Popular Summary

More and more time is spent in front of displays, be it mobile phones, TVs, or computer monitors. The vast amounts of content made available through displays are able to entertain, teach, and act as an efficient way for people to communicate. Yet, an everyday user might not spend much time thinking about how the content is created and how the displays are able to present it. What goes on under the hood, so to speak? One key component in the setup, which will be an overarching topic in this thesis, is knowledge about how human vision works. That knowledge is used to create, share, and present incredible visual experiences.

When a lamp is turned on and it lights up a room, what happens when the light reaches the eye? Researchers have explored this and figured out how the eye works. They have discovered that, because of particular properties of the eye, many, though unfortunately not all, of the visible colors can be produced by combining red, green, and blue light. This finding has lead to displays often being constructed to only emit light of those three colors. In simple terms, displays can be thought of as having three types of lamps. One type emits red light, one emits green light, and one emits blue light. The light from the three types of lamps is then combined into colors by the human visual system. How much light each of the three types of lamps should emit to give the impression of a certain color is decided by the image that is to be presented. Each color in the image is described by three numbers: one number corresponds to the amounts of red, one to the amount of green, and one to the amount of blue that the display should use to produce the color. To produce a yellow color, for example, the numbers would result in the display showing red and green light—a combination which the human visual system interprets as yellow—while the blue lamp could be turned off. The second chapter in this thesis describes what happens when light reaches the human eye while the process of how displays present colors is described in the thesis' third chapter.

Another question is how one can determine the three color numbers and, with them, the color that should be shown on the display. When a photograph is taken, sensors inside the camera record the light coming from what the camera is directed toward. The light is then converted into the numbers that correspond to the color it is perceived to have. Finally, the photograph can be shown to the photographer on the camera's display. But what if the photographer wanted to capture a photograph of an environment when they themselves are not physically at the location? This could be the case if they wanted a photograph of an environment that is not real or the location in some way is out of reach for the photographer. If they had a *very* detailed descrip-

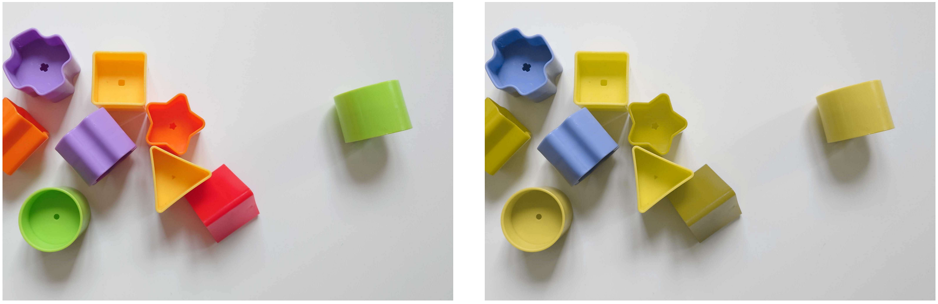
Illustration of a rendering setup including a teapot and a cheese slicer. **Left**: overview of the scene and its light sources. Behind the teapot is the virtual camera. The user can move it around the scene as desired. **Right**: image rendered from the perspective of the virtual camera. Note that both images are created using computer graphics—they are not photographs. *Images provided by Tizian Zeltner. Scene created by Andrea Weidlich, Toni Bratincevic, Davide Di Giannantonio Potente, and Kevin Margo.*

tion of the environment, however, such a photograph would be achievable through *photorealistic rendering*. Rendering is the process of taking the specific description of the environment, which is more commonly known as the *scene*, and producing an image of it. The description could, for example, include the location and extent of objects, as well as details about properties of the materials that constitute them. It could also include the positions of lamps, the sun, or other sources of light. The images above show an example of the setup in rendering and the finished, rendered image. The setup is similar to how photographs are taken by real cameras. The renderer—a computer program—uses the description of the scene and *simulates* how the scene's lights and objects interact. Once the simulation is done, the renderer knows what light reaches the camera, or eye, and can use that information to produce an image similar to a photograph. This process is explained in more detail in the third chapter of the thesis.

It is surprising how close rendered images can look to real photographs. The images above, for example, were generated through computer graphics and are not photographs. However, while renderers are capable of simulating light to some degree, doing so is a complicated task. In particular, it takes time for the renderer to do the computations necessary to precisely simulate how light interacts with its surroundings. For applications that need to render images quickly, such as video games, the time that an accurate simulation takes could be prohibitively long. Like any human in a hurry would do, the renderer then takes various shortcuts. The shortcuts lead to the images being completed faster, but the simulations are no longer exact. Because of this, the images may look less realistic, or otherwise distorted. Minimizing the negative effects of the renderer taking shortcuts is the goal of significant research efforts. A related research question is how the negative effects can be measured. With such measures, it is easier to build solutions that minimize the negative effects. Part of this thesis explores this in attempts to give users a visualization of which parts of quickly rendered images look distorted when compared to versions that were rendered with more accurate lighting simulations.

Another part of the scene-to-screen pipeline is storing and sharing the generated images. How is it possible for, e.g., a phone to store all the photographs that are captured with it? Each photograph could contain millions of numbers that describe its colors and phones could store many thousands of photographs—the amount of storage space required is vast! Fortunately, there are ways to represent each image with significantly fewer numbers. This *compression* unlocks the possibility of storing the large amount of numbers that a collection of photographs contains. Similar to how faster rendering could have negative effects on image quality, the cost of compressing images can be a quality degradation. One example of where this could occur is in chatting programs. In those, beautiful photographs are shared between users, but what might have looked like a high-quality, crisp image to the sender could look like it includes several blocks of the same color and is of generally worse quality to the person receiving it. Again like in rendering, significant efforts are being, and have been, focused on minimizing the negative visual effects of compression. Remarkably, it is possible to achieve compression without causing any degradation in quality. Under such a constraint, however, the compressed image might require more storage space than if quality-preservation was not a requirement. Compression is another topic discussed in this thesis.

While most viewers will have a similar visual experience when watching the same display, some people see the content on the display differently. In particular, some people see fewer colors than others—they suffer from *color vision deficiency* (CVD), which is a condition that is more commonly known as color blindness. The term color blindness is somewhat misleading, however, as most "color blind" people are able to see a large amount of colors, though not as many as someone with normal color vision. The effect of severe CVD can be seen in the images below. About one in 12 men and one in 200 women have some form of CVD. The condition is often genetic and cannot currently be treated. However, what can be done to help improve the visual experience of those with a CVD is being mindful when creating content that could be presented to them. Green and red colors, for example, are often confusing to someone with a CVD. The images demonstrate this: in the right image, which corresponds to what the person with a CVD would see when viewing the left image, it is difficult to tell the red, green, yellow, and orange colors apart because all have a yellow-brownish look. Therefore, it would help someone with a color vision deficiency if images did not contain content that can only be appreciated or correctly interpreted if the red and green colors in them can be told apart. Unfortunately, this



An image of differently colored toys. The **left** image is the original and the **right** image shows what someone with severe color vision deficiency might see when looking at the original. Notice that, from the perspective of the color-deficient person, it is difficult to tell which toys are red, which are green, and which are yellow.

cannot always be achieved. In such cases, it could instead be helpful to alter the image in a way that makes someone with a CVD able to separate red objects from green ones. One way to do that would be to make red objects blue. However, doing so would lead to confusion between red and blue objects instead of red and green ones, which illustrates one of the key difficulties with solutions like these—it could be that new color confusions arise as the changes are applied. Algorithms that change images in an attempt to alleviate the visual issues that a color-deficient observer might suffer are called *daltonization* algorithms. In this thesis, a new daltonization technique is proposed. The thesis also includes an introduction to the fields of daltonization and CVD simulation research and how the eye of someone with a color vision deficiency differs from the eye of someone with normal color vision.

Throughout this thesis, it becomes evident that knowledge about the human visual system has had a major impact on how images are created, processed, analyzed, stored, and displayed. One particular aspect that is considered is the limitations of the visual system. The human eye and brain are not able to interpret all the light that reaches them. This has several implications, including that small details, such as the squares that make up the screens of smartphones, go unnoticed. It is also why people generally do not notice that their lamps are constantly flickering. The scientific publications contained in this thesis consider phenomena such as those mentioned, as well as perceptual issues such as color vision deficiency, and provides ways to evaluate and improve on the quality of the output produced by renderers.