steerable filters will be evaluated. Therefore a reference image with 1000 cd/m² is chosen as this luminance level is nearest to the level of an uncompensated virtual image. Taking this image as source, the other images at 10, 0.04 and 0.001 cd/m² are computed. The reduction of sharpness is effected naturally by any normalized Gauss or Box filter, controlled by the parameter size. The pixel intensity can also be easily influenced by multiplying the whole kernel by a factor *scale*. Chrominance is more complex as the pixel luminance must be preserved. Therefore another factor e is introduced as the different color channels contribute differently to luminance perception in the human eye, e.g. green is much better perceived by the cone receptors (the empirically found final values are: $e_r = 0.35$, $e_g = 0.59$ and $e_b = 0.06$).

A simple Euclidean distance function between reference image pixel and filtered image pixel gives a criteria for the similarity of the images and thus for the accuracy of the current parameters.

Fig. 4. This series of colorimetrically compensated images shows the color dampening effect from full monochromatic level (left) in a scotopic environment to fully restored color (right) in a photopic environment

This approximation works remarkably well for Gauss and Box kernels, especially with the 10 cd/m² image, where no visual difference can be detected by observers. While at the darker images brightness and color intensity can be achieved to a not noticeable degree, sharpness cannot be reduced well as the form of the filter kernels leads to visible artifacts and higher distances.

One essential parameter is the kernel size. For the 0.001 cd/m² image it reaches for both Gauss and Box kernels values around 20. With optimizations techniques the computational complexity of Box filters can be made independent of the kernel size, which is the deciding argument for Box and against Gauss filters in the final system.

From the extracted luminance values the parameters which steer the image filter have to be computed. From the filter experiment the parameters size, scale and damp for every of the source images are known. For every measured luminance value between these parameters has to be interpolated. As final solution a linear interpolation is chosen, which considers only the image with the next higher and next lower luminance.

In the final step of colorimetric compensation the parameters from the former step are applied to a Box filter which is used to change the virtual backbuffer image to be adapted to the environment luminance. Though Gauss filters give better visual results the possibility to accelerate box filters and to decouple their runtime complexity from the kernel size is essential for a real-time application. Therefore box filters are chosen for the final system.

3.2 Photometric Compensation

The photometric compensation takes place after colorimetric compensation. Inputs are the (colorimetrically compensated) virtual image and the background image which have to be aligned with pixel accuracy.

We developed several compensation approaches: the trivial compensation, the subtraction compensation, the simple and the advanced smooth compensation. The trivial compensation approach is based on work for background compensation for video projectors ([1]). The computation step is to subtract the background from the virtual part pixel per pixel.

The main advantage of this technique is its simplicity and therefore its speed. The worst-case occurs when all background color channel values are larger than the virtual values (figure 5), which renders the resulting pixel black. Though this can hardly be seen for single pixels, a bright background area can cause the virtual part to be completely black and thus invisible. Hence the trivial approach has serious problems with preservation of the virtual part with both luminance and contrast. Background luminance can be compensated well for dark backgrounds but this effect decreases with increasing luminance

To avoid the shortcomings of the former approach, in the subtraction compensation the virtual part is introduced a base and only a fixed part is made dependent of the background. The constant parameter k denotes the fraction of the base. This small change weakens the disadvantages of normal Trivial Compensation and provides a decent compensation operator already. Background luminance is compensated to a lesser degree as in Trivial Compensation due to the base k. Contrast and luminance of the virtual part are maintained to some extent, yet the luminance of the virtual image is darkened for every $k < 1$ and $b_i > 0$. Thus if the virtual image is dark itself and the background is bright, r_i may become not noticeable anymore. To cope with this problem, the operator is rewritten to:

$$r_i = k_v * v_i - k_b * b_i \tag{1}$$

Here both parameter k are made independent and can also be greater than 1. Adequate values found are 1.15 for k_v and 0.25 for k_b. Yet optimal values can vary dependent from virtual scene and environment. The difficulty of this approach lies in finding good parameter k that matchs not only one single setup of virtual objects and surrounding but many. In the experiments of this work no suitable parameter k are found that are robust to all setups.

Fig. 5. This image shows the weakness of Trivial Compensation: Parts of the face disappear due to the brighter background

The simple smooth compensation arisen from the experiences with the former methods fulfill several demands on a compensation algorithm: It has to fulfill the four quality criteria and additionally must have smooth transitions between its adjustments to all parameters. Thus a new formula is created as fraction depending on the background value combined with the maximum intensity value max. This results in a smooth transition factor s which served as multiplier for the virtual part.

$$s_i = \frac{\max}{b_i + \max}$$

$$r_i = s_i * v_i$$

(2)

This compensation formula increases linearly with v_i and due to s decreases smoothly with rising b_i. By this behavior higher visual comfort for the user is achieved. Because of this smoothness the compensation of the background is weaker compared to e.g. Subtraction Compensation.

Virtual luminance and contrast are preserved with reasonable success. The complexity of the formula is as about as simple as in the former algorithms One disadvantage of the used equation is that the pixel value is always decreased. Another point is that s is independent of the virtual part as v_i is not included in the formula, thus there is no additional scaling with respect to v_i.

The Simple Smooth Compensation method already achieves reasonable results, but has the main weakness in not including the virtual part in s. This is changed with the advanced version as the new formula is a fraction of sums of differently weighted virtual and background values, resulting in a more complex s:

$$s_i = \frac{c * (\max - v_i) + b_i}{(\max - v_i) + d * b_i}$$

$$r_i = s_i * v_i$$

(3)

c and d are weights used to adjust the slope of s. In this work c = 1.4 and d = 2.0 are found appropriate values. For c > 1 the factor s can exceed 1 which is an intended increase of the virtual part. By inverting the virtual value in s at (max - v_i) the slope of the function decreases for higher values of v_i, which is a desired effect.

This compensation responds smoothly to changing v and b both, which achieves good visual comfort for the user. The main advantage of this operator is its robustness

as it responds smoothly to the environment luminance and scales with the virtual pixel intensity v_i. Nevertheless to achieve best results the parameters c and d can be adjusted to a specific setup.

For c = 1.4 and d = 2.0 virtual luminance and contrast are preserved well except for extreme values of v_i and b_i. Despite its complexity the formula does not significantly slow down the system compared to the former algorithms.

Weaknesses originating from the formula appear at the minimum and maximum values. For b ≈ 0 the factor s reaches its peak c, which leads to a increase of the virtual part for a c > 1. Experiments showed that even through this raised virtual value the falsification is unnoticed during use and virtual contrast is preserved. Nevertheless for high-luminance virtual images there is the possibility for contrast loss. For b → max s reaches 1/d which results in a reduction of r. To avoid exaggerated darkening of the image d should be chosen between 1 and approximately 4.

4 Results

The colorimetric compensation system described in this paper revises the visual part of an image to the adapted eye and thus improves visual quality for the user.

The delay occurring from updating the ring buffer effects the time function only on slower machines as it is linked to the frame rate. Future versions may couple it instead to the passed time made independent from the kernel size which allows application of large kernel without time slowdown. This acceleration method is a major key to real-time. The exaggeration of edges in the image coming from the square form of the box filter is barely visible to the user. Irawan's model of adaptation over time [3] could not be reconstructed completely, but from its form a similar own model is derived. One disadvantage of all time adaptation models is a small frame-rate dependent error.

The different photometric compensation operators introduced differ in quality of compensation for different virtual contrast ranges, level of background luminance and applicability. Trivial Compensation is a basic approach that can work for video projectors [1], but can cause severe image information loss for see-through displays. Especially for bright backgrounds parts of the virtual content can vanish.

Though it shares form of the equation with Trivial Compensation, Subtraction Compensation gives the best results and satisfies all quality criteria well. On the contrary it has to be configured for a specific setup. Virtual objects and scene luminance are supposed not to change much as otherwise quality decreases. Therefore Subtraction Compensation is the best choice for setups with known and limited content.

Figure 6 shows the difference between a compensated and an uncompensated image. In the uncompensated image (on the left side) the colored pattern is visible all over the face region and the grid shines through the image. Especially the yellow field behind the mouth and the light red field behind the left cheek appear disturbing.

Subtraction Compensation is applied to the same virtual image and shown on the right side. One major compensation effect can be seen in the forehead, eye and cheek region where the background grid is no longer apparent. The colored background fields are also barely visible anymore. These regions are all bright and there the compensation operator functions best. In the mouth region virtual luminance is lower and the background becomes more visible. For the bright yellow field at the mouth red
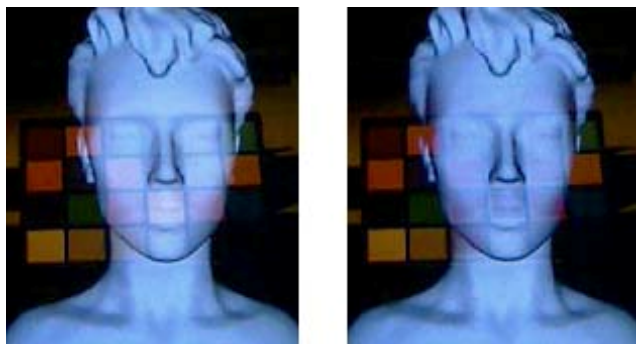
Fig. 6. Images captured by a webcam through the see-through display. Left is the uncompensated image, the right one is compensated with Subtraction Compensation

and green (which are composite colors of yellow) are subtracted partially. Together with the strong blue shift of the webcam this causes the field to appear purplish. Another effect is a gain in the virtual image's contrast. Especially darker regions like the temples are enhanced. This effect originates in the fact that the compensation operator subtracts the background independent of the virtual pixel value and thus darker pixels are reduced more.

Both Smooth Compensation operators preserve virtual luminance and contrast and compensate the background but not as strong as a scene-optimized Subtraction Compensation operator. Their advantage is that they function over a wide range of luminance levels and need no recalibration. In direct comparison the normal version is faster while the advanced version compensates better due to inclusion of the virtual value in the equation.

5   Conclusion

For colorimetric compensation more measurements are needed to provide more sample values for interpolation. Usage of a luminance meter that can be configured for scotopic, mesopic and photopic range would lead to very exact pixel-to-luminance mapping and allows independent computation of rod and cone adaptation.

Another task is the development of the hardware system. Integration of the camera into the see-through display could be implemented using a beam splitter. This improvement would not only increase weight of the system and thus wearing comfort, but also solve the problem that the viewing distance has to be known.

Adaptation over time is another task for future research. More psychophysical experiments, models and equations for the time course are needed to improve the colorimetric compensation system. A special focus has to be laid on rapid luminance changes and their influence on later eye adaptation effects. This focus is necessary for a quickly changing environment as it may occur in real-time applications with a moving user. An advance is also possible by new time course equations which reduce the per-frame error.

For photometric compensation there are two directions of research. The first one is to develop further advanced per-pixel operators. Both smooth operators are expected to be good examples as they are the most robust operators tested. A second way is to evaluate effects and speed of operators that include neighborhood information or even information about the whole image. A strong decrease of speed is expected.

Acceleration can not only be done with parallel processing, but also on the graphics processor. Its parallel shader units and the native vector operations have great capability to accelerate the system. This would also omit reading and writing from the GPU backbuffer.

A distribution on several computers is possible as well, though the network speed is assumed as bottleneck. Communication has to be bidirectional. For colorimetric compensation scene information and environment luminance have to be send to the slave computers and the results read back. For photometric compensation part of the background has to be send to each slave as they need the background information for the compensation operator and the results read back. The transfer is supposed to slow down the system and limit the frame rate.

References

1. Bimber, O., Emmerling, A., Klemmer, T.: Embedded entertainment with smart projectors. Computer 38(1), 48–55 (2005)
2. Ferwerda, J., Pattanaik, S., Shirley, P., Greenberg, D.: A Model of Visual Adaptation for Realistic Image Synthesis. ACM Transactions on Graphics, 249–258 (1996)
3. Irawan, P., Ferwerda, J., Marschner, S.: Perceptually Based Tone Mapping of High Dynamic Range Image Streams. In: The Eurographics Association 2005: Eurographics Symposium on Rendering (2005)
4. Jarosz, W.: Fast Image Convolutions. In: SIGGRAPH Workshop (2001)
5. Kiyokawa, K., Kurata, Y., Ohno, H.: An optical see-through display for mutual occlusion of real and virtual environments. In: Proceedings. IEEE and ACM International Symposium on Augmented Reality 2000 (ISAR 2000) (2000)
6. Ohlenburg, J., Braun, A., Broll, W.: Morgan: A Framework for Realizing Interactive Realtime AR and VR Applications. In: Proceedings of the Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS) at IEEE Virtual Reality 2008 (VR 2008) (2008)
7. Pattanaik, S., Ferwerda, J., Fairchild, M.: A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display. In: Siggraph 1998 Conference Proceeding (1998)
8. Stamminger, A., Scheeland, M., Seidel, H.-P.: Tone Reproduction for Interactive Walkthroughs. In: The Eurographics Association 2000 (2000)