A TOTAL VARIATION BASED COLOR IMAGE QUALITY METRIC WITH PERCEPTUAL CONTRAST FILTERING

Marius Pedersen[*,**], *Gabriele Simone*[*], *Mingming Gong*[*,***], *Ivar Farup*[*]

[*]Gjøvik University College, Gjøvik Norway.
[**] Océ Print Logic Technologies, Creteil, France
[***] Huazhong University of Science and Technology, Wuhan, China

ABSTRACT

In the last two decades, the evaluation between an original image and its reproductions has been widely considered by many researchers. Recent studies have shown that contrast is one of the most important image features falling under the umbrella of image quality factors. Total variation has shown to be a useful tool in different areas of computer vision. In this paper we introduce a novel image quality metric, named Total Variation of Difference (TVD), combining the total variation method with a local band-limited contrast filtering. Extensive tests and analysis of different pooling methods are carried out on two different databases. Results show a particular high correlation on the second database using Minkowski pooling.

***Index Terms*—** Image Quality, Metrics, Contrast, Filtering, Sensitivity, Image Difference

1. INTRODUCTION

Measuring Image Quality (IQ) has become more and more important as new technologies emerge. A popular and efficient way for measuring IQ is by using metrics. An impressive number of metrics have been proposed in the literature [1]. However, an efficient color IQ metric using spatial filtering has not been developed yet [2, 3].

Since the introduction of Total Variation (TV) in image processing in 1992 by Rudin et al. [4], TV has become increasingly popular. In their pioneering work on edge preserving image denoising, the use of variational image processing has been extended to several areas of computer vision, such as inpainting, segmentation, and deblurring. Originally developed for intensity images, TV has been extended to color images by Blomgreen and Chan [5]. Furthermore, in the last decades several efforts have been done for developing fast and robust TV solvers such as the one by Chen and Tai [6]. Due to page limitation we address the reader to Chan and Shen [7] for a detailed overview of variational image processing methods. In this paper we introduce TV in the field of IQ metrics.

Human observers are sensitive to various frequencies of a visual stimuli; the Contrast Sensitivity Function (CSF) tells us how sensitive. If the frequency of visual stimuli is too high, we will not be able to differentiate between stimuli patterns. The use of CSFs have been popular in image quality metrics, such as the Spatial-CIELAB (S-CIELAB) [8] and Spatial-ΔE_E (S-DEE) [9]. In these metrics, the CSFs are commonly used to modulate frequencies that are less perceptible [10]. The common way to do this is to use convolution kernels to ”blur” the spatial frequencies that observers cannot perceive [8]. This method is fast, but does not result in the most precise filtering of the image [11]. Recent studies have shown that contrast is one of the most relevant perceptual and IQ factors [12]. The history of contrast is one century long, and measuring perceived contrast is not a trivial task [13]. An important milestone was given by Peli [14] in 1990, who defines a local band-limited contrast for complex images. This work will be explained in details later in the paper as it will be relevant for our proposal.

The rest of this paper will be organized as follows: first we provide the description of a new image difference metric. Then we introduce background information on contrast filtering. Next, we describe how we evaluated the new metric and we will present the results and discuss how the metric reflects perceived quality. At last, conclusions are drawn.

2. BACKGROUND

Peli [14] introduced a method to simulate the human visual system, where contrast at each point in an image is calculated separately to account for variations across the image, and since contrast sensitivity depends on frequency, contrast is also calculated for different frequency bands.

Peli [14] proposes the idea of a pyramidal image-contrast structure where for each frequency band, the contrast is defined as the ratio of the bandpass-filtered image at that frequency to the low-pass image filtered to an octave below the same frequency (local luminance mean).

To define local band-limited contrast for a complex image, he obtains a band-limited version of the image in the frequency domain $A(u, v)$:

$$A(u, v) \equiv A(r, \theta) \equiv F(r, \theta)G(r), \tag{1}$$

Fig. 1. Original on the left, simulated filtered image in the center, and CSF based filtered image on the right. The images simulate a distance of 200 cm.

where u and v are the respective horizontal and vertical spatial frequency coordinates, $G(r)$ is a band-pass filter, and r and θ represent the respective polar spatial frequency coordinates: $r = \sqrt{u^2 + v^2}$, $\theta = \tan^{-1}(u/v)$, and $F(r, \theta)$ is the Fourier transform of the image $I(x, y)$.

In the spatial domain the filtered image $a(x, y)$ can be represented similarly, that is, as:

$$a(x, y) = I(x, y) * g(x, y), \tag{2}$$

where $*$ is the convolution, and $g(x, y)$ is the inverse Fourier transform of the band-pass filter $G(r)$. In Peli's approach of measuring local contrast, the pyramid is obtained as follows:

$$A_i(u, v) \equiv A_i(r, \theta) \equiv F(r, \theta) G_i(r), \tag{3}$$

where $G_i(r)$ is a cosine log filter centered at frequency of 2^icycles/picture, expressed as:

$$G_i(r) = \frac{1}{2}\left(1 + \cos\left(\pi \log_2 r - \pi i\right)\right). \tag{4}$$

The resulting contrast at the band of spatial frequencies can be represented as a two-dimensional array $c_i(x, y)$:

$$c_i(x, y) = \frac{a_i(x, y)}{l_i(x, y)}, \tag{5}$$

where $a_i(x, y)$ is the corresponding local luminance mean image and $l_i(x, y)$ is a low-pass-filtered version of the image containing all energy below the band.

This filtering differs from other types of filtering because suprathreshold features retain contrast and are not washed out [14] as seen in Figure 1.

3. THE NEW COLOR IMAGE QUALITY METRIC

We propose a new color IQ metric based on contrast filtering and TV. First the original I_O and reproduction I_R are converted into the $CIEXYZ$ color space. For each channel independently, the contrast of each pixel is calculated as described in Equation 5 in Section 2. The contrast c of each pixel is then compared against the contrast sensitivity threshold (T) for the corresponding channel for each band. If the contrast is suprathreshold the information is perceptible and should be kept, if the contrast is subthreshold the information is discarded. The contrast of each pixel is calculated for each band $L_i(x, y)$:

$$L_i(x, y) = \begin{cases} c(x, y) & \text{if } c(x, y) > T \\ 0 & else \end{cases}. \tag{6}$$

The final filtered image L_f is the sum over the n bands:

$$L_f(x, y) = \sum_{i=1}^{n} L_i(x, y). \tag{7}$$

For the luminance contrast sensitivity thresholds we use the same as Peli [15] while for the chrominance thresholds we use the ones from Johnson and Fairchild [11].

Since the $CIEXYZ$ color space is not orthogonal, i.e. the X and Z channels contain luminance information, we separate these channels into a color part and a luminance part, filtered with their respective contrast sensitivity thresholds. To obtain the luminance bandpass information in the color channel (X_{BL}), the lowpass information in the color channel (X_L) is divided by the lowpass information in the luminance channel (Y_L), and further multiplied with the bandpass information in the luminance channel (Y_B): $X_{BL} = (X_L/Y_L)Y_B$. The color information in the color channel (X_{BC}) is found by subtracting the luminance bandpass information in the color channel (C_{BL}) from the bandpass information in the same color channel (X_B): $X_{BC} = X_B - X_{BL}$.

After the filtering, the original and reproduction are converted to the log-compressed OSA-UCS color space as proposed by Oleari et al. [16]. Euclidean color differences in the OSA-UCS color space are shown to correlate well with perceived differences [17].

The new Total Variation of Difference (TVD) metric, given the original contrast filtered image L_O and its filtered reproduction L_R, is defined as following:

$$TVD = \sqrt{\sum_j \left(\int_\Omega \mid \nabla L_{O_j} - \nabla L_{R_j} \mid dA \right)^2} + \lambda \int \sqrt{\sum_j (L_{O_j} - L_{R_j})^2} dA, \tag{8}$$

where $\sqrt{\sum_i (\int_\Omega \mid \nabla L_{O_j} - \nabla L_{R_j} \mid dA)^2}$ is the TV term, while $\lambda \int \sqrt{\sum_j (L_{O_j} - L_{R_j})^2} dA$ is the Color Difference (CD) term. Ω is the image domain, λ is the weighting parameter for the CD term, and j indicates the color channel. The TV term is similar to the Color TV defined by [5], except that we take the gradient of the difference between the original and reproduction, and the CD term is the Euclidean color difference.

We will also investigate other methods to reduce the number of IQ values into a single number representing quality, so called pooling strategies. For the TV term we will replace the standard outer norm (L2) over the color channels with the L1 norm, minimum, median and maximum. For the CD term we will replace the standard outer norm (average) over the image space with Minkowski (M) [18], Monotonic Function (MF) [18], information content [18], and two different information content based pooling methods based on saliency. We will also test 100 λ values from 0 to 5, with equal steps.

4. EVALUATION

To evaluate the performance of the proposed IQ metric will compare its results against the results of human observers. Two different data sets have been selected for the comparison.

4.1. Test data sets

The first test was proposed by Pedersen et al. [2]. The database contains 24 reference images (Figure 2). The images were printed on an Oce Colorwave 600 CMYK wide format printer using three different rendering intents: perceptual, relative colorimetric, and relative colorimetric with black point compensation. Each printed image were judged by 15 observers. For details we refer to Pedersen et al. [2].

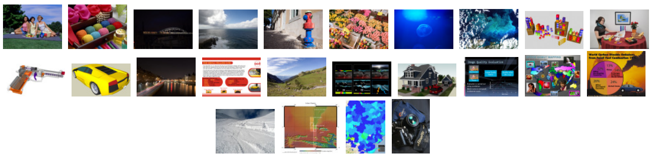

Fig. 2. The 24 reference images in the first test set.

The second test test consists of ten images (Figure 3) from Pedersen et al. [3]. The images were printed by a HP DesignJet 10ps printer using four different modes: the best print mode and the perceptual intent, the best mode and relative colormetic intent, normal print mode and the perceptual intent, and the last with normal print mode and relative colorimetic intent. Ten observers judged the images according to color quality, from which z-scores were calculated. For details we refer to Pedersen et al. [3].

Fig. 3. The ten test images in the second test set.

In order to apply objective IQ metrics to these printed images, these images are scanned into digital images and stored without compression using the framework by Pedersen and Amirshahi [19].

In both data sets the observers were asked to judge the color quality of the images, and the ratings have been quantified as z-scores [20].

4.2. Evaluation procedure

Three state of the art IQ metrics have been chosen for comparison in the evaluation: S-CIELAB [8], S-DEE [9], and ABF [21]. These are made for evaluating color, and other more traditional metrics, such as PSNR and MSE, are not, and therefore not included. We use three evaluation methods to compare the performance of different IQ metrics. Pearson Linear Correlation Coefficient (PLCC) is calculated for each image between the metric scores and subjective z-scores, and the first measure of performance is the mean PLCC of the whole database calculated as the average of PLCCs of each image. The second measure is the Percentage Of Images (POI) with PLCC higher than 60%. The last measure is the Rank Correlation (RC) [22], which is the PLCC correlation between objective rank order z-score and subjective z-score.

4.3. Results and discussion

We will show the results for the following configurations of the TVD metric; TV term with L1 and L2 pooling, λ equal to 0.1, and the best λ, and for the color term we will show the mean pooling together with the best pooling method. The results from the evaluation can be seen in Tables 1 and 2 for the first and second dataset. For the first test set a combination of the TV and CD terms, where the L1 pooling for the TV term, $\lambda = 4$ and the MF pooling with $p = 4$ gives the highest correlation with the perceived color quality. With a $\lambda = 1$ we obtain results similar to existing metrics. TV shows the results for the TV term (Eq. 8), without the spatial filtering and the CD term ($\lambda = 0$). We see that it has similar performance for PLCC and POI, but a higher RC. Nonetheless, the performance of the new metric is not great, most likely since the visual differences of the first test is small, making the task very difficult for IQ metrics.

Table 1. Results for the first test set. The highest PLCC and POI is found with L1 pooling for the TV term, $\lambda = 4$, and using MF pooling with $p = 4$ for the CD term.

Metric	PLCC	POI	RC
S-CIELAB	−0.29	13%	−0.95
S-DEE	−0.34	13%	−0.92
ABF	−0.39	8%	−0.99
TVD (L1/$\lambda = 1$/mean)	−0.31	8%	−0.94
TVD (L2/$\lambda = 1$/mean)	−0.30	13%	−0.44
TVD (L1/$\lambda = 0$)	−0.15	13%	−1.00
TVD (L2/$\lambda = 0$)	−0.16	13%	−0.95
TVD (L1/$\lambda = 4$/$MF_{p=4}$)	0.18	29%	−0.93
TV (L1)	−0.26	21%	−0.12

For the second test set an equal weighting of the TV and CD term gives similar results to the state of the art metrics (Table 2). However, by reducing the importance or removing the CD term the performance of the TVD greatly improves. TV without the spatial filtering and without the CD term, gives slightly lower performance, indicating that the spatial filtering adds value to the TVD metric.

Table 2. Results for the second test set. The best results are found with a $\lambda = 0$, and with $\lambda = 0.5$ and using the Minkowski (M) pooling with $p = 1/8$.

Metric	PLCC	POI	RC
S-CIELAB	-0.27	0%	-0.23
S-DEE	-0.42	0%	-0.42
ABF	0.07	0%	0.23
TVD (L1/λ = 1/mean)	-0.25	0%	-0.15
TVD (L2/λ = 1/mean)	-0.31	0%	-0.29
TVD (L1/$\lambda = 0$)	0.59	70%	0.98
TVD (L2/$\lambda = 0$)	0.56	60%	0.92
TVD (L1/λ = 0.5/M$_{p=1/8}$)	0.59	70%	0.98
TV (L1)	0.53	60%	0.92

5. CONCLUSION

We have developed a novel image quality metric, named Total Variation of Difference (TVD), based on the local band-limited contrast filtering proposed by Peli [14] and the total variation method. This novel metric has been compared with a selection of state-of-the-art metrics on two different databases. On the first database TVD and state-of-the art metrics show low correlation, due by very small visual difference between the original image and its reproductions. On the second database TVD show high correlation and outperforms state-of-the-art metrics using L1-norm for the variational term and Minkowski pooling for the data-attachment term in the total variation method.

References

[1] M. Pedersen and J.Y. Hardeberg. Survey of full-reference image quality metrics. Høgskolen i Gjøviks rapportserie 5, Gjøvik University College, Jun 2009.

[2] M. Pedersen, N. Bonnier, J. Y. Hardeberg, and F. Albregtsen. Estimating print quality attributes by image quality metrics. In *Color and Imaging Conference*, pages 68–73, San Antonio, TX, USA, Nov 2010.

[3] M. Pedersen, Y. Zheng, and J. Y. Hardeberg. Evaluation of image quality metrics for color prints. In *Scandinavian Conference on Image Analysis*, volume 6688 of *Lecture Notes in Computer Science*, pages 317–326, Ystad Saltsjobad, Sweden, May 2011.

[4] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D Nonlinear Phenomena*, 60:259–268, Nov 1992.

[5] P. Blomgren and T. F. Chan. Color TV: Total variation methods for restoration of vector-valued images. *IEEE Trans. Image Process*, 7:304–309, 1998.

[6] K. Chen and X-C. Tai. A nonlinear multigrid method for total variation minimization from image restoration. *Journal of Scientific Computing*, 33:115–138, 2007.

[7] T. Chan and J. Shen. *Image Processing And Analysis: Variational, PDE, Wavelet, And Stochastic Methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005.

[8] X. Zhang and B.A. Wandell. A spatial extension of CIELAB for digital color image reproduction. In *Soc. Inform. Display 96 Digest*, pages 731–734, 1996.

[9] G. Simone, C. Oleari, and I. Farup. Performance of the euclidean color-difference formula in log-compressed OSA-UCS space applied to modified-image-difference metrics. In *11th Congress of the International Colour Association (AIC)*, Sydney, Australia, Oct 2009.

[10] G.M. Johnson and M. D. Fairchild. On contrast sensitivity in an image difference model. In *Image Processing, Image Quality, Image Capture, Systems Conference (PICS)*, pages 18–23, Portland, OR, Apr 2002.

[11] G. M. Johnson and M. D. Fairchild. Darwinism of color image difference models. In *Color Imaging Conference*, pages 108–112, Scottsdale, AZ, Nov 2001.

[12] M. Pedersen, N. Bonnier, J. Y. Hardeberg, and F. Albregtsen. Attributes of image quality for color prints. *J. Electron. Imaging*, 19(1):011016–1–13, Jan 2010.

[13] G. Simone, M. Pedersen, and J. Y. Hardeberg. Measuring perceptual contrast in digital images. *Journal of Visual Communication and Image Representation*, 2010. under 2nd review.

[14] E. Peli. Contrast in complex images. *JOSA A*, 7:2032–2040, 1990.

[15] E. Peli. Contrast sensitivity function and image discrimination. *JOSA A*, 18(2):283–193, 2001.

[16] C. Oleari, M. Melgosa, and R. Huertas. Euclidean color-difference formula for small-medium color differences in log-compressed OSA-UCS space. *Journal of the Optical Society of America A*, 26(1):121–134, 2009.

[17] D. R. Pant and I. Farup. Riemannian formulation and comparison of color difference formulas. *Color Research and Application*, 2011. (Accepted).

[18] Z. Wang and X. Shang. Spatial pooling strategies for perceptual image quality assessment. In *IEEE International Conference on Image Processing*, pages 2945–2948, Atlanta, GA, Oct 2006. IEEE.

[19] M. Pedersen and S.A. Amirshahi. A modified framework the evaluation of color prints using image quality metrics. In *5th European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, pages 75–82, Joensuu, Finland, Jun. 2010. IS&T.

[20] P. G. Engeldrum. *Psychometric Scaling, a toolkit for imaging systems development*. Imcotek Press Winchester USA, 2000.

[21] Z. Wang and J. Y. Hardeberg. An adaptive bilateral filter for predicting color image difference. In *Color Imaging Conference*, pages 27–31. IS&T/SID, Nov 2009.

[22] M. Pedersen and J. Y. Hardeberg. Rank order and image difference metrics. In *4th European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, pages 120–125, Terrassa, Spain, Jun 2008. IS&T.