Achieving Perceptual Consistency of Contrast in Physical Images Viewed at Different Distances

Francisco Ayala Le Brun
Supervisor(s): Baran Usta, Michael Weinmann, Elmar Eisemann
EEMCS, Delft University of Technology, The Netherlands

27 june 2022

A Dissertation Submitted to EEMCS faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering

Achieving Perceptual Consistency of Contrast in Physical Images Viewed at Different Distances

FRANCISCO AYALA LE BRUN, BARAN USTA, MICHAEL WEINMANN, and ELMAR EISEMANN

1 ABSTRACT

According to the Contrast Sensitivity Function, the contrast of different spatial frequency elements of an image will be seen differently depending on distance. We propose a method for compensating for these differences by acting on an image's frequency representation and manipulating the magnitudes of the individual wave functions composing the image. We provide an open source implementation of the algorithm which can be used to process videos and images in real time.

2 INTRODUCTION

The Contrast Sensitivity Function (CSF) has important applications on our perception of media. This function, when given a perceived contrast, maps spatial density to a true contrast [Nadenau et al. 2003]. Critically, experimental results have found human's contrast discrimination capacity to be generally inversely proportional to spatial density. However, very low spatial densities also have been found to have an adverse effect on contrast discrimination [Whittle 1986] [Carney et al. 2000].

The CSF is an important part for describing the Human Visual System, with various applications in computer graphics and image processing. By using it to calculate contrast discrimination for image segments, it has been applied to image compression [Nadenau et al. 2003]. Applications have also been found in contrast enhancement by means of deriving invariants from its general shape [Majumder and Irani 2006].

An important factor in determining the effect of the CSF is distance. This is because an individual's total visual angle devoted to a physical surface, and thus spatial density as measured for the CSF, decreases linearly with distance. Thus distance, often changing in contexts such as cinema or where the viewer is moving, can have unintended consequences on the perception of a visual medium.

In this work, a new technique is proposed for maintaining a uniform contrast perception of an image presented in a physical screen by exploiting the properties of the Human Visual System. The algorithm uses the effect distance has on the CSF in order to dynamically adjust contrast according to the current position of the viewer. Calculations are done fragment-by-fragment in such a way that all of its processes are easily parallelizable. The algorithm runs in real time, thus opening the door for the use on dynamic content and situations when viewer position changes often.

Several examples are presented, where unmodified images are compared to modified images. Furthermore, a metric is used to visually quantify the differences between the modified image and the original image viewed at the target distance.

Authors' address: Francisco Ayala Le Brun; Baran Usta; Michael Weinmann; Elmar Eisemann.

3 RELATED WORK

In the past, the CSF has proved useful for several image processing techniques. One approach utilizes the properties of the CSF in order to change the apparent contents of an image depending on distance, by applying filters on two images' frequencies and combining them to create one hybrid image[Oliva et al. 2006].

Closely related to compensating for changes in contrast sensitivity related to distance, is detecting how changes are perceived by the Human Visual System. To this end, an approach was previously proposed in order to spatially filter images according to the CSF, in order to better quality of perception in a viewing distance-dependent manner [Zhang and Wandell 1997].

A more recent approach known as HDR VDP uses several effects from the Human Visual System such as non-linear response to luminance, contrast sensitivity and visual masking. This approach allows for both an absolute metric comparing two images and a fine-grained visibility map indicating which changes between two images are most visible [Mantiuk et al. 2005].

Another approach proposes a CNN-based metric, where a CNN is trained from a dataset of manually marked images with differences. In this case, the Human Visual System is emulated by the CNN [Wolski et al. 2018].

The method we propose, similarly to the first two works mentioned in this section, also utilizes the CSF directly to produce its output. However this method is used in order to replicate perception, instead of detecting differences in perception. Additionally, other works have not directly made use of the Fourier domain and its relationship with the CSF.

Compared to the previous techniques, our technique is only proposed for luminance images. The technique could be applied to the chrominance channels. However, the chrominance channels have a different CSF [Samu et al. 2002] and existing tools focus on measuring the luminance CSF [Canare 2021]. Thus modification of chrominance channels was decided to be out of the scope of this paper.

4 BACKGROUND

4.1 Contrast Sensitivity Function

The Contrast Sensitivity Function (CSF) models how the human visual system perceives contrast as a function of visual spatial density. Normally, the CSF is experimentally measured with the use of horizontally-varying Gabor patches. Gabor patches consist of a sinusoidal function modulated by a Gaussian function [Whittle 1986].

The CSF has been experimentally found to resemble the subtractive combination of an exponential and a Gaussian function [Carney et al. 2000]. Based on this model, the function increases until ≈ 3.5 cycles per degree, and subsequently approaches zero.

It is important to understand why the CSF implies that contrast is a function of viewing distance. Through simple trigonometric

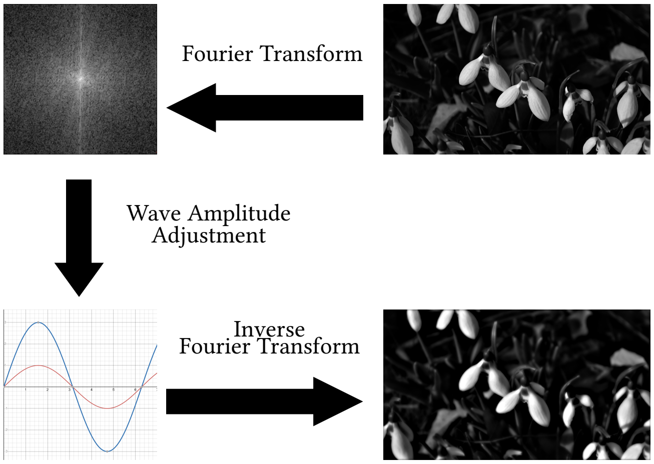


Fig. 1. Main steps of the algorithm proposed in this paper.

properties, it is possible to calculate a horizontal visual angle of a viewer towards a perpendicular screen of width w and distance d with the formula $v = 2arctan(\frac{w}{d})$, where v represents the visual angle.

If an image is a horizontal sinusoidal grating with n cycles, then the cycles per degree will be uniform across the image and exactly $\frac{n}{v}$. Substituting with the first function, we have $\frac{n}{v} = \frac{n}{2arctan(\frac{w}{d})}$. From this function, it is possible to see that as the distance d increases, cycles per degree increase. Relating this to the CSF, since cycles per degree increase, contrast discrimination decreases.

5 CONTRAST ADJUSTMENT ALGORITHM

The Contrast Sensitivity Function (CSF) tells us that human vision will respond differently to contrast depending on visual spatial density. This results in inconsistencies in how an image's contrast is perceived depending on changes in viewer's distance. Thus in this paper, we propose an algorithm for compensating for these inconsistencies in contrast.

The first step of the algorithm is to compute the two-dimensional Fourier transform of the image [Ukidave et al. 2014]. The second step is to interpret the frequency bins' magnitude and cycles per degree (cpd). Thirdly, the magnitude of the wave function in the frequency bin is scaled depending on viewing conditions and the CSF. Finally, the inverse Fourier transform is used in order to obtain the transformed image. These steps of the algorithm are illustrated in Figure 1.

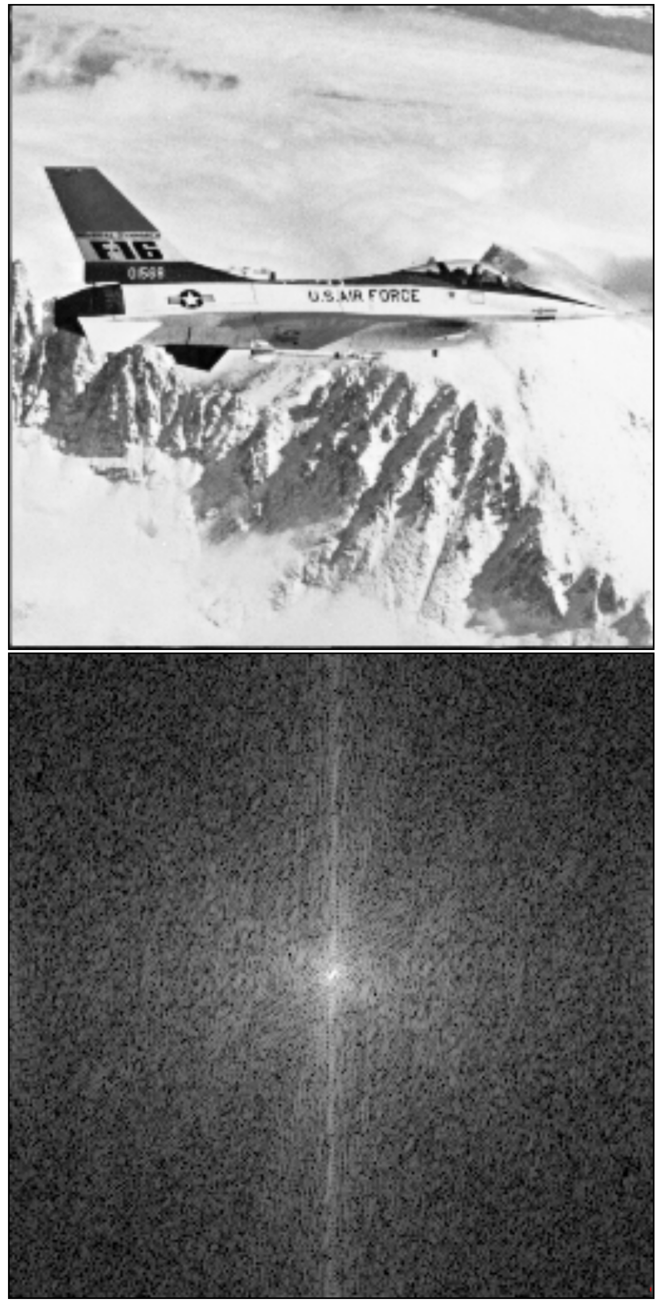


Fig. 2. Image of jet fighter in spatial domain and its frequency domain power spectrum.

5.1 FFT and Inverse FFT

The Fast Fourier Transform (FFT) and inverse FFT are used in order to transform the image from the spatial domain to the spatial frequency domain and vice versa (See Figure 2). Applying the FFT is the first step of the algorithm and applying the inverse FFT is the last.

5.2 2D Fourier Transform Interpretation

In order to act on the output of the Fourier transform, it is important to interpret its data points. Each data point represents a wave function with magnitude, phase, and a defined frequency, where frequency in this case is defined as cycles per visual degree (cpd). In this manner, each data point's frequency may be directly used as input to the CSF as a basis for adjustment. We thus first measure the average amount of pixels per visual angle of the screen related to the viewer, and then use this to obtain the correct visual spatial density related to this data point. The same procedure is repeated, but for a 'desired' viewing distance perception.

The first quantity to measure is the amount of pixels per visual degree relating to the viewer. The viewer's position is taken to be at the perpendicular to the screen's center. The horizontal visual angle can be derived to equal $2arctan(\frac{w}{2d})$, where w is the screen width and d is the distance to the screen. To obtain the average amount of pixels per degree, the total amount of pixels is divided by this equation, $\frac{pixels}{2arctan(\frac{w}{d})}$. Since pixels are square, it is taken as an assumption that the amount is approximately valid for the vertical and diagonal directions as well. This is then repeated to obtain pixels per degree at the 'desired' viewing distance.

With pixels per visual angle known, the next step is to translate this information to concrete cycles per visual degrees (cpd). The frequency of an FFT data point is defined according to

$$f = \frac{i * S_r}{N} \tag{1}$$

where i is the index of this data point, S_r the sampling rate and N the total number of points. Since our sample points are the pixels themselves, and we want to measure in terms of cycles per visual degree (cpd), we can take S_r to be the number of pixels per visual degree as measured earlier.

However, the issue of obtaining i and N still remains. The main challenge in this case is that the frequency is defined in terms of one-dimensional quantities but our Fourier space has two dimensions. To overcome this, the projection-slice theorem may be used. The projection-slice theorem states that using the Radon transform to transform a 2D function into a 1D function and taking its Fourier transform is equivalent to slicing the Fourier transform through its origin with an equivalent line [Bracewell 1990] (See Figure 3). Thus, with the image in the Fourier domain and its origin at $(0, 0)$, we can compute i and N for any point (x, y).

In order to compute i, we simply use the length of the vector, $|(x, y)|$. For N, we first compute the point of intersection of the ray with the edge of the image, and the length to this point. Then, this length doubled is the total extent of this slice, N. This is given by

$$\begin{aligned} v\_intersect &= (\frac{xh}{2y}, \frac{h}{2}) \\ h\_intersect &= (\frac{w}{2}, \frac{yw}{2x}) \\ N &= 2min(|v\_intersect|, |h\_intersect|) \end{aligned} \tag{2}$$

where w and h are the width and height of the transformed image.

5.3 Magnitude Adjustment

The next step is to adjust the magnitude of the wave function represented by each data point. Towards this end, the wave function is first translated from rectangular form to polar form. Then, a compensation factor is calculated based on the CSF. Finally, this compensation factor is applied to the magnitude of the wave function and the result is stored back in rectangular form.

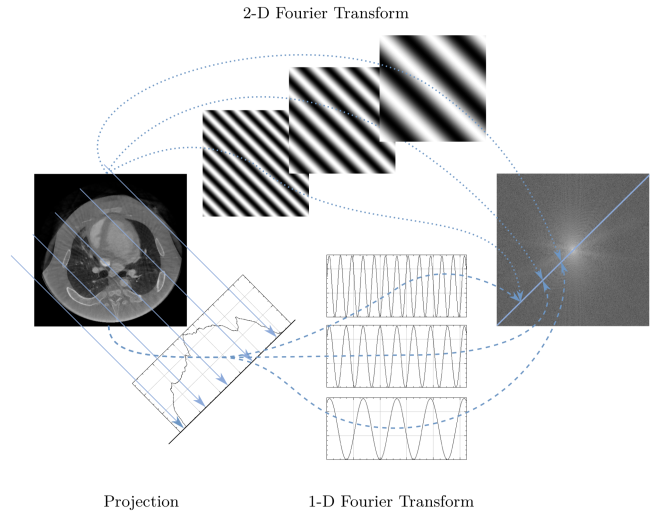


Fig. 3. Visualized projection-slice theorem [Maier et al. 2018]. A slice is taken through a 2d Fourier transform, whose power spectrum can be seen on the right. On the bottom, the frequencies corresponding to each point are represented as sinusoids with a corresponding frequency, while on the top they are represented as gratings with a corresponding direction as well as frequency.

The compensation factor is derived from the wave function and two uniform inputs, $\frac{p}{vd_{current}}$ and $\frac{p}{vd_{desired}}$, where p refers to number of pixels and vd the total visual degrees. Each of this is used in Eq. 1 to obtain $cpd_{current}$ and $cpd_{desired}$. These represent the cycles per visual degree (cpd) formed by each wave function, with $cpd_{current}$ representing the current observation scenario and $cpd_{desired}$ representing the desired observation scenario. Both of these are then fed into the CSF to obtain the compensation factor c, according to

$$c = \frac{csf(cpd_{desired})}{csf(cpd_{current})} \tag{3}$$

Finally, c is used in order to modulate the magnitude of the wave function, as expressed by $r_{new} = cr_{old}$, where r is the magnitude of the wave function. The magnitude in this case relates directly to the amplitude of the wave function. Thus, by modulating the magnitude, the distance between the largest and smallest points produced by the wave function is being changed, hence the contrast.

After this operation, the next step is to perform the inverse FFT step before producing the final image.

6 IMPLEMENTATION

The program was implemented in the Rust programming language with OpenGL compute shaders. The source code is publicly accessible [1].

[1] https://github.com/fayalalebrun/CSFContrast

6.1 Fast Fourier Transform

The Fast Fourier Transform is a family of algorithms for efficiently computing the Fourier transform [Ukidave et al. 2014]. Due to its lack of an ordering stage the Stockman variant is especially well-suited for use in GPUs [Swarztrauber 1984]. In this work, we use a Stockham radix-2 FFT implementation in order to achieve real-time performance. Specifically, this work uses an adaptation of the open-source OpenGLFFT [bane9 2022] implementation.

6.2 Runtime statistics

FFT	14.830 ms
Interpretation and adjustment	0.860 ms
Inverse FFT	10.804 ms

Table 1. Runtime costs.

CPU	AMD Ryzen 5950HX
GPU	AMD RX 6700M
RAM	16 GB
OpenGL Driver	Mesa RadeonSI 22.0.3

Table 2. System specifications.

In Table 1 the runtime costs of the algorithm running in a system with specifications as detailed in Table 2 is shown. Here it can be seen that the costs of the interpretation and adjustment phase are an order of magnitude lower compared to the other stages. Thus, in general the algorithm may be run in real-time if the FFT and inverse FFT can be run in real time, which is the case with the tested specifications.

7 RESULTS

Fig. 4. Artifacts in the first row of Figure 6. Due to shape of the Contrast Sensitivity Function, attempting to manipulate an image to make it seem as though it is much closer to the viewer makes it necessary to greatly increase the wave amplitude of the higher frequency wave functions. Thus in areas with high frequency components, particularly edges, contrast values are clipped.

Fig. 5. Practical application of the algorithm. The second image is manipulated to appear as if it is being viewed from the same distance as the first. This is assumed to be viewed from a distance of 10 cm in printed form.

Figure 6 shows the results of applying the algorithm to a static grayscale 1920x1080 image with several parameters. The current distance represents the actual distance of the viewer to the screen, while the target distance represents the perception which should be emulated. We use the HDR-VDP-2 [Narwaria et al. 2015] algorithm in order to evaluate the results quantitatively. This metric compares a reference image to an image viewed at a certain distance, and computes the difference. In this case, each manipulated image is compared to the reference image being viewed from the desired distance. In other words, we use our manipulated image as the reference image in the metric's terms. Due to our objective of achieving similar perception to the image at this desired distance, in this case it is desirable to have the least amount of differences as possible.

When the desired distance is significantly less than the actual distance, artifacts may arise. This is related to the shape of the Contrast Sensitivity Function, as in this case generally wave functions have to compensate by increasing their magnitude, which can result in values exceeding the image's maximum value. The artifacts present in the first row of Figure 6 are showcased in Figure 4.

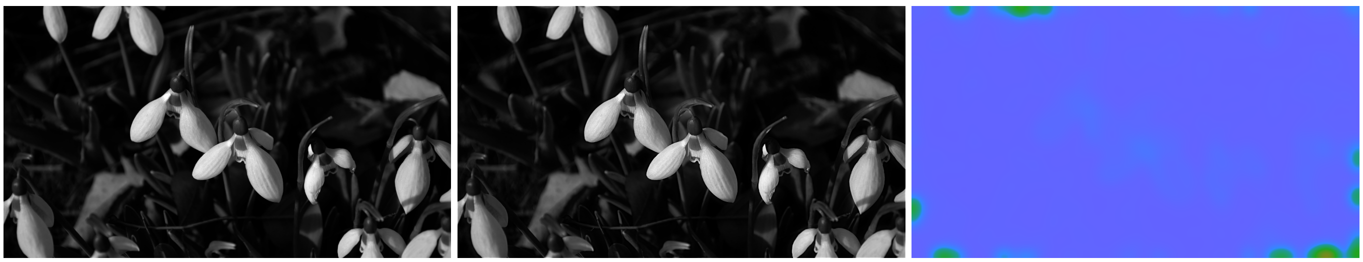
Figure 7 shows the algorithm being applied to a variety of images with different parameters.

Figure 5 illustrates the algorithm being used in practice. Real applications could also be found in theaters, signage screens, or anywhere else where predictable perception of an image is important.

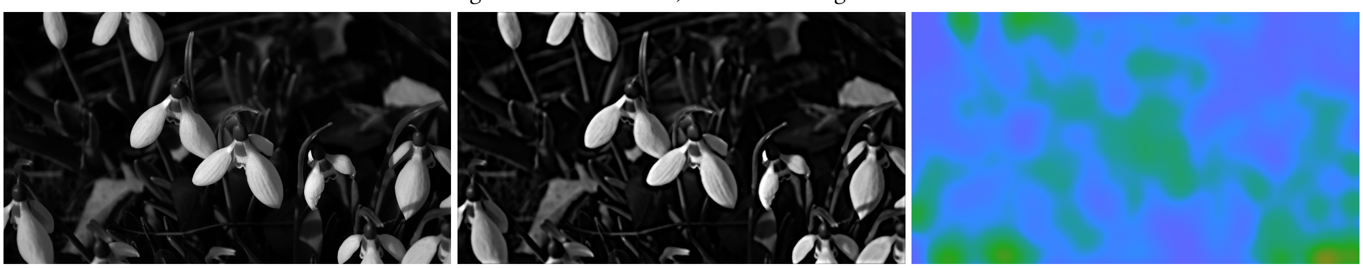
Several works have previously used the Contrast Sensitivity Function (CSF) for image processing [Nadenau et al. 2003] [Majumder and Irani 2006]. However, these use properties of the CSF, and do not use the CSF itself. In contrast, this work utilizes values from a pre-defined CSF. This means results could potentially be catered to the viewing characteristics of a single person.

Traditional media is subject to the effects of the CSF. That is to say, perception of contrast changes depending on distance to the

Actual viewing distance: 112.53 cm; Desired viewing distance: 28.49 cm



Actual viewing distance: 170.06 cm; Desired viewing distance: 88.19 cm



Actual viewing distance: 24.83 cm; Desired viewing distance: 140.29 cm

Actual viewing distance: 24.83 cm; Desired viewing distance: 424.68 cm

Fig. 6. The algorithm is applied to an image with several different parameters. The images from left to right represent the reference image, the processed image, and the visible difference using HDR-VDP-2 [Narwaria et al. 2015] set to a viewing distance equal to the desired distance. For the visible difference, green represents minor differences, and red represents major differences.

medium. This work proposes a real-time algorithm for compensating for changes in contrast perception due to distance.

8 CONCLUSION AND FUTURE WORK

Changes in contrast perception due to distance can pose an important challenge for media. In this paper we presented a novel approach for achieving perceptual consistency of contrast across distance for a medium. We provide a real-time open-source program for performing this algorithm on arbitrary images and videos. With this type of tool, media creators can better ensure their content is presented as they intend.

In terms of future work, the use of wavelets could be interesting due to their localized nature. Wavelets are defined in both frequency and space. This would make it possible to define spatial frequency depending on the location on-screen, which could create a better compensation effect depending on the circumstances. Of particular interest are Gabor wavelets since they resemble the Gabor patches used to measure the CSF [Whittle 1986]. A variant of the Gabor

Actual viewing distance: 75.00 cm; Desired viewing distance: 1000.00 cm

Actual viewing distance: 4000.00 cm; Desired viewing distance: 3000.00 cm

Actual viewing distance: 4000.00 cm; Desired viewing distance: 2300.00 cm

Fig. 7. The algorithm is applied to several images. From left to right is the reference image and modified image. Below each pair, the parameters for the algorithm are listed.

wavelet, the Log-Gabor wavelet, could be better suited for fast decomposition to its orthogonality [Fischer et al. 2007].

9 RESPONSIBLE RESEARCH

The source code for the proposed algorithm is publicly available [2]. All reference images used can be obtained there and parameters are provided within the paper.

[2]https://github.com/fayalalebrun/CSFContrast

REFERENCES

bane9. 2022. OpenGLFFT. https://github.com/bane9/OpenGLFFT original-date: 2020-10-09T13:37:53Z.

R. N. Bracewell. 1990. Numerical Transforms. *Science* 248, 4956 (May 1990), 697–704. https://doi.org/10.1126/science.248.4956.697 Publisher: American Association for the Advancement of Science.

Dominic Canare. 2021. QuickCSF. https://github.com/vpclab/QuickCSF original-date: 2019-04-13T20:42:51Z.

Thom Carney, Christopher W. Tyler, Andrew B. Watson, Walter Makous, Brent Beutter, Chien-Chung Chen, Anthony M. Norcia, and Stanley A. Klein. 2000. Modelfest: year one results and plans for future years, Bernice E. Rogowitz and Thrasyvoulos N. Pappas (Eds.). San Jose, CA, 140–151. https://doi.org/10.1117/12.387150

Sylvain Fischer, Filip Šroubek, Laurent Perrinet, Rafael Redondo, and Gabriel Cristóbal. 2007. Self-Invertible 2D Log-Gabor Wavelets. *International Journal of Computer Vision* 75, 2 (Nov. 2007), 231–246. https://doi.org/10.1007/s11263-006-0026-8

Andreas Maier, Stefan Steidl, Vincent Christlein, and Joachim Hornegger. 2018. *Medical Imaging Systems*. https://link-springer-com.tudelft.idm.oclc.org/book/10.1007/978-3-319-96520-8

Aditi Majumder and Sandy Irani. 2006. Contrast enhancement of images using human contrast sensitivity. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization (APGV '06)*. Association for Computing Machinery, New York, NY, USA, 69–76. https://doi.org/10.1145/1140491.1140506

Rafal Mantiuk, Scott J. Daly, Karol Myszkowski, and Hans-Peter Seidel. 2005. Predicting visible differences in high dynamic range images: model and its calibration. In *Human Vision and Electronic Imaging X*, Vol. 5666. SPIE, 204–214. https://doi.org/10.1117/12.586757

M.J. Nadenau, J. Reichel, and M. Kunt. 2003. Wavelet-based color image compression: exploiting the contrast sensitivity function. *IEEE Transactions on Image Processing* 12, 1 (Jan. 2003), 58–70. https://doi.org/10.1109/TIP.2002.807358 Conference Name: IEEE Transactions on Image Processing.

Manish Narwaria, Rafal Mantiuk, Matthieu Perreira Da Silva, and Patrick Le Callet. 2015. HDR-VDP-2.2: A calibrated method for objective quality prediction of high-dynamic range and standard images. *Journal of Electronic Imaging* 24 (Jan. 2015), 010501. https://doi.org/10.1117/1.JEI.24.1.010501

Aude Oliva, Antonio Torralba, and Philippe Schyns. 2006. Hybrid images. *ACM Trans. Graph.* 25 (July 2006), 527–532. https://doi.org/10.1145/1141911.1141919

Krisztián Samu, Klara Wenzel, and Karoly Ladunga. 2002. Color and luminance contrast sensitivity function of people with anomalous color vision. *Proc SPIE* (June 2002). https://doi.org/10.1117/12.464652

Paul N. Swarztrauber. 1984. FFT algorithms for vector computers. *Parallel Comput.* 1, 1 (Aug. 1984), 45–63. https://doi.org/10.1016/S0167-8191(84)90413-7

Yash Ukidave, Gunar Schirner, and David Kaeli. 2014. Fast Fourier Transform (FFT) on GPUs. In *Numerical Computations with GPUs*, Volodymyr Kindratenko (Ed.). Springer International Publishing, Cham, 339–361. https://doi.org/10.1007/978-3-319-06548-9_16

Paul Whittle. 1986. Increments and decrements: Luminance discrimination. *Vision Research* 26, 10 (Jan. 1986), 1677–1691. https://doi.org/10.1016/0042-6989(86)90055-6

Krzysztof Wolski, Daniele Giunchi, Nanyang Ye, Piotr Didyk, Karol Myszkowski, Radosław Mantiuk, Hans-Peter Seidel, Anthony Steed, and Rafał K. Mantiuk. 2018. Dataset and Metrics for Predicting Local Visible Differences. *ACM Transactions on Graphics* 37, 5 (Nov. 2018), 172:1–172:14. https://doi.org/10.1145/3196493

X. Zhang and B. A. Wandell. 1997. A spatial extension of CIELAB for digital color-image reproduction. *Journal of the Society for Information Display* 5, 1 (1997), 61–63. https://doi.org/10.1889/1.1985127 _eprint: https://sid.onlinelibrary.wiley.com/doi/pdf/10.1889/1.1985127.