# Image Processing
# for Improved Perception and Interaction

**Author** **Ing. Michal Seeman**

**Supervisor** **Doc. Dr. Ing. PAVEL ZEMČÍK**

**2012**

# Author's c.v.

Born: 1979-04-12 in Brno
Education

- 1994 - 1998 Gymnázium Vídeňská, Brno, informatics branch
- 1998 - 2005 FIT BUT Brno, Master's Degree Study
- 2005 - 2012 FIT BUT Brno, Ph.D. Degree Study

Activity

- 1996 - 1997 Ext. teaching of programming, CVČ Lužánky Brno
- 1997 Unofficial Czech programming mastership below age of 18 years, 11. place
- 2001 - 2006 Mobile applications development, OS Epoc (Symbian)
- 2006 - 2011 Project Biomarker, FIT BUT

# Abstract

Image reproduction ought to provide subjective sensation possibly closest to the one, where the original image is observed. Digital image reproduction involves image capture, image processing and rendering. Several techniques in this process are not ideal. This work proposes improvement of speed and accuracy of some state-of-the-art methods.

# Introduction

Digital image reproduction involves mainly image capture and image rendering. Between these two techniques, the data are digitally processed.

Meaning of the image processing might seem to be insignificant. In fact, if the image had been captured by an ideal camera and rendered via an ideal display device, no data processing would be necessary for perfect reproduction. Unfortunately, the available devices are certainly not ideal.

The scanning devices suffer of geometry distortion, luminance non-linearity and limited contrast (dynamic range). Although all of these imperfections has been overcome, in some of the cases it is at certain price. High dynamic range can be captured by multi-exposure, which does not allow for taking photographs of non static objects or capturing of motion pictures. Geometry correction can be measured and corrected, but the algorithms are rather slow for real-time processing.

Commonly used LCD display devices have pixel matrix fixed by construction so they do not suffer from geometry distortion. But the pixel density is still low, the matrix is visible and causes disturbing artifacts. Either the highest displayable contrast is still limiting. Despite of the marketing claims, most of the common displays are not capable of rendering much higher contrast than 1:1000. The ordinary practice is to scale range of the digital image to fit the display range. The procedure is so frequent, that many users do not consider it as an image processing at all. Yet the image is certainly changed. And as will be explained further in this work, substantial change in the contrast causes noticeable change in the color perception.

In some cases, the user demands to alter the image instead of perfect reproduction. The operation is not easy, if the manipulations are to be precise. Human perception of the luminance scale is neither simple nor linear. The essentials of the visual perception and its aspects to the practical image manipulation are also explained in detail.

# 1 Human Vision Physiology

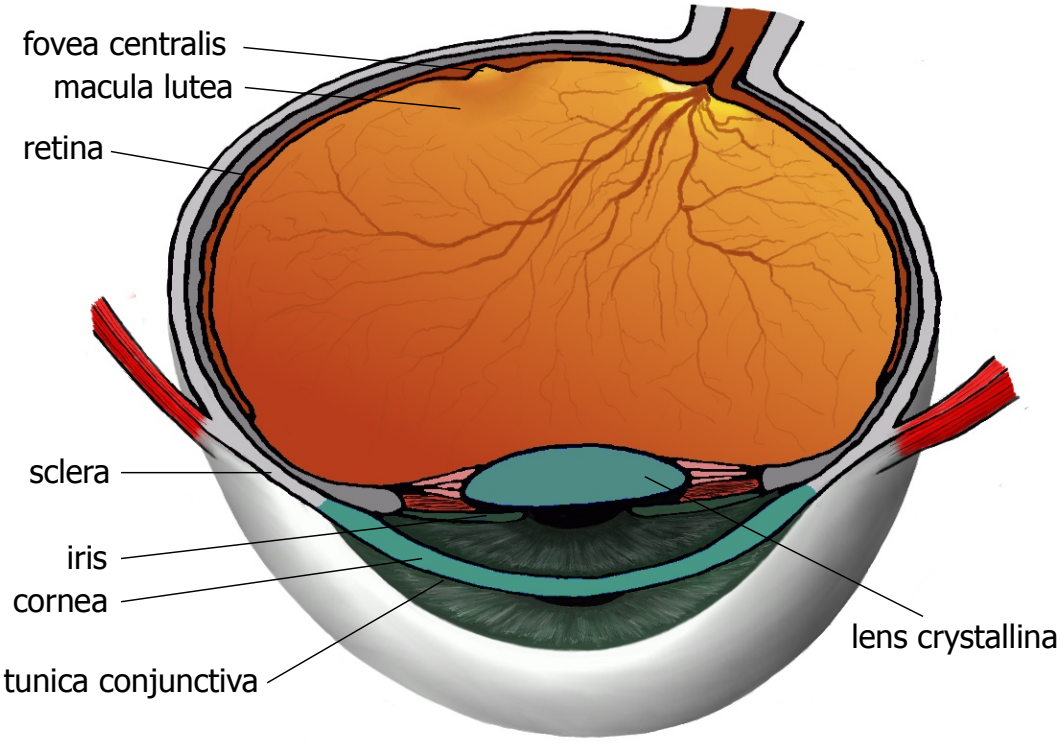


*Figure 1: Transverse section of the eye, drawing by M. Seeman*

As primates, we are equipped with the best kind of light sensitive organs in Animalia. Primate eye is equipped by complex optical part with several environments of different refractive indices, enables for accommodation, luminous flux control and three axis individual rotation.

## 1.1 Retina

The retina is highly vascularized layer of tissue, approximately 0.4mm thick. It contains photoreceptors and neural cells. The photoreceptors respond to the light stimuli and their signal is processed by several tens of specialized cell types [14]. Retina is full of neurons and therefore it is sometimes considered as a part of brain.

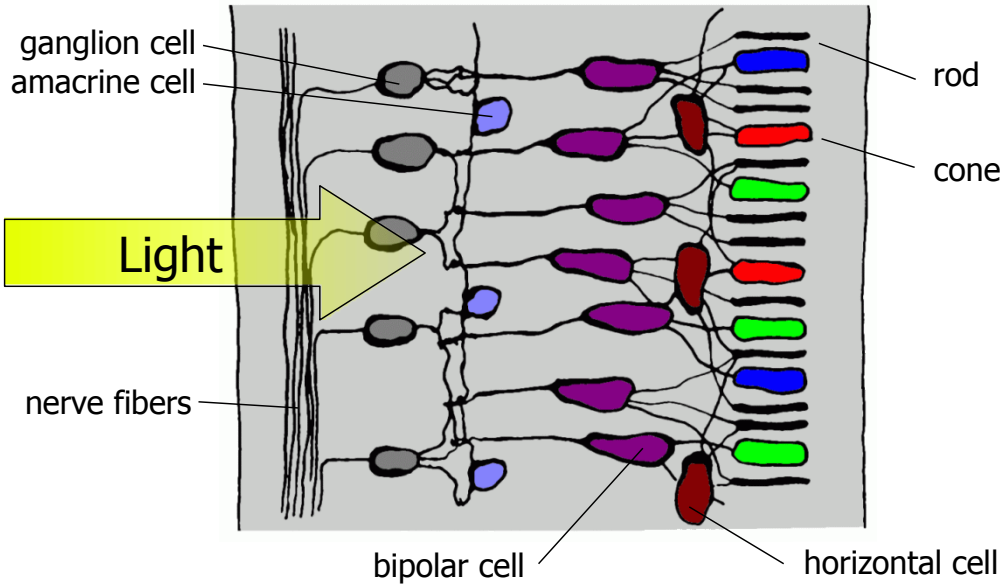


*Figure 2: Basic scheme of retina neural connections*

## 1.2 Photoreceptors

Two kinds of photoreceptors have been described, rods and cones. Rods are more sensitive and take part in nocturnal vision. Less sensitive cones work in daylight. Three types, L-, M- and S-cones in human retina are sensitive to different spectrum bands.

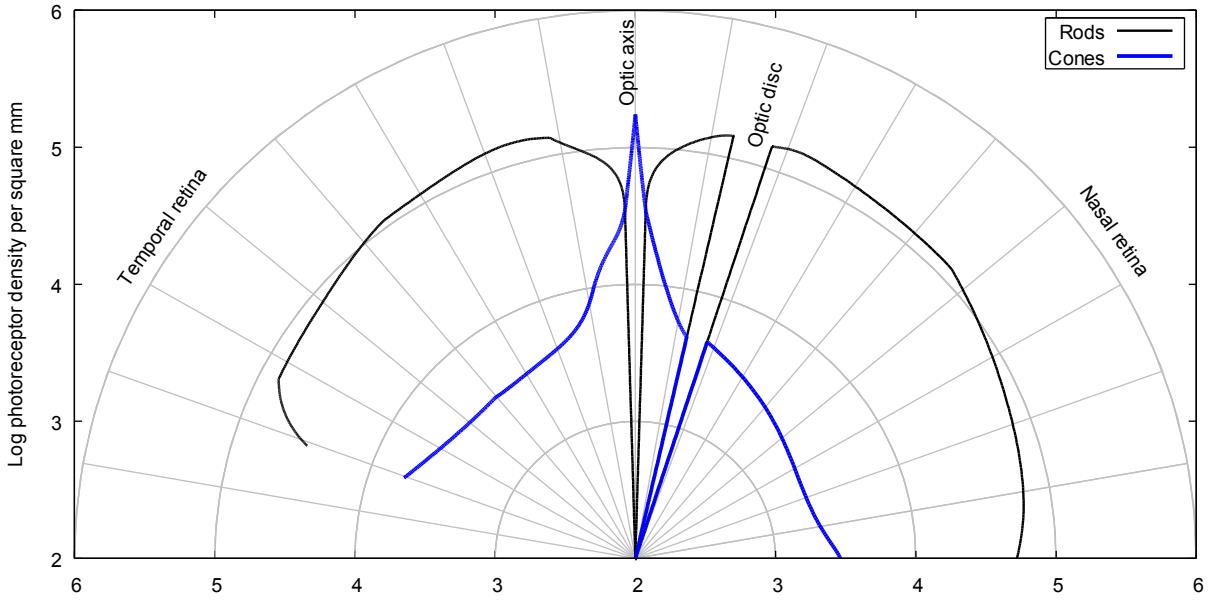


*Figure 3: Photoreceptor density across transverse section [9]*

The photoreceptors are not spread uniformly across the retina (Figure 3).

The very center of the retina is a specialized area called fovea. The fovea provides the finest vision and we use it to examine details.

## 1.3 Cell Signals

The retina neural cells form a complex structure. The essential principle of the human vision spatial characteristics is the signal fusion in horizontal and bipolar cells.

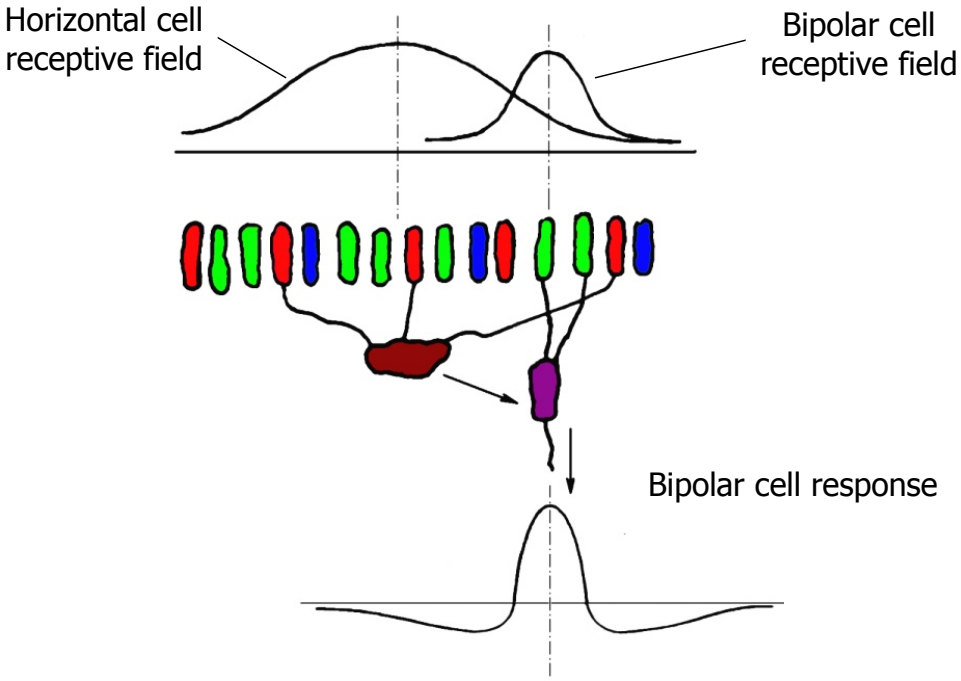


*Figure 4: Bipolar and horizontal cells spatial responses*

The result is a difference between two Gaussian-like responses (see Figure 4). The plot is similar to the common high-pass filter response.

## 1.4 Color Perception

The brightness information is passed separately and the color channel acuity is approximately one third of the brightness channel acuity [18]. The color perception dependency on the luminance is described as the Bezold-Brücke phenomenon [15] (see Figure 5).

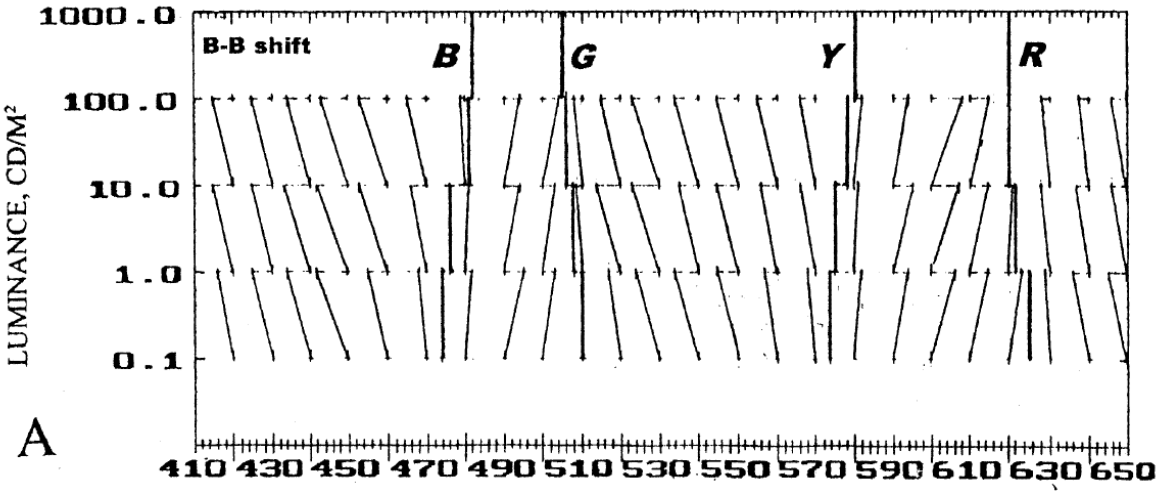


*Figure 5: Bezold-Brücke Hue-Shift [15] in the extended wavelength scale [23] reprinted with kind permission of Dr. Pridmore*

# 2 Perception Optimizing on a Display

The following scheme describes the contributions to the image processing process presented in this work.

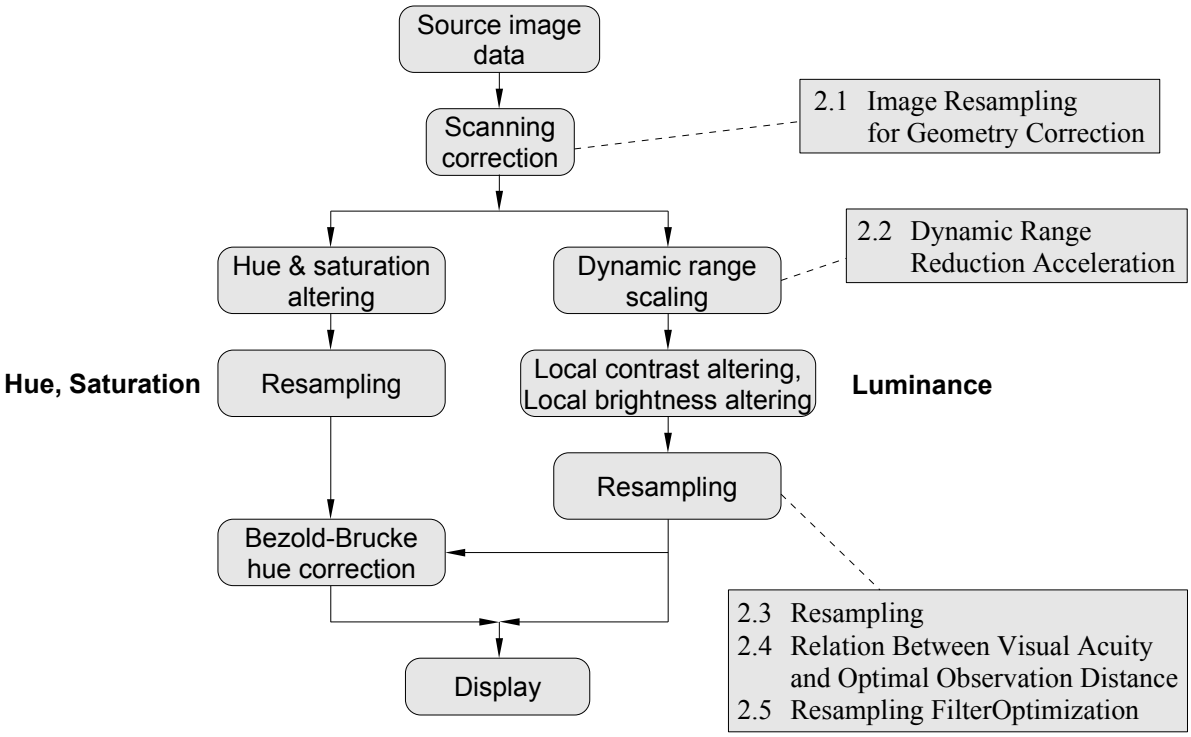


*Figure 6: General framework for image reproducing and manipulation*

## 2.1 Image Resampling for Geometry Correction

The geometrical distortion may be unacceptable in some applications. Therefore it is desirable to acquire geometrically correct image. The presented algorithm helps in correcting such images. The algorithm provides high performance at the price of certain limits. The displacement and rotation should stay in some constraints.

The algorithm exploits separable resampling via FIR filter bank. The set of filters is for selection of the subpixel displacement.

This approach enables for implementation using a pipeline with low consumption of resources in a programmable hardware. Although the implementation proposed in the presented approach is simple, it preserves the image, as well as the more complex implementations of the filters given the constraints of the approach are respected.

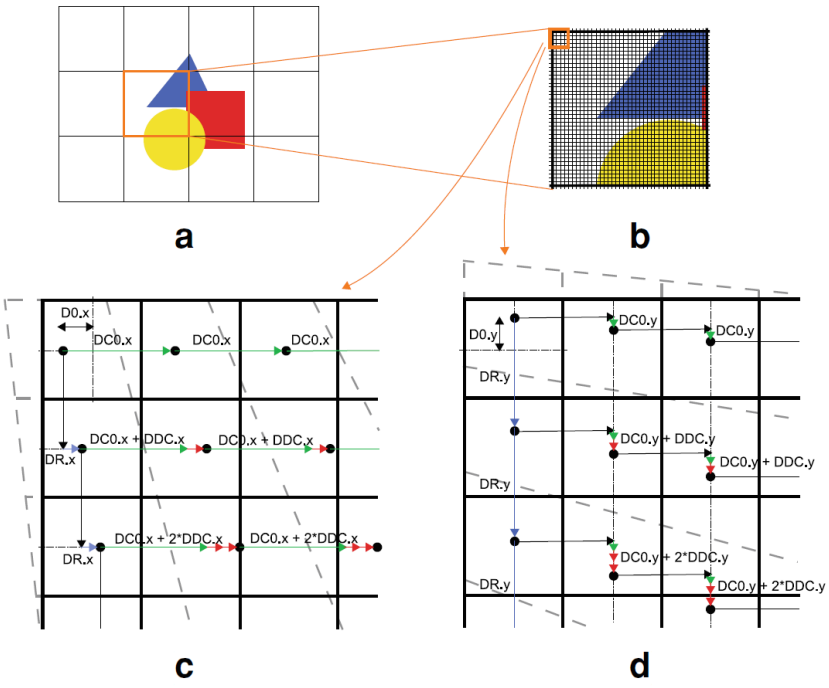


*Figure 7: Displacement interpolation in squares. Pixels of original distorted image are plotted with gray dashed line, pixels of output image are plotted with black solid line. Meaning of precalculated coefficients is marked with coloured vectors*

## 2.2 Dynamic Range Reduction Acceleration

The tone mapping operators for dynamic range reduction has been rapidly improved during last decade. One of the most complex physiologically influenced methods is [5]. This method (and many others, e.g. [6]) uses the bilateral filter for computing of the light adaptation. The filter is a bottle-neck in fast image processing. Though several attempts were made to accelerate the filtering, in 2011 we designed an approximation method with very small error and fastest computation so far. The method [7] is presented here.

Bilateral filtering is a nonlinear filtering computed as a weighted average of each pixel's surrounding. The weight is based on the spatial distance and the intensity difference. In most cases, the maximum weight is centered at zero differences of position and intensity.

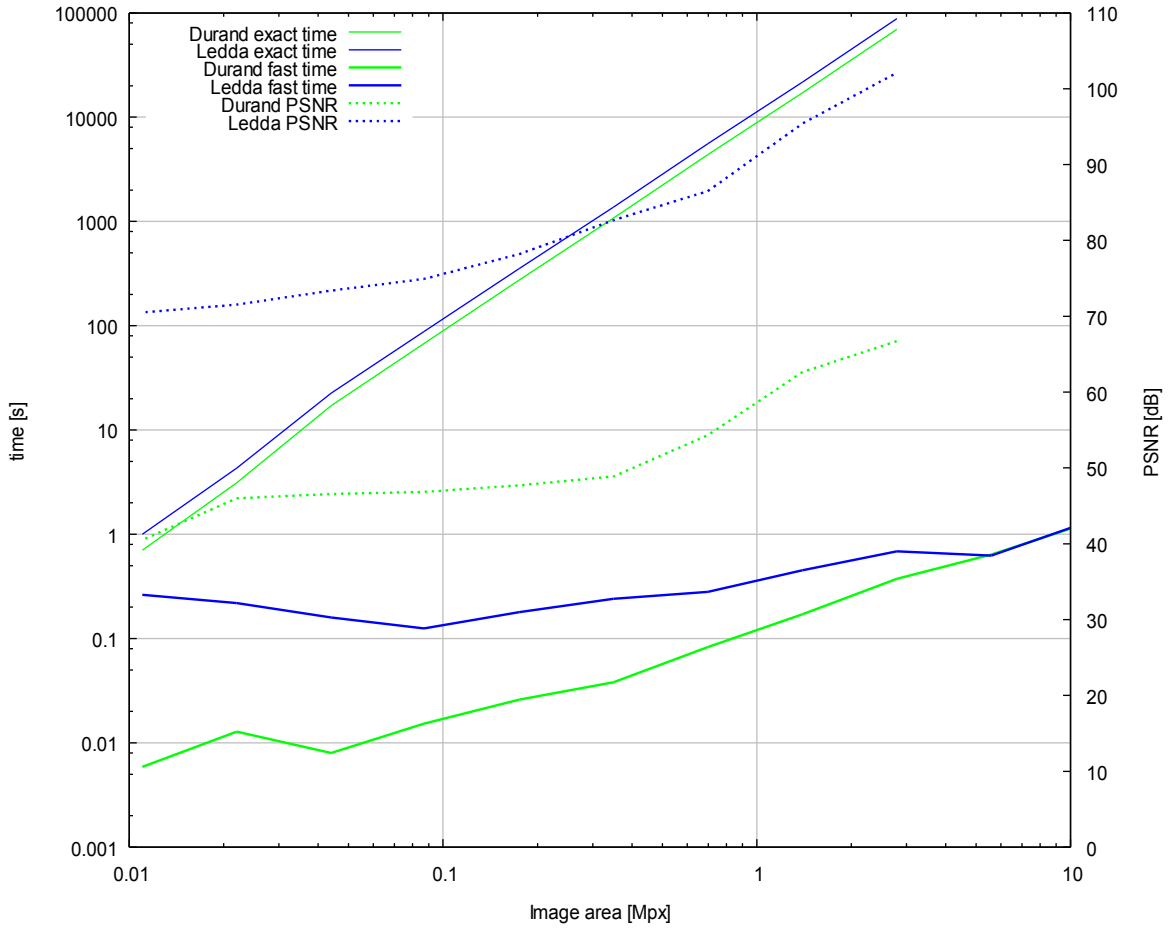


*Figure 8: Dependency of a computation time and PSNR on the image area. Time of the exact bilateral filter computation (exact), time of the accelerated algorithm (fast) –Durand, triple EMA, bilinear tile filtering and Ledda, single EMA, bilinear tile filtering and PSNR for both filter settings.*

1. The image is split into tiles. Two different histograms are computed for each tile: histogram of the pixel intensity values and the same histogram where each count is multiplied by the intensity.
2. The histograms are convolved with a function close to intensity domain Gaussian $G_{\sigma i}$.
3. A spatial filter close to convolution with a space domain Gaussian $G_{\sigma s}$ is applied to the histograms. It means that the signal value is spread among the histograms in space, but not across each of the histograms.
4. The result image value is computed as the two histograms value ratio. An interpolation has to be applied.

The filter results were compared to the exact bilateral filter implementation limited to the radius $5\sigma$. The precision was measured on twenty-nine different images, each with the area approximately 0.7megapixels. The intensity sigma was set according to two state-of-the-art approaches [5, 6]. PSNR did not drop below 43dB for $\sigma$i = 4dB [6] or below 69dB for $\sigma$i = 0.6dB [5]. While the exact bilateral filter time dependency on the image area is almost exactly quadratic, in the approximation method the dependency is close to linear.

*a* *b*

*Figure 9: (a) Tone-mapped input image, (d) differential image (Ledda's σ, single EMA[5])*

## 2.3 Resampling

It has to be said, that ideal resampling is not necessarily best for display devices. The reasons are:

1. Display pixels are not ideal samples. Ideal sample would be close to Gaussian or Sinc function with very low frequency domain response above half of the sampling frequency.
2. In the HVS there is no lo-pass filter, which would have suppressed high frequency harmonic signal caused by inadequate samples. Or to be more specific, there is a filter suppressing high frequencies in HVS, but the inhibition is not very steep and it depends on the observation distance.

The resampling should respect shape of the display pixel, spatial response of the visual system and observation distance. These problems are discussed on the following pages.

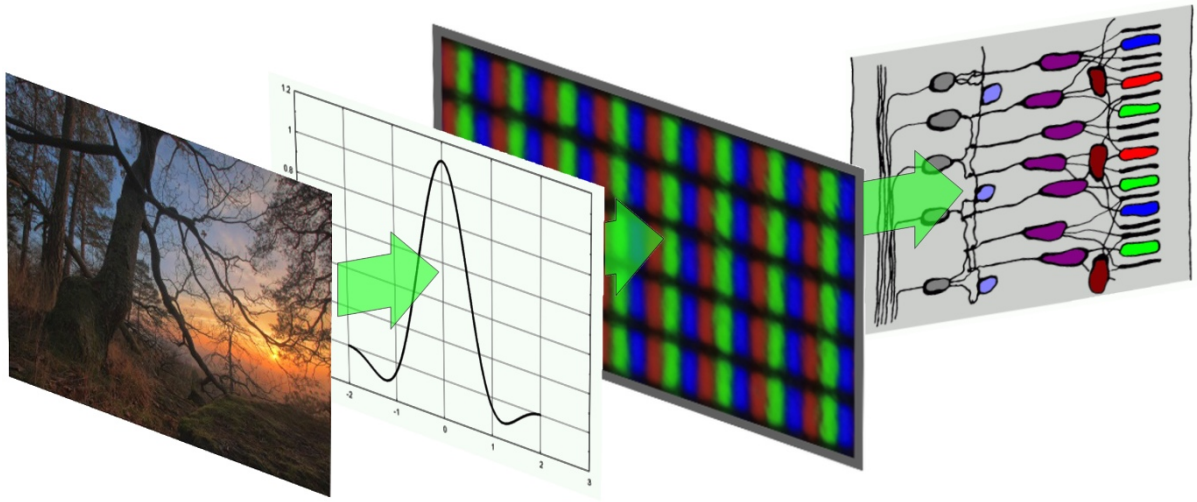

*Figure 10: Perception of an image on a display: Image is filtered and sampled, rendered via the display pixels and processed by the HSV*

## 2.4 Relation Between Visual Acuity and Optimal Observation Distance

Users tend to view the display from the so-called "comfortable distance". The question is how does the comfortable observation distance correspond to the visual acuity. The proposed approach was used to measure correlation between optimal observation distance from the display device and the visual acuity.

Users were to compare image post-processing methods. Tiny differences forced them to carefully choose optimal distance. The distance was then measured by triangulation. Visual acuity was measured at the display surface, so the accommodation conditions were comparable.

A standard monitor with the pixel spacing 0.270mm was used. The correlation was measured on 20 subjects.

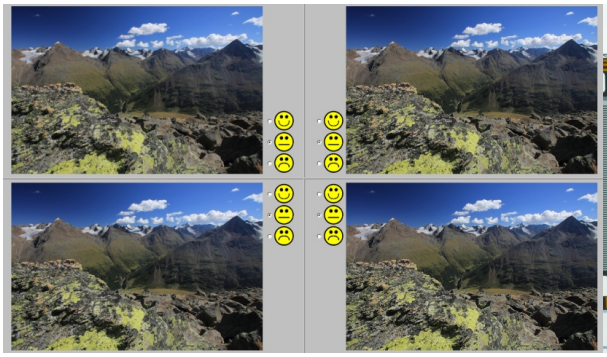

*Figure 11: Testing screen*

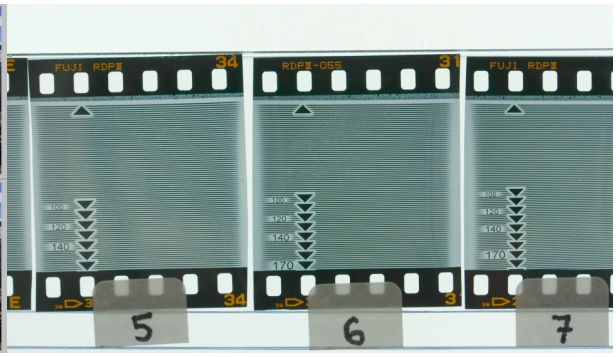


*Figure 12: Optotypes*

| | |
|---|---|
| Angular acuity | 32.3 cycles·deg$^{-1}$ |
| Ang. ac. deviation | 0.118 $\log_{10}$ cycles·deg$^{-1}$ |
| Relative acuity | 3.98 cycles·mm$^{-1}$ |
| Rel. ac. deviation | 0.024 $\log_{10}$ cycles·mm$^{-1}$ |

Although the acuity varies, it shows strong correlation with the preferred observing distance (Figure 13 and Figure 14). The results show that the relative spatial acuity in preferred distance has much smaller deviation than the angular acuity.

The statistics were used to project the retina cell receptive field to the display plane.

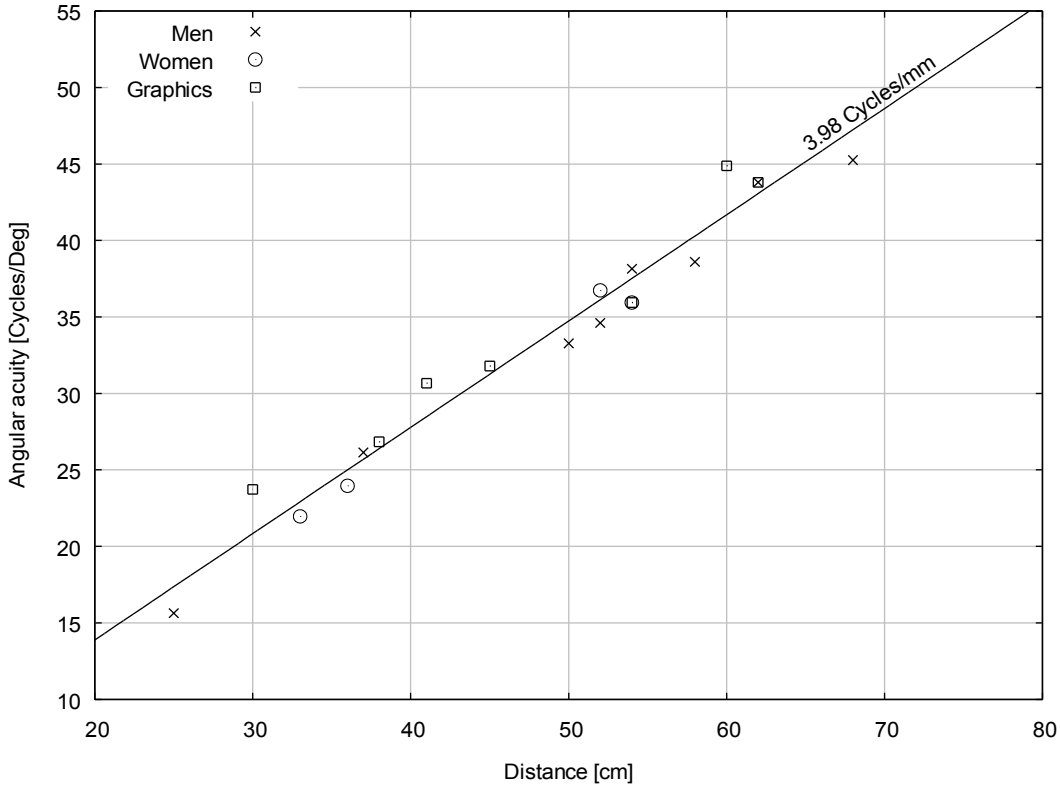


*Figure 13: Correlation between angular acuity and preferred observing distance*

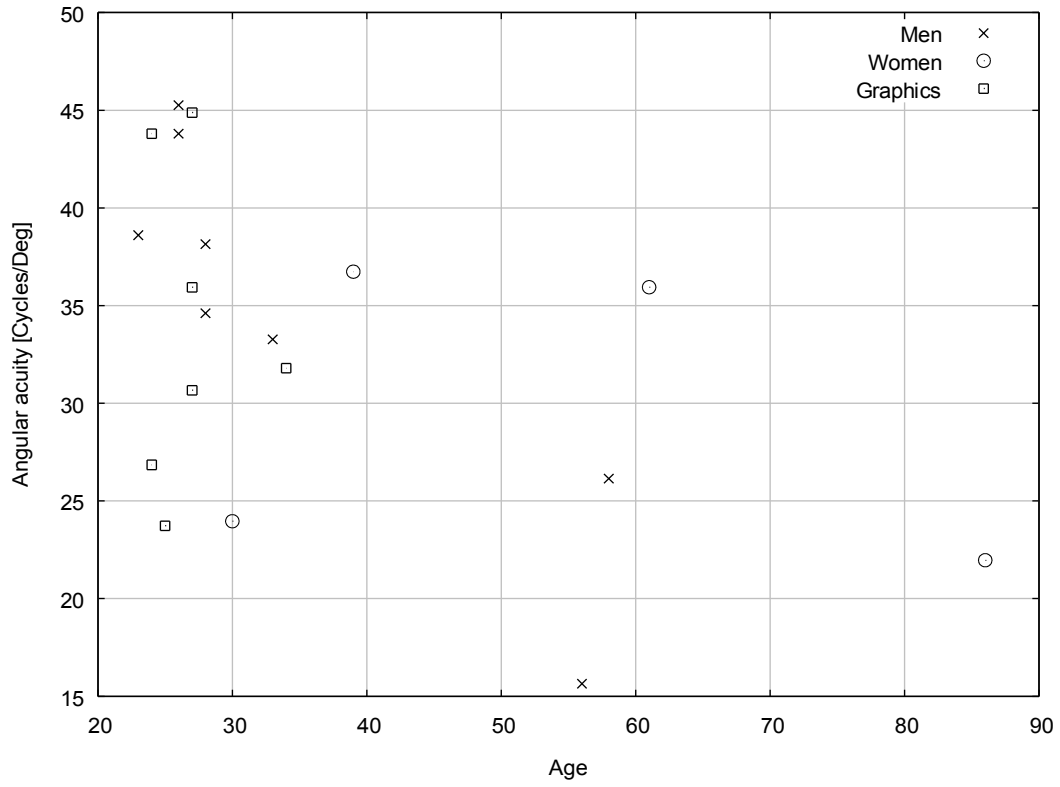


*Figure 14: Distribution of angular acuity across the age*

## 2.5 Resampling Filter Optimization

When the image is post-processed, rendered on a display and observed, the whole process can be described as shown in Figure 15.

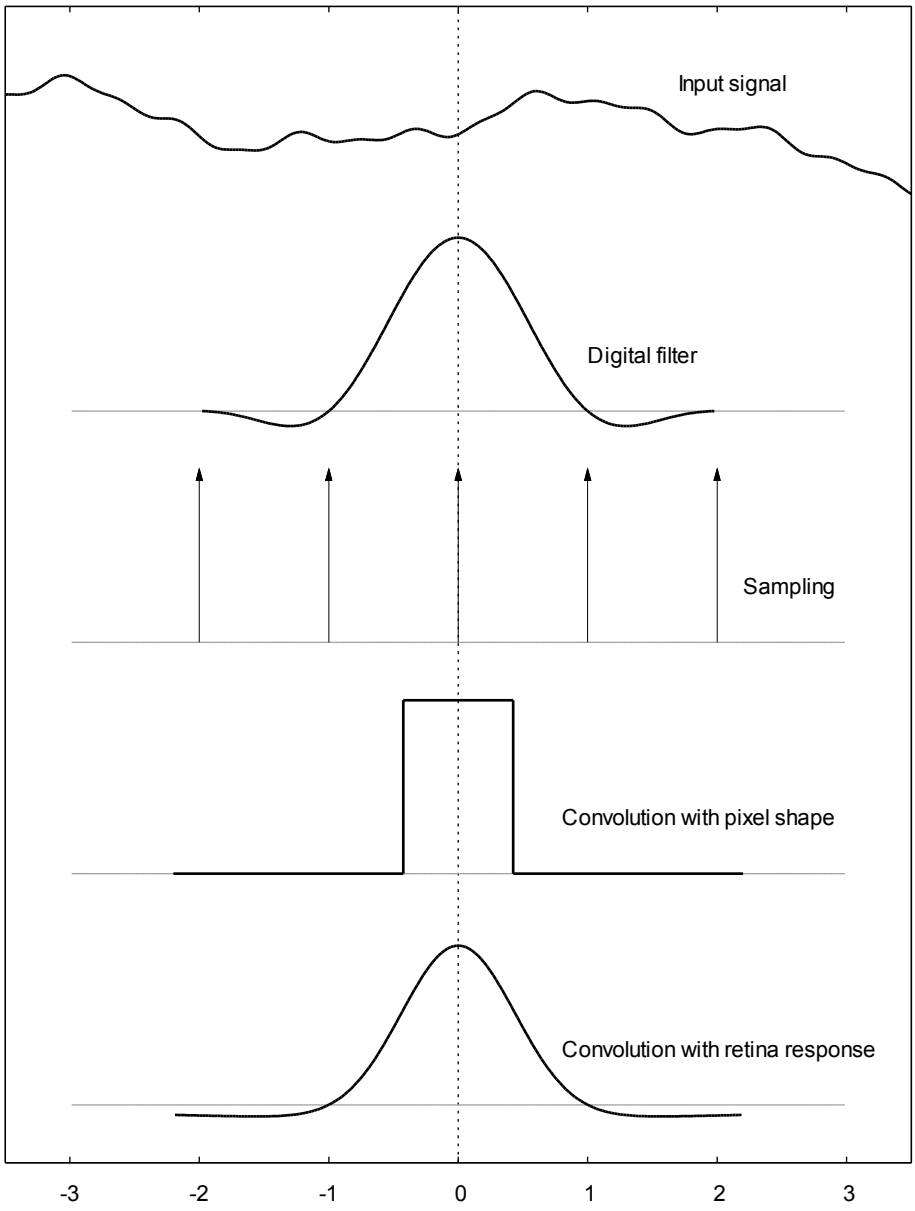


*Figure 15: Scheme of the post-processing and observing process*

The best filter would give the same result as direct observing. However this is not possible due to the loss of information by sampling. The filter can by only optimized. Unfortunately the optimal *Filter* with minimal error can not be generally expressed.

The equation contains both convolution and multiplying, so minimization can not be solved in spatial nor in frequency domain. But the problem could be split into different sub-pixel positions *s*. The result needs to be expressed as a convolution for any case of s. Each case gives a different *Kernel*, so the complete operation is not convolution. But later we can minimize the error across all kernels.

The space of all Filters has to be searched by brute-force algorithm. For this purpose the space has to be reduced reasonably. Following method was used:

- *Filter* is designed via Fourier transform. The highest harmonic frequency is not above the spatial frequency recognizable by the HVS. The amplitudes are not complex numbers. Filter should be symmetrical (even function), so only cosine harmonics are contained. This reduces the parameters to a relatively small amount of numbers.
- The parameter quantization step was selected as 1/1000. For the 8-bit displays the precision is sufficient. The *Eye* absolute values below $10^{-4}$ were ignored.
- Filter area (integral) should be 1. It gives that the zero harmonic value $F_0$ is inverse of the Filter size.
- The Filter value at both ends should be 0. It gives that $F_0-F_1+F_2-F_3+ \ldots = 0$
- The spatial domain quantization was 11 samples per pixel. It is dense enough, so that the highest sampled frequency according to Nyquist is well beyond HVS recognition and odd number made some of the numeric computations easier.

The error minimization process was run twice, for the *Filter* of radius 2 and 3 pixels.

**Optimization Result**

The results are compared in Figure 16. It is clear that the function is close to zero from the eccentricity about 2 pixels.

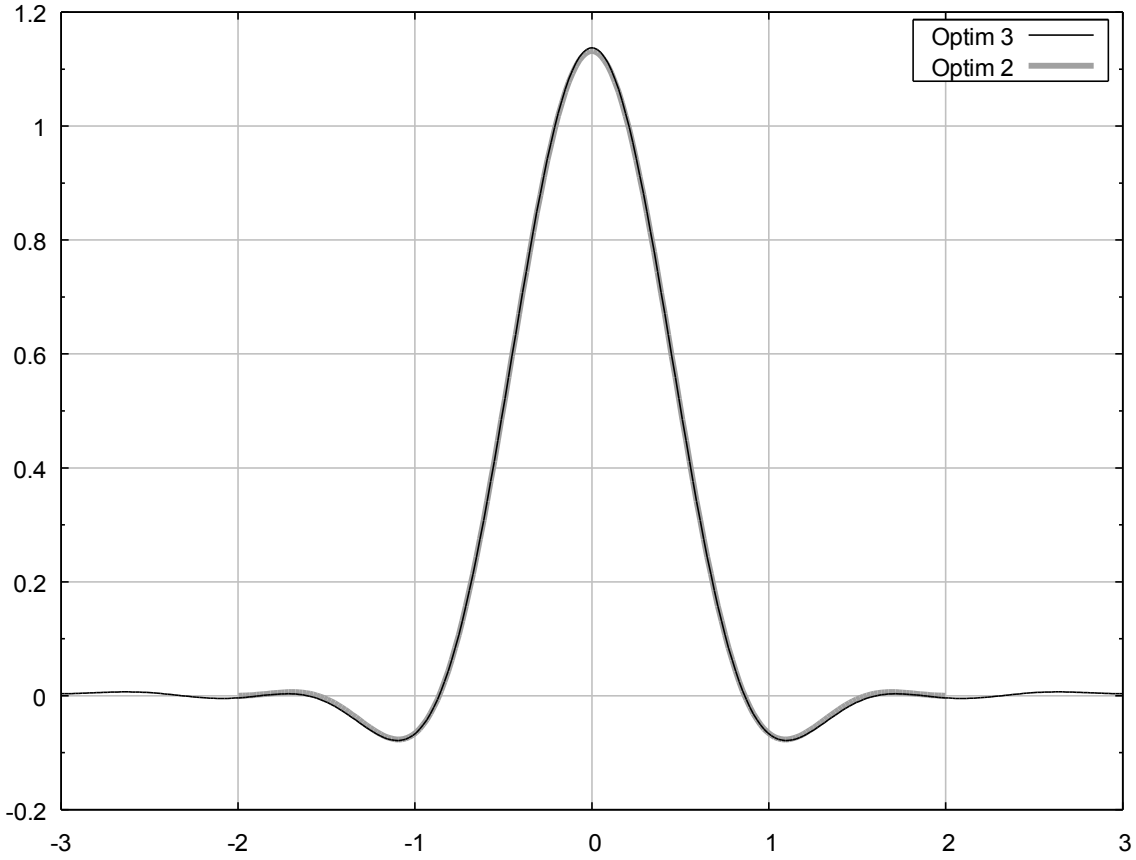


*Figure 16: Optimization results for different filter sizes*

The amplitudes of the filter with 2 pixels radius given by the optimization are listed below:

$$\begin{aligned} &Filter(x) = \\ &1/4 + \\ &0.504 \cdot \cos(1 \cdot \pi/2) + \\ &0.302 \cdot \cos(2 \cdot \pi/2) + \\ &0.048 \cdot \cos(3 \cdot \pi/2) \end{aligned} \tag{2.1}$$

# 3 Conclusions

The reconstruction of an image via digital display is a complex problem. The result is still not ideal with commonly available devices and current methods. The aim of this work was to identify the weak points in the whole system and improve them. Some of the methods already provide transfer with error undetectable by human vision, but the processing is too slow for interactive view or real-time video processing. Several improvements were achieved and described in this work. The framework with proposed changes enhances both performance and perceived image quality.

# References


[1] Croner, L.J., Kaplan, E.: Receptive Fields of P and M Ganglion Cells Across the Primate Retina, Vision Res. Vol. 35, No. I, pp. 7 24, 1995

[2] Purpura, K., Kaplan, E., Shapley, R.M.: Background Light and the Contrast Gain of Primate P and M Retinal Ganglion Cells, Proc. Natl. Acad. Sci., Vol. 85, pp. 4534-4537, Jun 1988

[3] Schwartz, S.H.: Visual Perception: a Clinical Orientation, 3rd ed., McGraw-Hill, ISBN: 0-07-141187-9, 2004

[5] Ledda, P., Santos, L.P., Chalmers, A.: A Local Model of Eye Adaptation for High Dynamic Range Images, Proc. of the International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction, AFRIGRAPH '04, pp. 151-160, ISBN: 1-58113-863-6, 2004

[6] Durand, F., Dorsey, J.: Fast Bilateral Filtering for the Display of High-Dynamic Range Images, Proc. of the Conference on Computer Graphics and Interactive Techniques, USA, pp. 257-266, ISSN: 0730-0301, 2002

[7] Seeman, M., Zemčík, P., Juránek, R., Herout, A.: Fast Bilateral Filter for HDR Imaging, J. Vis. Commun. Image R., Vol. 23, pp. 12-17, 2012

[8] Ferwerda, J.A.: Elements of Early Vision for Computer Graphics, IEEE Computer Graphics and Applications, pp. 22-33, ISSN: 0272-1716, Sep-Oct 2001

[9] Osterberg, G.: Topography of the layer of rods and cones in the human retina, Acta Ophthal Vol. 6, pp. 1-103, 1935

[10] Peichl, L., Wassle, H.: Size, Scatter and Coverage of Ganglion Cell Receptive Field Centres in the Cat Retina, J. Physiol. 291, pp. 117-141, 1979

[11] Hammond, P.: Cat Retinal Ganglion Cells: Size And Shape of Receptive Field Centres, J. Physiol. 242, pp. 99-118, 1974

[12] Dacey, D., Packer, O. S., Diller, L., Brainard, D., Peterson, B., Lee, B.: Center Surround Receptive Field Structure of Cone Bipolar Cells in Primate Retina, Vision Res. Vol. 40, Issue 14, pp. 1801-1811, Jun 2000

[13] Nelson, R., Kolb, H., Freed, M.: Off-Alpha and Off-Beta Ganglion Cells in Cat Retina I: Intracellular Electrophysiology and HRP Stains. J. Comp. Neurol. 329, pp. 68-84, 1993

[14] Kolb, H.: How the Retina Works, Am. Scientist, 91, pp. 28-35, Jan-Feb 2003

[15] Pridmore, R.W.: Bezold-Brucke Hue-Shift as Functions of Luminance Level, Luminance Ratio, Interstimulus Interval and Adapting White for Aperture and Object Colors, Vis. Research, Vol. 39, Issue 23, pp. 3873-3891, Nov 1999

[16] Lamming D.: Spatial Frequency Channels. Chapter 8. In: Cronly-Dillon, J., Vision and Visual Dysfunction, Vol 5., 1991

[17] Van Nes F.L., Bouman, M.A.: Spatial Modulation Transfer in the Human Eye, J. Opt. J. Optical Society of America, Vol. 57, Issue 3, pp. 401-406, 1967

[18] Anderson, S.J., Mullen, K.T., Hess, R.F.: Human Peripheral Spatial Resolution for Achromatic

and Chromatic Stimuli: Limits Imposed by Optical and Retinal Factors, J. Physiol. 442, pp. 47-64, 1991
[19] Dowling, J. E., Boycott, B.B.: Organization of the Primate Retina: Electron Microscopy, Proc. R. Soc. 166, pp. 80-111, 1966
[20] Kolb, H.: Organization of the Outer Plexiform Layer of the Primate Retina: Electron Microscopy of Golgi-Impregnated Cells, Phil. Trans. R. Soc. 258, pp. 261-283, 1970
[21] Liang, J., Williams, D.R., Miller, D.T.: Supernormal Vision and High-Resolution Retinal Imaging Through Adaptive Optics, J. Opt. Soc. Am. A14, pp. 2884-2892, 1997
[22] Derrington, A.M., Krauskopf J., Lennie, P.: Chromatic Mechanisms in Lateral Geniculate Nucleus of Macaque, J. Physiol. 357, pp. 241-265, 1984
[23] Pridmore, R.W.: Extension of Dominant Wavelength Scale to the Full Hue Cycle and Evidence of Fundamental Color Symmetry. Color Res. Appl., 18, pp. 47–57, 1993
[24] Daly, S.: The Visible Differences Predictor: an Algorithm for the Assessment of Image Fidelity, Digital Images and Human Vision, pp. 179-206, MIT Press 1993
[25] Seeman, M., Zemčík, P.: Vision Physiology Survey for Image Reproduction and Manipulation, submitted to Computers & Graphics
[26] Seeman, M., Zemčík, P.: Visual Acuity and Comfortable Distance from a Display, Poster Proc. of the 20th Int. Conf. in Central Europe on Computer Graphics, Visualization and Computer Vision WSCG, 2012
[27] Seeman, M., Zemčík, P.: Improving Image Perception on Display Devices, submitted to J. Vis. Commun. Image R.
[28] Hill, W., Duggan, M., Keely, L.B., Hitchcock, G.C., Whitted, J.T.: Methods and Apparatus for Performing Image Rendering and Rasterization Operation, US Patent 6307566B1, 2001
[29] Yamada, E.: Some Structural Features of the Fovea Centralis in the Human Retina, Arch. Ophthal. 82, pp. 151-159, 1969
[30] Peichl, L., Wassle, H.: Size, Scatter and Coverage of Ganglion Cell Receptive Field Centres in the Cat Retina, J. Physiol. 291, pp. 117-141, 1979
[31] Seeman, M., Zemčík, P.: Histogram Smoothing for Bilateral Filter, Proc. GraVisMa, Plzen, Czech Rep., pp. 145–148, 2009
[32] Yoshizawa, S., Belyaev, A.G. Yokota, H.: Fast Gauss Bilateral Filtering, Computer Graphics Forum, 29, pp. 60–74., 2010
[33] Reinhard, E., Devlin, K.: Dynamic Range Reduction Inspired by Photoreceptor Physiology, IEEE Trans. on Visualization and Computer Graphics, Vol. 11, No. 1, pp. 13-24, 2005
[34] Zemčík, P., Přibyl, B., Herout, A., Seeman, M.: Accelerated Image Resampling for Geometry Correction, J. Real-Time Image Proc., Vol. 6, No. 3, ISSN 1861-8200, 2011
[35] Gouras, P.: Color Opponency From Fovea to Striate Cortex, Investigative Ophthalmology, Vol. 11, No. 6, Jun 1972
[36] Gallagher, A.C.: Detection of Linear and Cubic Interpolation in JPEG Compressed Images, Proc. of the 2nd Canadian Conference on Computer and Robot Vision, Vol. 171, pp. 65-72, 2005
[37] Zemčík, P., Herout, A., Seeman, M., Přibyl, B.: Způsob a Zařízení pro Digitální Korekci Obrazu, Czech Republic Patent 2010-650, 2010