Human visual system model for Earth observation system's informative parameters estimation

A. Ivanov [1),]M. Kotov [2)]

1) Public corporation "Peleng", engeneer; Minsk, Makaenka str., 23a; Api@inbox.ru, Peleng.by

2) Public corporation "Peleng", engeneer; Minsk, Makaenka str., 23a; Api@inbox.ru, Peleng.by

Abstract: *Human eye is a complex nonlinear part of Earth observation system's optoelectronic route. To estimate an observation system's parameters correctly one needs an adequate human visual system (HVS) model. In this paper a HVS model designed for limiting spatial frequency resolution estimation is presented. Modeling results lie within the 95% confidence interval of subjective evaluations.*

***Keywords*:** image quality assessment, human visual system, eye spatial resolution.

1. INTRODUCTION

Full Earth observation system's optoelectronic route model includes HVS. A number of studies devoted to HVS model design for different purposes are done. [1-2].

The main informative parameter of an Earth observation system is its limiting spatial frequency resolution. Usually, this parameter is estimated with models without HVS or subjectively with the standard resolving power test image (fig. 1). These methods are not optimal, require a number of human resources and usually subjective. The main goal of this work was to develop a n objective HVS model, corresponding to subjective limiting spatial frequency resolution estimation.

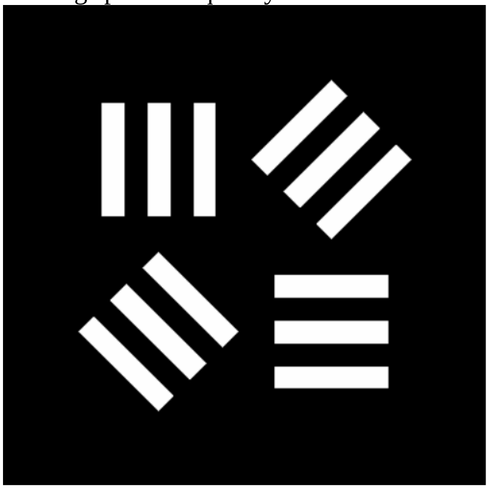

Fig.1 - Standard resolution power test image.

2. HVS MODEL

Developed HVS model's input (fig. 2) is 8-bit depth grayscale test-images.

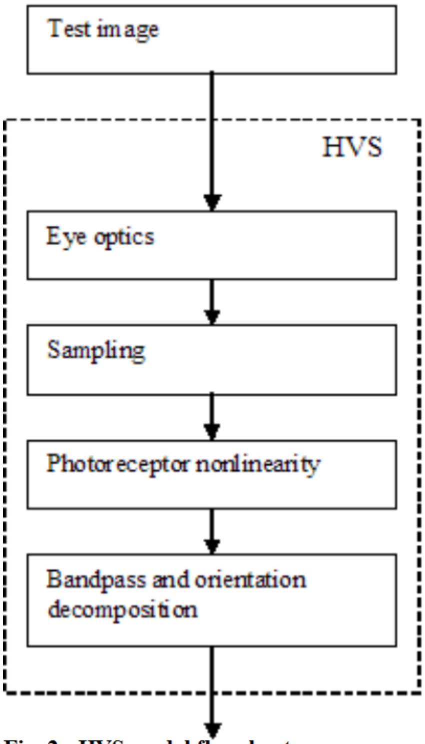

Fig. 2 - HVS model flowchart.

Eye optics is usually modeled using a Point Spread Function (PSF) in the spatial domain or Optical Transfer Function (OTF) in the Fourier domain. The model uses PSF approximation:

$$\begin{aligned} Q(\rho) &= 0.952\exp(-2.59|\rho|^{1.36}) + \\ &+ 0.048\exp(-2.43|\rho|^{1.74}), \end{aligned} \tag{1}$$

where ρ - the distance from the geometrical center of image is in minutes of visual angle arc [3].

The photosensitive cells of the eye retina perform sampling of the image. Photoreceptor's structure can be easily reproduced [4]. In the fovea photoreceptors are densely packed on a hexagonal sampling array (Fig. 3). To perform such a sampling, input image is enlarged about twenty times first. Then it is sampled using a hexagonal mask. In a resulting image (Fig. 4) each hexagonal cell is processed separately as a single photoreceptor.

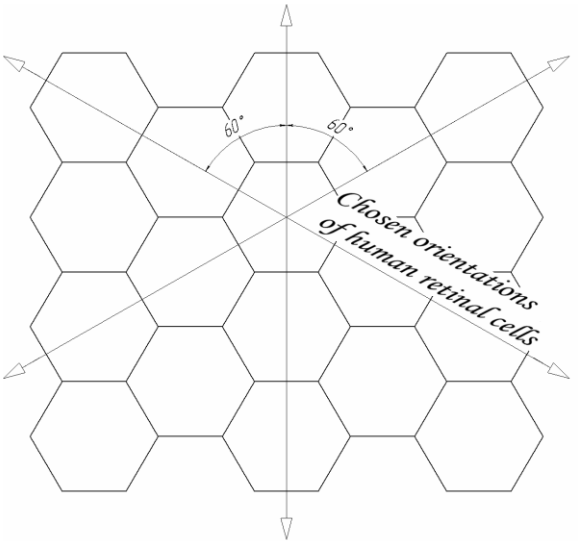

Fig. 3 - Human retina model.

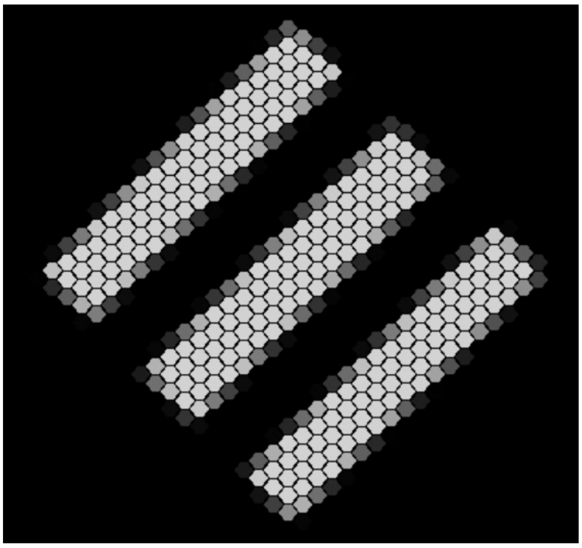

Fig. 4 – Sampled standart power test image.

Photosensitive cells convert light into neural signals, which travel further to the different areas of human brain. Photoreceptor output signal is not linear to the input and depends strongly on eye adaptation to light level. That's why HVS is mostly sensitive to the relative light levels. The response of the photoreceptors is usually modelled as an S-shaped function (on log-linear plot), known as the Michaelis-Menten or Naka-Rushton equation:

$$\frac{R}{R_{\max}} = \frac{Y^n}{Y^n + \sigma^n}, \qquad (2)$$

where R - photoreceptor response, $R_{\max}$ - maximum response, Y - luminance, σ - half-saturation constant, and n - sensitivity control exponent that has value between 0.7 and 1.0.

If high complexity of those models can not be afforded or lower accuracy of the visual model is acceptable, this equation can be replaced with simpler formulas that do not depend on the adaptation state. This is possible by introducing a simplifying assumption that the eye can perfectly adapt to very small patches, such as single pixels. Assuming that n = 1, Daly proposes a shift invariant model of photoreceptor response [5]:

$$\frac{R}{R_{\max}} = \frac{Y}{Y + c_1 Y^b}, \qquad (3)$$

where b=0,63 and c_1 =12,6.

For 8-bit images, $R_{\max} = 255$.

The widely accepted multiresolution theory claims that the images registered by the retina are transmitted to the brain via several visual channels, each one carying information about different spatial/temporal frequency band, orientation and color. An example of such multi-resolution representation, excluding temporal and color aspects, is known as Cortex Transform [6].

There are several computational models of visual channels, which differ from each other by the conformity with the psychophysical measurements and their computational cost [5-7].

According to experimental results, human monochromatic vision has from 3 to 6 dedicated orientations. The decomposition in model presented has 3 orientations (fig. 3) and a variable number of spatial bands. All channels are processed separately.

3. Experiments

All the experiments were held using standard power resolution test image (fig.1) as a test object. At first, test object was rotated on a random angle. Then it undergone sampling with step Δ, which corresponds to sampling on observation system photosensitive matrix. The goal of observations was to evaluate the maximum power of Gaussian noise when test object can be seen for a range of spatial frequencies. Spatial frequencies were calculated by multiplying the frequency of lines of test object on sampling step of photosensitive matrix Δ.

In subjective experiment the maximum power of Gaussian noise was determined by 5 observers of different level of preparation. Observers were placed on a 3 meter distance from display. The test object scaled the way its angular size remained constant for observer. For each sampling step experiment was repeated 20 times to achieve better accuracy.

In the second experiment we've replaced subjective observer with described HVS model. Model took a separate line triples as its input. This triple was oriented perpendicular to one of three chosen orientations of the model, which corresponds to the orientation adaptation of the human eye. Then this orientation frequency channel was analyzed. According to the presence or absence of a frequency peak on a frequency corresponding to test object spatial frequency (fig. 5) model reach a decision whether it is possible to resolve test object or not.

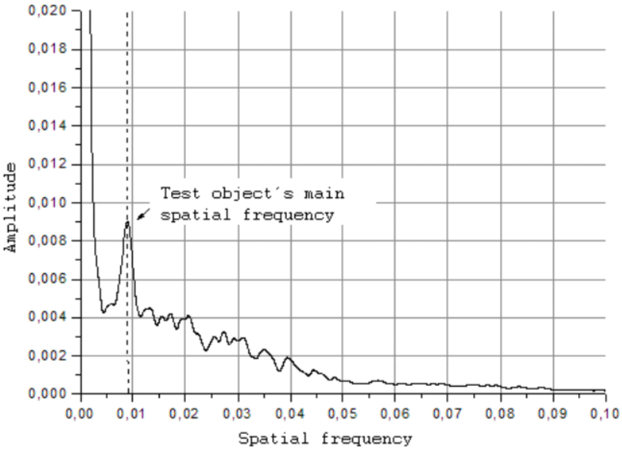

Fig.5 - Test object specter for chosen HVS model orientation.

4. RESULTS

Experimental results are presented on fig. 5. Crosses refer to subjective experimental results, solid line – simulation results. Dots show the boundaries of a 95% confidence intervals for all 5 observers. Signal to noise ratio was chosen as more common metric rather then Gaussian noise power.

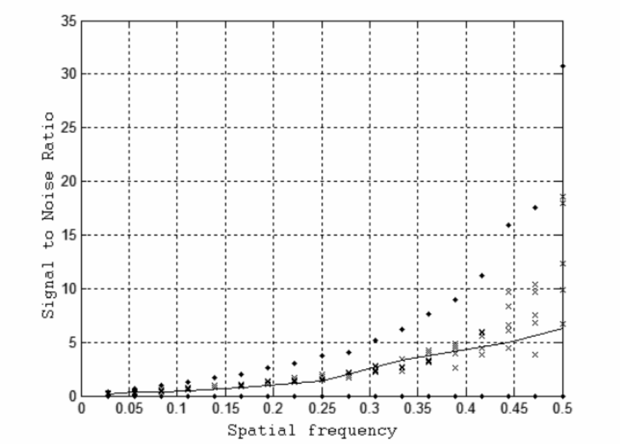

Fig. 5 - Experimental results.

The modeling results correspond to the subjective experimental results. The abnormal difference in subjective estimations at high frequencies can be explained by the observer's different preparation level. Modeling results lie in 95% confidence intervals of subjective evaluations.

5. CONCLUSION

A new HVS model for Earth observation system's informative parameters estimation proposed. Model is based on the PSF, photon accurate model of the human eye, photoreceptor nonlinearity equation, orientation and bandpass transforms. A resolution power experiments with subjective observers and objective model have being conducted. As a result, a minimum signal to noise ratios were achieved for a range of spatial frequencies. Modeling results correspond to the subjective evaluations, which prove the model's fidelity.

6. REFERENCES

[1] Winkler S., Digital video quality. Vision models and metrics. John Wiley & Sons Ltd, 2005. – 175 p.

[2] Wang Z., Hamid R. Sheikh and Alan C. Bovik OBJECTIVE VIDEO QUALITY ASSESSMENT Chapter 41 in The Handbook of Video Databases: Design and Applications, B. Furht and O. Marqure, ed., CRC Press, pp. 1041-1078, 2003.

[3] Westheimer, G. The eye as an optical instrument. In Boff, K. R., Kaufman, L.,Thomas J. P. (eds), Handbook of Perception and Human Performance, vol. 1, chap. 4,John Wiley, 1986.

[4] DEERING, M. F.. A photon accurate model of the human eye. ACM Transactions on Graphics 24, 3, 2005.

[5] DALY, S. The Visible Differences Predictor: An algorithm for the assessment of image fidelity. In Digital Image and Human Vision, Cambridge, MA:MIT Press, A. Watson, Ed., 1993.

[6] Watson A., The cortex transform: Rapid computation of simulated neural images. Computer vision, graphics and image processing 39, 311-327, 1987.

[7] Bradley A., A wavelet visible difference predictor. IEEE transactions on image processing, vol. 8, no. 5, 1999 .