Masking in Chrominance Channels of Natural Images – Data, Analysis, and Prediction

Vlado Kitanovski, Marius Pedersen
Colourlab, Department of Computer Science
Norwegian University of Science and Technology, Gjøvik, Norway
{vlado.kitanovski, marius.pedersen}@ntnu.no

***Abstract*— This paper addresses the visual masking that occurs in the chrominance channels of natural images. We present results from a psychophysical experiment designed to obtain local thresholds of just noticeable log-Gabor distortion in the Cr and Cb channels of natural images. We analyzed the data and investigated the correlation between several low-level image features and the collected thresholds. As expected, features like variance, entropy, or edge density were correlated relatively high with the thresholds. We evaluated the performance of linear and non-linear regression (using neural networks and support vector machines) for thresholds prediction from multiple global image features; we also fitted a modified Watson-Solomon's computational model (based on log-Gabor features) for thresholds prediction. The evaluation showed that neural networks and support vector machines are most suitable for thresholds prediction. The computational model performed reasonably well, with further prospects of its improvement.**

Keywords—visual masking; natural images; masking models;

I. INTRODUCTION

Visual masking occurs when the visibility of a visual target is affected by the presence of another visual stimulus (mask). It can be quantified with a threshold - the amount of particular distortion introduced to the mask from the target, which becomes just noticeable by a human observer. As the image content changes locally, for many applications it is beneficial to have a masking map that is simply a map of thresholds for perceptible distortion in different image regions. In image quality assessment, masking maps are utilized to weigh the objectively measurable distortions according to their visibility [1]. In image compression, different schemes for visual masking have been used to create perceptually uniform quantization tables [2]. In data hiding/watermarking, masking/visibility maps are used to properly distribute the watermark energy and achieve perceptually equalized data embedding [3].

Traditionally, studies related to the sensitivity of the Human Visual System (HVS) to image distortions have used unnatural setups for measuring the distortion detection thresholds. Red-green and blue-yellow sinusoidal gratings on homogenous background have been used to measure the detection thresholds for different spatial frequencies [4], in different color spaces like CIELAB and YCbCr [5], for different background luminance [6] and across the whole visual field [7]. Using isolated chromatic gratings on uniform background masks allows precise investigation of its effects on the detection thresholds, and provides useful but only global understandings of HVS. However, due to the high non-linearity and complexity of the HVS, results obtained using these artificial setups may have limited usage for predicting the thresholds in natural masks [8]. Several researchers have measured detection thresholds in natural images. Sensitivity to phase distortions was measured in grayscale natural images [9] and in color images [10], [11]. Alam *et al*. measured the local detection thresholds for a log-Gabor noise target in natural grayscale images [12]. The images were block-segmented and a masking map was obtained for all of the images from the CSIQ image database [13]. They found that the thresholds depend on visual complexity, fineness of texture, sharpness and overall luminance. However, the masking maps for grayscale images they have provided may not be the most suitable for applications where the image distortion is distributed mainly in the chrominance channels, for example, like in data hiding applications [14].

Despite the obvious need for masking maps in various color image-processing applications, we were unable to find a large dataset of masking thresholds in chrominance channels of natural images. To address this issue, in this paper we present results of a visual experiment where we collected thresholds for detecting log-Gabor noise targets in the chrominance channels of the YCbCr-represented natural images. The details of the experiment setup and the analysis of collected data are given in the next section. In Section III, we investigate the potential of thresholds prediction using both regression approaches and a computational model for predicting the perceptibility of image differences.

II. VISUAL EXPERIMENT AND DATA ANALYSIS

The first part of this section describes the visual experiment we performed for collecting just noticeable distortion thresholds in the Cr and Cb chrominance channels. In the second part, we analyze the collected thresholds and their relations with some common statistical image features. This analysis is performed in a similar way as it is in [12]. The data from our visual experiment can be downloaded from [15].

A. Experiment Setup

For our visual experiment, we used a total of 480 image patches of 80×80 size as masks: 160 patches were obtained

from each of the Kodak [16] and CID:IQ databases [17], 120 patches were selected from the CSIQ database [13], and another 40 patches were obtained from random natural images. The patches were selected so they have consistent texture, and they include wide variety in terms of luminance levels, hues, saturation, and texture types. The visual stimuli were displayed on a Dell U2412M LED/LCD monitor using its native 1920×1200 resolution. The display was setup to conform to the sRGB standard (calibrated to a Gamma curve γ = 2.2, D65 white point, the minimum and maximum luminance were set to 1cd/m^2 and 80cd/m^2 respectively). The subjects viewed the stimuli in a darkened room at a distance of around 90cm. The displayed stimuli consisted of two masks and one mask with added target in one of the chrominance channels. The noise target was cropped from the same 510×510 noise target used in the previous gray-scale experiment by Alam et al. [12]. It is a normalized log-Gabor noise patch with vertical orientation and 5 cycles per degree (cpd) spatial frequency for the selected viewing distance and monitor resolution. The image that contained the target was generated as follows: the RGB mask image was transformed to the YCbCr space; the noise target was multiplied by a constant (the magnitude) and added to either the Cr or the Cb channel. The modified YCbCr image was transformed back to the RGB space. The resulting images were padded with 60 pixels of content from all sides to 200x200 pixels, and then were blended with the background using a circular-shaped 2D window. The visual angle of the target patch is 1.37°, while the whole padded stimuli is 3.41°. An example of the displayed stimuli in our visual experiment is shown in Fig. 1.

The thresholds were collected using the method of adjustments. The subjects viewed the three images placed next to each other on a 17cd/m^2 neutral background. The subjects used keyboard input to increase or decrease the visibility of the target (by increasing/decreasing the magnitude of the inserted target). For each subject there were two separate runs of the experiment, denoted as run A and run B. In the run A, the starting visibility of the target was very low. Subjects were instructed to increase the visibility of the target until they can correctly identify which of the three images contains it. In the run B, the starting visibility of the target was very high so it was easy to identify which of the images initially contains it. Subjects were instructed to decrease the visibility of the target until they can no longer identify which of the three images contains the target. After every increase/decrease, the three images disappeared (were substituted with the background) for 0.25 seconds. During this time, the target was multiplied with the new magnitude, and added to one of the three images (randomly selected). Around one-half of all of the subjects had run A first, the other half had run B first.

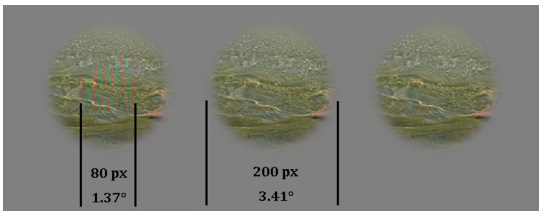

Figure 1. Stimuli display setup; log-Gabor target is inserted in the central part of the Cr channel of the left image.

The total number of observers was 24. Each of the 480 images was observed by three observers. While one observer, Subject1, observed all of the 480 images for both Cr and Cb channels, the results for Subject2 and Subject3 were effectively consisted of the results from 23 subjects. Each of these 23 persons observed at least 48 images for at least one channel (Cr or Cb). All of the participating observers had normal or corrected-to-normal color vision.

The thresholds were recorded in terms of magnitude of the log-Gabor target, as well as RMS chrominance channel difference between the mask and the mask with added log-Gabor target. The distortion detection threshold per image per chrominance channel was calculated as the average threshold from the two runs, A and B.

B. Analysis of Collected Thresholds

In this subsection, we provide analysis of the collected thresholds, their consistency across subjects and their correlation with different image features. The results are provided in terms of Pearson correlation coefficient (CC) or Spearman rank-order correlation coefficient (SROCC).

The average inter-subject correlation of the collected (RMS-based) thresholds, in terms of CC, was 0.76 and 0.88 for the Cr and the Cb thresholds, respectively. Table 1 shows the correlation between the two runs, A and B, which can be used as an indicator of intra-subject consistency. From these results, it can be seen that both the inter- and intra-subject consistencies are higher for the Cb thresholds. The lower intra-subject correlation for the Cr thresholds implies that the Cr thresholds (the average of the two runs) may have higher variance per subject and partly explain why the inter-subject correlation for the Cr thresholds has also been lower. Among the three subjects, the intra-subject correlation is highest for Subject1, which may be expected as the results for Subject1 are collected from only one observer.

The thresholds in the Cr channel are considerably lower than the thresholds in the Cb channel. Even though the YCbCr space is not considered to have high perceptual uniformity, the fact that Cr thresholds are around three times lower than the Cb thresholds confirm what has been previously known - the sensitivity of the HVS to blue-yellow distortions is lower when compared to the sensitivity to red-green distortions. The images that have no texture and are virtually “single-color” gave the lowest thresholds. Generally, the thresholds are increasing as the complexity of the texture increases. The perceived sharpness of the image is also related to the thresholds – increased sharpness leads to higher thresholds, and some heavily blurred images had low thresholds despite the obvious complex texture. Regarding the mean luminance, the thresholds were higher for the very dark images and for the very bright images. Figure 2 shows an example set of images that have both the Cr and Cb thresholds in the same percentile group.

TABLE 1. Pearson correlation (CC) between the two runs, A and B.

CC	Subject1	Subject2	Subject3
Cr channel	0.846	0.758	0.757
Cb channel	0.896	0.777	0.813

Figure 2. Example of images (80x80) with different levels for both Cb and Cr thresholds, grouped from lowest (top row) to highest thresholds (bottom row).

We examined the correlation between the thresholds and several global and commonly used image features like: average luminance, variance, RMS contrast, edge density, entropy, mean saturation, and energy in spatial frequencies occupied by the inserted target. We report the results in terms of SROCC.

We calculated separate variance and entropy for each of the three YCbCr channels of the mask images of size 80x80 pixels. We used first-order entropy, where the probability distribution was approximated with a 256-bins histogram. The RMS contrast of the masks was calculated as in [12] for each of the three YCbCr channels. To calculate edge density we used Laplacian of Gaussian (LoG) edge detector (of size 13x13 pixels) for the Y channel, with the threshold set to 0.002 and the standard deviation of the Gaussian set to σ=2. The edge density feature is simply the percentage of edge pixels in the resulting binary image. The mean saturation of the mask image was calculated in the HSV color space. For each of the YCbCr channels we also calculated the energy of the spatial frequencies occupied by the log-Gabor target, specifically from 3.75 cpd to 6 cpd, – this is denoted in the later text simply as sub-band energy. The actual calculation of this feature was performed in the FFT domain: it equals the sum of the squared amplitudes of all FFT coefficients that correspond to spatial frequencies between 3.75 and 6 cpd, divided by the mask size (80x80). The SROCC values between these mask features and the collected thresholds for both channels are shown in Fig. 3.

From the results in Fig. 3, several conclusions can be made. Overall, the correlations between mask features and the (RMS-based) thresholds of perceivable distortion are higher for the Cb channel thresholds. The reason for this may be the higher sensitivity to noise in the collected thresholds for the Cr channel, as well as the lower dynamic range of Cr channel thresholds - they are roughly a third of the Cb channel thresholds, and their variance is an order of magnitude lower than the variance of Cb thresholds. As expected, the mean luminance and mean saturation and are poorly correlated with the thresholds. While they may have influence on the thresholds, the relation is non-linear and non-monotonic, thus unable to be revealed by linear or ranked correlation. The variance, entropy, RMS contrast, and sub-band energy in general show good correlation with the thresholds. For all of them the SROCC was higher than the CC. Among the calculated three channels, the variance in the Y channel shows highest correlation with the thresholds. Similar pattern can be observed for the entropy, the RMS contrast and the sub-band energy – the values calculated for the Y channel correlate best with both Cr and Cb thresholds. The edge density, as expected, was also highly correlated with the thresholds for both Cr and Cb channels, with SROCC values of around 0.7. When both Cr and Cb thresholds are considered together, the sub-band energy feature performed best – the sum of the two SROCC values (for Cr and Cb thresholds) was highest. The scatter plots of the collected thresholds versus the sub-band energy in the Y channel are shown in Fig. 4, where a noisy but clear positive correlation can be observed. While certain Y-channel features showed relatively high correlation with the RMS-thresholds, combining multiple features in a multiple regression approach may further improve the correlation, and consequently, lead to better thresholds prediction. This is investigated in the next section.

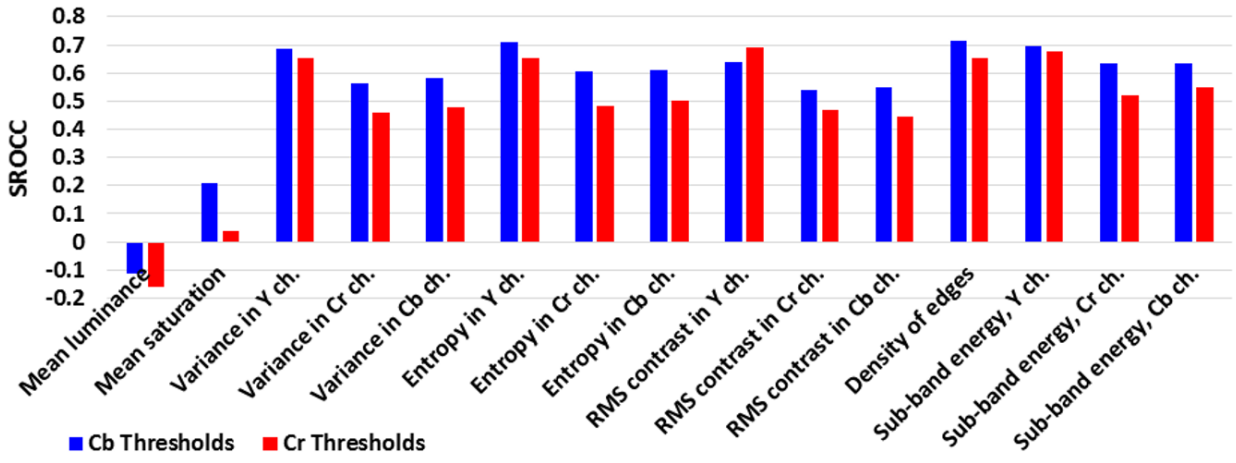

Figure 3. SROCC values between various image (mask) features and the Cb and Cr thresholds.

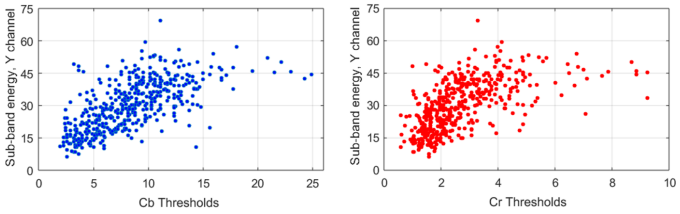

Figure 4. Scatter plots of the collected Cb (left) and Cr (right) RMS-thresholds versus the sub-band energy in the Y channel.

III. Prediction of Masking Thresholds

In this section, we investigate approaches for predicting the thresholds obtained in our visual experiment. We evaluate a multiple linear regression approach as well as non-linear regression approaches such as neural networks and Support Vector Regression (SVR). We also present a modified version of the Watson-Solomon's model [18] that is suitable for thresholds prediction.

A. Prediction using Multiple Linear Regression

Our first choice for thresholds prediction is Multiple Linear Regression (MLR). The thresholds are modelled as a linear combination of selected image features. In the previous section, we used 15 different features to evaluate how they correlate with the collected thresholds. For all of them, the feature calculated for the Y channel showed highest correlation. One way to choose features as predictors in our regression models is to choose the features with highest correlation with the thresholds. As the variance and entropy are very similar to each other, with SROCC between them of 0.94, we include only the entropy (the variance is also part of the RMS-contrast formula [12]). To account for luminance masking, we use the average luminance as input feature. We also test whether mean saturation has significant impact on the model's prediction. Thus, the selected two feature sets for our regression models are given in Table 2. The regression model parameters are fitted by minimizing the least squares (LS) error. While this linear model will not be able to capture highly non-linear relation between the mask features and the thresholds, we are using it here as a baseline regression approach.

B. Prediction using Neural Network

We choose to use a feed-forward fully-connected two-layered neural network (NN), a scheme which has been proven to perform non-linear regression reasonably well [19]. The output layer has only one node since the network is performing regression. The number of input nodes (image features) is either five or six - we are using the same two feature sets that are given in Table 2. The number of nodes in the hidden layer has been set to be one plus the number of input nodes – which was selected as a good choice using empirical tests. We used the Levenberg–Marquardt algorithm [20] for training the network; the training set consisted of features/thresholds from 312 images (65%), we used 48 (10%) for validation and training termination, and the rest 120 (25%) images (not presented during training to the network) were used for testing the network's prediction performance.

C. Prediction using Support Vector Regression

Our second choice is to use Support Vector Machines (SVM) for non-linear regression, specifically the ε–SVR method [21]. For the SVR implementation, we used the publicly available libSVM library [22]. The error tolerance ε was set to 0.001. We used radial basis function kernel that, as expected, provided best results when compared to other types of kernels. The γ parameter of the radial basis function was set to $\gamma = 0.5$, while the regularization constant was empirically chosen to be $C = 100$. We used the same two feature sets as in the previous two regression approaches. The training data consisted of thresholds/features from 360 images (75%), whereas the rest 120 (25%) were used for testing.

D. Prediction using Modified Watson-Solomon's Model

The Watson and Solomon's computational model of pattern masking can be used for predicting the perceptibility of differences in grayscale images. Even though it was published around twenty years ago, its modular structure together with its extensive parameters set make it flexible and potentially capable of incorporating new findings about the human visual system [12], [18].

There are different ways to extend this model so it could predict the perceptibility of difference between color images. We trialed few extensions of the model to images represented in the opponent-colors YCbCr color space, by introducing parallel branches for the additional chrominance channels at various points in the original model. For each model structure, we trialed different settings by sampling the parameters space in the region close to what had been previously selected as nearly optimal [12], [23]. The model structure that resulted in substantially better predictions is shown in Fig. 5, and the parameter set we used in this work is given in Table 3. In Fig. 5, the two input images to be compared by the model follow identical paths, of which only the path for the second image is shown. The input images are transformed to the YCbCr space, and fed to a bank of 24 real log-Gabor filters - with four different passbands (three band-pass and one high-pass) and six different orientations. All of the log-Gabor filters are normalized to unit energy. Each of the filter-bank outputs are summed across the three color channels using YCbCr contrast-sensitivity weights [24] calculated at the central frequencies of the four passbands. The resulting 24 responses are split into non-linear excitatory and inhibitory paths. The responses in the inhibitory path are pooled across space (5x5 neighborhood) and orientation (the closest orientation from each side). The 24 responses from the two non-linear paths are divided, and then the result is subtracted from the one for the other image. Before the division, a saturation constant, b^q, is used to prevent very high responses in the model.

TABLE 2. Two sets of image features used in the regression models.

Set1	Set2
Mean luminance	Mean luminance
Entropy in Y ch.	Entropy in Y ch.
RMS contrast in Y ch.	RMS contrast in Y ch.
Density of edges	Density of edges
Sub-band En. in Y ch.	Sub-band En. in Y ch.
	Mean saturation

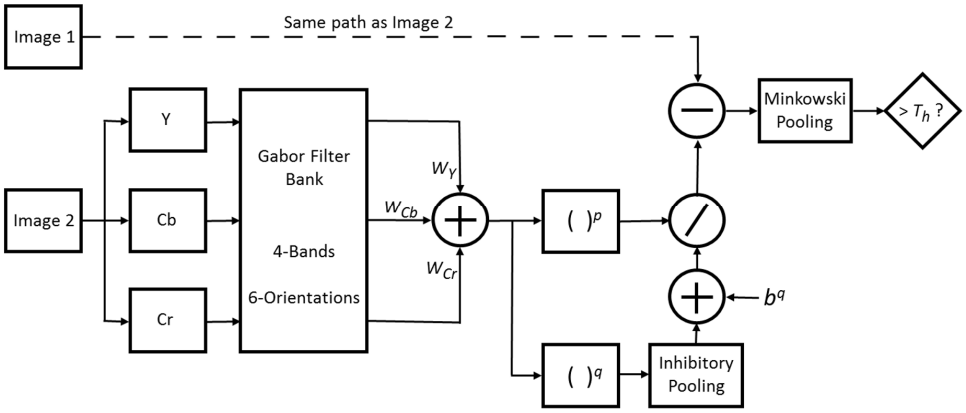

Figure 5. Modified Watson-Solomon's model for predicting the perceptibility of image differences in the chrominance channels.

The phenomenon of visual masking is modelled mainly by the pooling in the inhibitory path and the division of the responses in the two paths – this effectively simulates the reduction in the HVS response for spatially co-located responses that are close in orientation and frequency. The obtained image differences are finally pooled over space (the whole image), frequencies (the four bands) and orientations (the six orientations), using Minkowski pooling. The obtained value is compared to a threshold to decide whether the image difference is perceptible. For more detailed explanations of the elements in this model, readers are directed to the original published work [18].

In order to predict local image distortion threshold of just noticeable difference, the modified Watson-Solomon's model is used iteratively: the amount of distortion added to the second image (or image patch) is increased until the model's response, the Minkowski-pooled difference from the referent first image, becomes higher than the model's threshold T_h. The actual distortion threshold is then calculated as a weighted average of the highest distortion value that results in model's response lower than T_h, and the lowest distortion value that results in model's response higher than T_h. The value of the model's threshold, T_h, can be obtained by calibration with data from subjective experiments. In our case, all of the images distorted at the collected Cb and Cr thresholds from our experiment, were fed to the model (paired with their undistorted version), and the average model's response was used as T_h.

E. Evaluation of the four methods for thresholds prediction

In this subsection, we present evaluation of the four different approaches for thresholds prediction. The accuracy of prediction was measured in terms of correlation (CC and SROCC) and RMS difference between predicted and collected thresholds. For the three regression approaches, we split the data into training (model-fitting) set (75%) and testing set (25%); the results are averages from 100 regression models obtained from 100 different pseudo-random training/testing set separations (that were the same for the three regression approaches). The results for the modified Watson-Solomon's computational model are from a single run on all of the images, as the model's threshold T_h was obtained using all of the experiment data.

The performance of the different threshold prediction methods are given in Table 4. The values corresponding to the best results (highest correlations and lowest RMSE) are in bold. Regarding the three regression methods, the correlation between the predictions and the actual thresholds has improved from using multiple features. Using the mean saturation had no impact on predicting the Cr thresholds, but improved the Cb thresholds prediction. This improvement is relatively small, and it is somewhat expected given that in Fig. 3, mean saturation showed small positive correlation with the Cb but not for the Cr thresholds. The neural networks performed consistently well, with best or next-to-best results. The multiple linear regression performed worst among the three regression methods. The modified Watson-Solomon's model compares relatively well with the non-linear regression methods, and it performed best in terms of SROCC for the Cb thresholds. However, the model's predictions on average had considerably highest RMSE, which can be attributed mainly to the very dark or very bright images – we suggest that this is because the model does not explicitly consider the mean luminance, so it leads to bigger errors for these certain types of images. As for the algorithm's complexity, this computational model is much more complex than the regression methods, due to the iterative implementation, the large Gabor filter bank, and the extensive pooling in the model.

TABLE 3. Parameters used for the modified Watson-Solomon's model.

Parameter	Value
Bandwidth of frequency bands of log-Gabor filters	1 octave
Center frequencies of the bands	2.9, 5.7, 10.6, 21.1 cpd
Bandwidth of orientation of log-Gabor filters	30°
Center angles of orientations of log-Gabor filters	0, ±30°, ±60°, 90°
Spatial pooling kernel	5x5 Gaussian, σ=1
Pooling across orientations	±30° with equal weights
Excitatory exponent p	2.3
Inhibitory exponent q	2
Semi-saturation constant b	0.05
Minkowski pooling exponent	4

TABLE 4. Performance of different threshold predictors.

	CC	SROCC	RMSE
Cr thresholds:			
MLR – Set1	0.68	**0.76**	1.06
MLR – Set2	0.68	0.75	1.06
NN – Set1	0.70	**0.76**	**1.00**
NN – Set2	**0.71**	**0.76**	1.01
SVR – Set1	0.70	0.75	1.01
SVR – Set2	**0.71**	0.74	1.02
Modified W.-S. model	0.68	0.74	1.21
Cb thresholds:			
MLR – Set1	0.72	0.76	2.67
MLR – Set2	0.74	0.78	2.59
NN – Set1	0.72	0.76	2.68
NN – Set2	**0.76**	0.79	**2.54**
SVR – Set1	0.73	0.76	2.65
SVR – Set2	**0.76**	0.77	2.57
Modified W.-S. model	0.74	**0.80**	3.18

IV. CONCLUSION

In this paper, we presented results of a visual experiment for obtaining thresholds of perceptible distortion in the chrominance channels of the YCbCr color space of natural images. A total of 480 80×80 images with a variety of natural content, texture, hue, luminance and saturation levels were used in the experiment as a mask. The distortion target used was a log-Gabor patch, and it was inserted in the Cr and Cb channels of the mask. The analysis of the experiment data showed that the thresholds are influenced by the visual complexity of the mask, the texture type and the mean luminance levels. We examined the correlation of different low-level global image features with the thresholds; several features like variance, entropy, edge density, or energy in spatial frequencies occupied by the target, have shown high correlation with the thresholds – above 0.6. Both linear and non-linear regression approaches were investigated for improving the threshold prediction from the low-level image features. We presented a modified Watson-Solomon's computational model for prediction of the perceptibility of image differences in the chrominance channels. Among the four different methods for thresholds prediction, the non-linear regression methods, especially the neural networks, provided marginally better results. Given their low computational complexity, we select the neural networks as preferable choice for thresholds prediction. The CC/SROCC correlations with the collected thresholds improved for around 0.1 percentage points when using multiple features in the NN/SVR models. The modified Watson-Solomon's model performed relatively well, given that it does not account for luminance masking. Even though the model is substantially more computationally intensive than the regression methods, its good performance should be emphasized because, apart from the final threshold T_h, the model did not explicitly use the experiment data for optimizing its parameters.

The future work will focus on integrating the masking thresholds prediction into the chrominance channels based data-hiding scheme [14], in order to achieve perceptual uniformity of the introduced distortion from data embedding in color images.

REFERENCES

[1] S. Hu, L. Jin, H. Wang, Y. Zhang, S. Kwong, "Compressed image quality metric based on perceptually weighted distortion", IEEE Trans. on Image Proc., Vol. 24, No. 12, pp. 5594-5608, Dec. 2015.

[2] A. B. Watson, "Perceptual optimization of DCT color quantization matrices", Proceedings of IEEE International Conference on Image Processing, pp. 100-104, Austin, Nov. 1994.

[3] A. Reed, D. Berfanger, Y. Bai, and K. Falkenstern, "Full-color visibility model using CSF which varies spatially with local luminance", SPIE Proc. Imaging and Multimedia Analytics in a Web and Mobile World 2014, vol. 9027, pp. 902705-902705-12, San Francisco, Feb. 2014.

[4] K. T. Mullen, "The contrast sensitivity of human colour vision to red/ green and blue/yellow chromatic gratings," Journal of Physiology, vol. 359, pp. 381–400, Feb. 1985.

[5] J. Yao, "Measurements of human vision contrast sensitivity to opposite colors using a cathode ray tube display", Chinese Science Bulletin, vol. 56, no. 23, pp. 2425-2432, Aug. 2011.

[6] K. J. Kim, R. Mantiuk, and K. H. Lee, "Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance", SPIE Proc. Human Vision and Electronic Imaging, vol. 8651, pp. 86511A-86511A-14, Burlingame, Feb. 2013.

[7] M. A. Diez-Ajenjo, P. Capilla, and M. J. Luque, "Red-green vs. blue-yellow spatio-temporal contrast sensitivity across the visual field", Journal of Modern Optics, vol. 58, pp. 1736-1748, 2011.

[8] P. J. Bex, S. G. Solomon, and S. C. Dakin, "Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure", Journal of Vision, vol. 9, no. 10, pp. 1.1-19, Sep. 2009.

[9] P. J. Bex, "(In)Sensitivity to spatial distortion in natural scenes", Journal of Vision, vol. 10, no. 2, pp. 23.1-15, Feb. 2010.

[10] A. Yoonessi, and F. Kingdom, "Comparison of sensitivity to color changes in natural and phase-scrambled scenes", Journal of Optical Society of America, vol. 25. no. 3, pp. 676-684, Mar. 2008.

[11] B. J. Jennings, and F. Kingdom, "Detection of chromatic and luminance distortions in natural scenes", Journal of Optical Society of America, vol. 32, no. 9, pp. 1613-1622, Sep. 2015.

[12] M. Alam, K. P. Vilankar, D. J. Field, and D. M. Chandler, "Local masking in natural images: A database and analysis", Journal of Vision, vol. 14, no. 8, pp. 22.1-38, Aug. 2014.

[13] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," Journal of Electronic Imaging, vol. 19, no. 11, pp. 011006- 011006-21, Mar. 2010.

[14] V. Kitanovski and M. Pedersen, "Orientation modulation for data hiding in chrominance channels of direct binary search halftone prints", Journal of Imaging Systems and Technology, Vol. 60, No. 5, pp. 50407-1-50407-9(9), Sept-Oct 2016.

[15] Visual experiment data, http://colourlab.no/cid

[16] Kodak lossless true color image suite, http://r0k.us/graphics/kodak

[17] X. Liu, M. Pedersen, and J. Y. Hardeberg, "CID:IQ - A new image quality database", Image and Signal Processing, vol. 8509, pp. 193–202, Springer, 2014.

[18] A. B. Watson, and J. A. Solomon, "A model of visual contrast gain control and pattern masking" The Journal of Optical Society of America, vol. 14, No.9, pp. 2379–2391, 1997.

[19] A. Landi, P. Piaggi, M. Laurino, D. Menicucci, "Artificial Neural Networks for nonlinear regression and classification", Proc. Intl. Conf. on Intell.. Systems Design and Appl., pp. 115-120, Cairo, Nov. 2010.

[20] D.W. Marquardt. "An algorithm for least-squares estimation of nonlinear parameters," Journal of the Society for Industrial and Applied Mathematics, Vol. 11, No.2, pp.431-441, 1963.

[21] V. Vapnik, Statistical Learning Theory, Wiley, New York, NY, 1998.

[22] C. Chang, and C. Lin, "LIBSVM: a library for support vector machines", ACM Transactions on Intelligent Systems and Technology, Vol. 2, No. 3, pp.27:1-27:27,Apr. 2011.

[23] M. J. Nadenau, J. Reichel, and M. Kunt, "Performance comparison of masking models based on a new psychovisual test method with natural scenery stimuli", Signal Processing: Image Communication, Vol. 17, No. 10, pp. 807-823, Nov. 2002.

[24] M. J. Nadenau, "Integration of human color vision models into high quality image compression" PhD thesis, Ecole Polytechnique Federale de Lausanne, 2000.