A MS-SSIM Optimal JPEG 2000 Encoder

Thomas Richter
(richter@rus.uni-stuttgart.de)
University of Stuttgart
RUS Computing Center
70550 Stuttgart, Germany

Kil Joong Kim
(kil210@snubhrad.snu.ac.kr)
Seoul Nat. Univ. College of Medicine
Department of Radiology
Seoul, Korea

Abstract

In this work, we present a SSIM [16] optimal JPEG 2000 rate allocation algorithm. However, our aim is less improving the visual performance of JPEG 2000, but more the study of the performance of the SSIM full reference metric by means beyond correlation measurements.

Full reference image quality metrics assign a quality index to a pair of a reference and distorted image. The performance of a metric is then measured by the degree of correlation between the scores obtained from the metric and those from subjective tests. It is the aim of a rate allocation algorithm to minimize the distortion created by a lossy image compression scheme under a rate constraint.

Noting this relation between objective function and performance evaluation allows us now to define an alternative approach to evaluate the usefulness of a candidate metric: we want to judge the quality of a metric by its ability to define an objective function for rate control purposes, and evaluate images compressed in this scheme subjectively. It turns out that deficiencies of image quality metrics become much easier visible — even in the literal sense — than under traditional correlation experiments.

Our candidate metric in this work is the SSIM index [16] proposed by Sheik and Bovik which is both simple enough to be implemented efficiently in rate control algorithms, but yet correlates better to visual quality than MSE; our candidate compression scheme is the highly flexible JPEG 2000 [3] standard.

1. Introduction

The most traditional objective function used in rate allocation algorithms is of course the mean square error (MSE) between reference and distorted image; this is mostly because MSE is simple and mathematically tractable, and many results have been derived in rate distortion theory for this metric. However, it is also known that MSE correlates only poorly to perceived quality; naturally, image compression algorithms employing MSE as their objective function yield sub-optimal subjective quality. For example, it has been observed [25] that MSE optimal JPEG 2000 [3] is often outperformed by traditional JPEG [2] in subjective tests. While a couple of suggestions exist to address this deficiency [5], [1], these proposals are not based on an existing tested quality metric, but rather employ known results from psycho-physics.

In this work, we want to fit the pieces together: Equipping JPEG 2000 with a multi-scale SSIM optimal rate allocator, showing that the resulting images show indeed an improvement in SSIM score, and evaluating the performance of SSIM by judging the quality of the resulting images — very much unlike the traditional approach that would evaluate the correlation of the SSIM score with a subjectively assessed score. We believe that this alternative approach is not only more fruitful in identifying metrics that are more successful for image compression purposes, but we also believe that this approach identifies weaknesses of metrics more easily than a correlation study since it allows the identification of image artifacts tolerated by the metric. This work is in line with the publication of Wang et al [13] where the SPIHT rate allocation had been used to maximize the SSIM, though the intend of the earlier work was less to study the performance of the metric. Finally, the analysis performed in this work is of some interest of its own: We find, similar to other recent works [19] that SSIM is mostly equivalent to more traditional metrics employing known phenomena such as the contrast-sensitivity function [14], [9] or visual masking [6], [11] of errors due to texture.

This work is structured as follows: In the next section, we present a brief overview on the multi-scale SSIM metric. Following that, we describe how rate allocation in the JPEG 2000 standard works. The next section develops a simplification of SSIM that is, within a high-bitrate approximation, equivalent to the original, but is easily implementable as objective function in the JPEG 2000 rate allocation framework. We provide experimental results of our implementation in the following section, presenting both numerical results and some example images, and conclude with an outlook on future works.

2. On Multi-scale SSIM

Very much unlike traditional image quality metrics that implement a physio-visual model of the human visual system, SSIM [15] follows a top-down approach by separating reference and distorted image into luminance, contrast and structure components, and pools their contribution into a single index. Luminance l, contrast c and structure s are defined as follows:

$$l(\vec{x},\vec{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2+\mu_y^2+C_1} \quad c(\vec{x},\vec{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2+\sigma_y^2+C_2} \quad s(\vec{x},\vec{y}) = \frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \; . \tag{1}$$

Here $\vec{x}$ and $\vec{y}$ denote samples in the same local neighborhood of reference and distorted image, μ_x denotes their mean, σ_x their variance and σ_{xy} their covariance, respectively. The constants C_1 to C_3 are small numbers that avoid numerical instabilities for denominators close to zero. For the single-scale SSIM, l, c and s are measured over small image blocks, then multiplied, assigning a local quality index to that block, and the total index is computed by averaging the local index over the complete image area. It is easy to see that this index is symmetric, and equals unity for identical images[1].

In the multi-scale version [16], this procedure is extended by computing the index on scales of the images, each of which is obtained by downscaling the original images by a power of two plus suitable filtering to avoid aliasing. For the multi-scale index, the contribution from scale j is taken to a power γ_j weighting the contribution of that scale, and the product of weighted contributions from all scales gives the full index. The weights γ_j are similar to contrast sensitivity weights — as we shall see — and have been found by extensive subjective tests. We picked the weights from [16]. Since the luminance of all scales is identical by definition, it is sufficient to measure it only at the lowest scale. That is,

$$\text{MS-SSIM}(\vec{x},\vec{y}) = l(\vec{x}_M,\vec{y}_M)^{\gamma_M} \prod_{j=1}^{M} c(\vec{x}_j,\vec{y}_j)^{\gamma_j} s(\vec{x}_j,\vec{y}_j)^{\gamma_j} \; . \tag{2}$$

Similar to the above, $\vec{x}_j$ and $\vec{y}_j$ denote the pixel values of reference and distorted image at scale j.

3. Rate Allocation in JPEG 2000

The JPEG 2000 compression algorithm first decomposes the image by a cascaded discrete wavelet transformation, giving one low-pass band and three differently oriented high-pass bands for each decomposition level. In the Mallat decomposition as used here, only the low-passes are transformed again, leaving the high-passes alone. Each band b is separated into so-called *code-blocks* c which are encoded independently by the EBCOT [3] algorithm: all coefficients from each codeblock are scanned bitplane by bitplane, and each bitplane is encoded in three coding passes. In each coding pass, the distortion $D_{j,b,c}(t_{j,b,c})$ and rate $R_{j,b,c}(t_{j,b,c})$ are measured, giving the rate distortion R over D of this codeblock, where t labels the coding pass and thus potential truncation positions of the encoding procedure.

Even though the JPEG 2000 standard does not require this, most implementations define the distortion as the mean square error between the image area identified with the codeblock in in the reference image, and the image the decoder would reconstruct from the data it received up to the truncation position t; rate is given by the output stream size of the entropy coding back-end.

1. SSIM is *not* a metric in the mathematical sense, though. The triangle inequality does not hold.

Rate allocation tries now to optimize the total distortion, that is the sum of all local contributions, under a total rate constraint. In a Lagrangian formulation, this is equivalent of finding stationary points of the following functional in the encoding truncation positions $t_{j,b,c}$:

$$J := \sum_{j,b,c} D_{j,b,c}(t_{j,b,c}) + \lambda \sum_{j,b,c} R_{j,b,c}(t_{j,b,c}) \,. \tag{3}$$

Since the contribution to rate and distortion in codeblock (j, b, c) depend only on the fractional bitplane where encoding stops in this bitplane, i.e. on $t_{j,b,c}$, one can now, by a short computation, identify λ with the *critical* slope of all rate distortion curves; J becomes then stationary if and only if the cut-off points $t_{j,b,c}$ are chosen such that the slope of all rate distortion curves at $t_{j,b,c}$ become equal. The target rate can then be hit by a bisection algorithm over λ. For details, we refer to [3], [4].

4. EBCOT Rate Allocation for SSIM

The first step in implementing a SSIM optimal rate allocation in JPEG 2000 is noting that the SSIM is multiplicative by nature, but eqn. (3) requires an additive distortion measure. However, this problem is easily resolved by noting that instead of optimizing MS-SSIM directly, one can optimize its logarithm as well. Naturally, we pick the block size of the SSIM index equal to the codeblock size of the JPEG 2000 encoder, which also means that the sliding window average of the SSIM index is replaced by the non-overlapping average over the codeblocks. Note that even though codeblocks are typically larger than the window size of the SSIM, their size can be adjusted by a coding option. The image scales of the SSIM computation are of course identified with the wavelet low-pass bands containing downscaled versions of the image.

The logarithm of the index then splits into a sum of contributions from all scales of the image, with multiplicative weights γ_j for each scale:

$$\begin{aligned} \log\big(\text{MS-SSIM}(\vec{x},\vec{y})\big) \quad &= \quad \gamma_M \log\left(\frac{1}{N_M}\sum_{c\in C_M} l(\vec{x}_{M,c},\vec{y}_{M,c})c(\vec{x}_{M,c},\vec{y}_{M,c})s(\vec{x}_{M,c},\vec{y}_{M,c})\right) \\ &\quad + \sum_{j=1}^{M-1}\gamma_j \log\left(\frac{1}{N_j}\sum_{c\in C_j} c(\vec{x}_{j,c},\vec{y}_{j,c})s(\vec{x}_{j,c},\vec{y}_{j,c})\right)\,, \end{aligned} \tag{4}$$

where the inner sum over c is the sum over codeblocks, and N_j is the number of codeblocks in scale j.

In the next step we note similar to other authors [17], [19] that for $C_2 = C_3$ the product of structure and contrast term can be expressed in terms of the MSE and a denominator that looks surprisingly similar to the visual masking model proposed by Zeng and Daly [1], here however with a larger exponent of 2:

$$\begin{aligned} 2M\sigma_{xy} \quad &= \quad 2\langle \vec{x}-\mu_x, \vec{y}-\mu_y\rangle = M\sigma_x^2 + M\sigma_y^2 - \|(\vec{x}-\mu_x)-(\vec{y}-\mu_y)\|^2 \\ \Rightarrow \quad c(\vec{x}_{j,c},\vec{y}_{j,c})s(\vec{x}_{j,c},\vec{y}_{j,c}) \quad &= \quad 1 - \frac{\|(\vec{x}_{j,c}-\mu_{x,j,c})-(\vec{y}_{j,c}-\mu_{y,j,c})\|^2}{\|\vec{x}_{j,c}-\mu_{x,j,c}\|^2 + \|\vec{y}_{j,c}-\mu_{y,j,c}\|^2 + C_2 M_j}\,, \end{aligned} \tag{5}$$

where M_j is the number of coefficients in a block at scale j, i.e. the number of coefficients of which the variance resp. covariance is computed.

This expression depends on the reconstruction error in the numerator, which is directly available in the JPEG 2000 EBCOT rate allocator [5], but also on the reconstructed coefficients itself. Furthermore, the logarithms do not allow an easy error pooling as required by the Lagrangian functional, eqn. (3). We therefore make now the assumption that the reconstructed image is close to the reference image, i.e. $\vec{x} - \vec{y}$ is rather small or the image quality is high. Under these assumptions, the fraction in eqn. (5) is close to zero because the numerator is small. We get for the first order linear term of its Taylor series the following expression:

$$\begin{aligned} c(\vec{x}_{j,c},\vec{y}_{j,c})s(\vec{x}_{j,c},\vec{y}_{j,c}) \quad &\approx \quad 1 - \frac{\|(\vec{x}_{j,c}-\mu_{x,j,c})-(\vec{y}_{j,c}-\mu_{y,j,c})\|^2}{2\|\vec{x}_{j,c}-\mu_{x,j,c}\|^2 + C_2 M_j} =: 1 - D_{j,c} \\ \Rightarrow \qquad \log\left(1 - \frac{1}{N_j}\sum_{c\in C_j} D_{j,c}\right) \quad &\approx \quad -\frac{1}{N_j}\sum_{c\in C_j} D_{j,c} \end{aligned} \tag{6}$$

It is readily calculated that the luminance term becomes unity in first order, and it can be simply ignored for our purposes. The fact that the luminance term has only minor impact on the overall performance of the SSIM metric has also been noted by Rouse and Hemami [18] in their experiments and is thus also justified heuristically.

One issue remains, namely that the numerator of the approximated SSIM distortion $D_{j,c}$ is to be computed in the *scale* j of the image domain, whereas the JPEG 2000 EBCOT process operates on the *wavelet high-passes*. To overcome this problem, we note that the error in scale j can be estimated by the sum of all errors made in the bands contributing to this scale in the reconstruction process, times a weight given by the inverse impulse response of the filter that reconstructs the scale from its sub-bands. A similar approximation is traditionally used in MSE optimal JPEG 2000 rate allocation processes; there, however, the inverse impulse response is computed in the image domain, i.e. for scale $j = 1$. This approximation, while only exact for orthogonal filters, is justified as follows: Denote by $T^{j,b}$ the wavelet reconstruction filter from band b to scale j with $(T^{j,b}\vec{x})_i = \sum_k t^{j,b}(i-k)x_k$, i the spatial location of the coefficient, and assume that $\vec{x}$ in the source band b is a representative of a decorrelated i.i.d. process with zero mean. Then the expectation of the MSE in scale j is given by

$$\mathbb{E}\big(\|T_{j,b}\vec{x}\|^2\big) = \sum_{k,l}\left(\sum_i t_{j,b}(i-k)t_{j,b}(i-l)\right)\mathbb{E}(x_k x_l)\ , \tag{7}$$

but the last term vanishes for $l \neq k$ because x is decorrelated, and we have

$$\mathbb{E}\|T_{j,b}\vec{x}\|^2 = \sum_k\sum_i t_{j,b}(i-k)t_{j,b}(i-k)\mathbb{E}(x_k^2) = \sum_i t_{j,b}(i)^2\,\mathbb{E}(\|\vec{x}\|^2)\ = \hat{\gamma}_{j,b}^2\mathbb{E}(\|\vec{x}\|^2) \tag{8}$$

as claimed, with a factor $\hat{\gamma}_{j,b}^2$ that depends on the source wavelet band b and the target scale j only.

Putting everything together, we have the following expression for the (approximate) SSIM distortion in terms of the MSE and the wavelet scale coefficients:

$$\log\big(\text{MS-SSIM}(\vec{x},\vec{y})\big) \approx -\sum_{j,b,c}\frac{\gamma_j\hat{\gamma}_{j,b}^2 M_j}{2N_j M_b}\cdot\frac{\|\vec{x}_{b,c}-\vec{y}_{b,c}\|^2}{\|\vec{x}_{j,c}-\mu_{x,j,c}\|^2+\frac{C_2 M_j}{2M_b}}\ , \tag{9}$$

where we again simplified the numerator by assuming that the mean values of reference and distorted image are approximately identical.

This expression is now suitable for a modified JPEG 2000 rate allocation process, and also admits a very nice interpretation: The numerator of the fraction is the classical MSE term already computed on the fly by the EBCOT rate allocation process in the band b and codeblock c. The denominator is a masking term which has to be computed in the scale j, i.e. in the downscaled image domain and not in the wavelet high-pass that gets encoded. Luckily, this term can be computed upfront in the wavelet decomposition process and remains constant within the encoder. The factor $\gamma_j\hat{\gamma}_{j,b}^2 M_j/2N_j M_b$, finally, is equivalent to a fixed contrast sensitivity masking term, and contains both the relative amplitude gain due to the wavelet filtering process, as well as the band visibility term obtained by Bovik et al obtained by MS-SSIM calibration [16] with subjective tests.

5. Experiments

As explained in the introduction, the focus of this work is to evaluate the performance of SSIM as the objective function of a rate allocation. To this end, we modified the EBCOT rate allocation procedure of the JJ2000 implementation [22] to measure eqn. (9) instead of the MSE. We compressed images from the JPEG image test core with the modified implementation, the unmodified MSE optimal JJ2000 implementation and a visually optimized version provided by Pegasus Imaging which both employs visual weighting and visual masking [26]. To have block sizes comparable to the 11×11 blocks suggested for SSIM, we also set the code block size to 8×8 consistently for all three implementations, though the results do not depend overly on the particular choice. Images then have been compressed to bit-rates between 0.5 and 1.75 bits per pixel. The quality of the resulting images is then estimated: To test the correctness of the implementation, we measured the multi-scale SSIM on all images, expecting that the SSIM-optimal implementation would outperform the two other codecs if the implementation is proper.

We also measured the PSNR, and the VDP score, defined as the relative number of pixels the standard observer of Daly's Visual Difference Predictor [8] would distinguish by a probability of at least 75%. We also applied an experimental DCT based metric, mDCT-PSNR, by the author [25] as an additional cross check. Unfortunately, we do not have the lab equipment to run subjective tests at the University of Stuttgart, but we reproduce some of the images here to allow the reader to judge the output himself and to compare it with our conclusions.

bpp	0.5	0.75	1.0	1.25	1.5	1.75	0.5	0.75	1.0	1.25	1.5	1.75
	$-20\log_{10}(1-\text{SSIM})$						$-20\log_{10}(\text{VDP})$					
PSNR-optimal	25.61	29.43	33.02	35.76	38.13	40.40	15.45	24.50	33.72	44.58	56.48	64.44
Visually	26.78	30.90	34.48	36.66	39.62	41.89	18.53	29.76	40.92	51.06	64.44	73.98
SSIM-optimal	28.31	34.35	38.41	41.08	43.22	45.10	19.00	28.00	34.52	41.21	48.41	57.08
PSNR-optimal	25.02	28.71	32.14	35.44	38.64	41.16	10.70	16.02	23.14	30.87	43.88	55.92
Visually	27.03	30.99	34.78	37.91	40.64	43.29	13.94	21.87	32.77	44.29	55.92	70.46
SSIM-optimal	29.67	35.88	39.73	42.60	45.04	47.08	16.08	22.96	29.71	37.59	45.52	53.98
PSNR-optimal	31.24	35.49	38.02	39.94	41.34	42.72	26.96	45.35	61.94	73.98	inf	inf
Visually	32.24	36.14	38.91	40.60	42.62	44.61	32.32	51.70	70.46	80.00	inf	inf
SSIM-optimal	34.07	39.12	41.94	43.77	45.06	45.87	31.28	39.66	45.85	52.40	60.00	64.44
PSNR-optimal	28.98	33.33	36.88	39.75	42.16	44.42	13.21	21.46	29.35	40.00	51.06	61.94
Visually	30.18	34.58	38.20	41.04	43.48	45.41	17.63	27.87	39.50	52.04	66.02	80.00
SSIM-optimal	32.93	39.22	42.73	45.23	47.09	48.42	19.58	26.80	33.81	42.98	54.90	67.96

Table 1. SSIM scores (left) and VDP scores (right) for the PSNR optimal, the visually tuned and the SSIM optimal implementation, measured on images from the JPEG test core: Oahu_Northcoast, Zoo, Rokunji and P04 (top to bottom).

Block Size	PSNR	SSIM	VDP
8×8	31.35	39.73	29.71
16×16	32.52	40.47	34.02
32×32	33.21	40.52	36.60
64×64	33.44	40.64	37.46

Table 2. PSNR, SSIM and VDP scores for the SSIM optimal implementation measured on the Zoo image at 1 bpp. As seen, the JPEG 2000 coding overhead for lower code blocks affects scores approximately equal for all metrics.

6. Results and Discussion

The plots in Fig. 1 and the numerical results in Table 1 show that the SSIM score is indeed consistently higher for the implementation using the approximated SSIM implementation as objective function; hence, the proposed linear version of the SSIM is indeed close, or close enough to the exact version.

The results are *not* consistent for other metrics, typically the SSIM optimized version performs worst under evaluation of other metrics. It is also no surprise that the visually optimized JPEG2000 code performs better in terms of the mDCT-PSNR metric because the visual masking model of the metric and the objective function of this implementation are related; both are based on the Zeng and Daly model [1].

The resulting subjective image quality is also debatable, and depends on the nature of the image region. Figures 2,3 show some typical results, again on an example image from the JPEG image test core: The first column of images has been compressed with the SSIM optimized version of the JJ2000 code, the second is a visually optimized code by the author [26].

In some image areas, specifically those with high contrast edges as the telegraph wires in the first pair of images, Gibbs phenomena are much more visible in the SSIM optimized version, and the overall impression here is that many high frequencies have been dropped. However, the SSIM optimized version performs better in textured image areas, very notable in the cypress trees in the lower row of figure 2.

Table 2 shows the scores for the zoo image coded with the SSIM optimal implementation at 1 bpp depending on the code block size. As it is seen here, the coding overhead required for the smaller code blocks causes a degration in the image quality that is mostly consistent over all metrics. The choice of

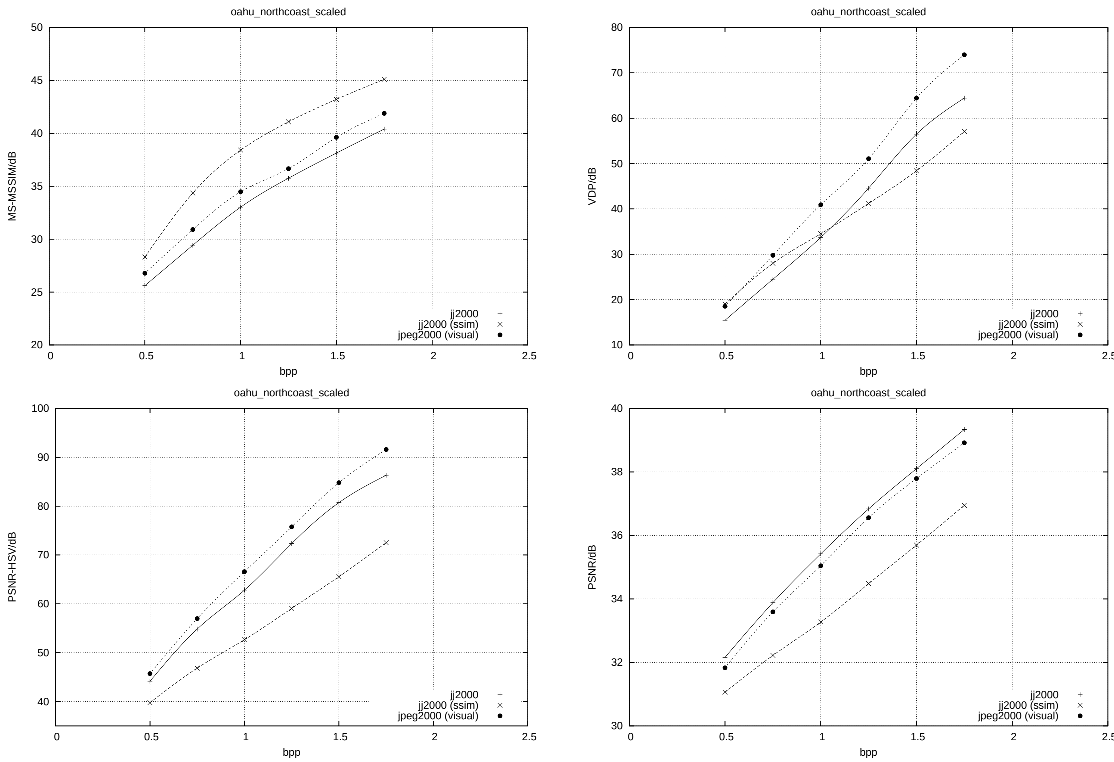

Figure 1. Objective image quality of the Oahu_Northcoast image, measured by the MS-SSIM itself, VDP, and mDCT-PSNR, an experimental DCT based metric recently proposed by the author [25], and PSNR (from left to right, top to bottom). The SSIM plot shows $-20\log_{10}(1-\text{SSIM})$ instead of the index itself, making the plots easier to compare.

8×8 code blocks to approximate the 11×11 block size of SSIM seems to be rather irrelevant compared to the encoding overhead, thus for any practical purposes chosing larger code blocks is beneficial.

7. Complexity of the Algorithm

It is not the aim of this work to present precise results on the complexity of the SSIM based rate allocation, but we can at least give a fair estimate on the overhead: First of all, the variances of the coefficients in all low-passes must be computed. While we currently simply buffer all data, a more realistic implementation would perform this operation within the wavelet filter stage, requiring one extra addition and one multiplication per pixel. The modification of the rate allocation stage itself is rather trivial and requires only a couple of look-up tables for the band-dependent factors, and the collection of variance data from all image scales the codeblock contributes to. This operation is, however, only required once per coding pass per codeblock, and shouldn't contribute much to the overall complexity. We believe that this is a rather mild overhead, and comparable to other techniques employing some type of masking, e.g. [5], [1].

8. Conclusion and Outlook

The aim of this work was less to implement an SSIM optimal JPEG 2000, but more to estimate the quality of a metric by the quality of images generated by employing the metric as objective function. Our results show that the SSIM metric is not completely consistent with other metrics, though it is consistent with its approximation made for purposes of rate allocation. The SSIM is simple and efficient enough

Figure 2. Two image regions from the Oahu_Northcoast image, left side is the SSIM optimal JJ2000 implementation, right side is the visually optimized JPEG2000 code, both compressed to 0.75 bpp. Note that the telegraph wires show Gibbs artifacts on the left, while the cypresses are blurry on the right.

to be deployed in practical applications, but shows potential for improvement should the aim be optimal subjective quality. Then, however, extensive subjective tests must be performed.

A notable side-result of our work is that the approximations made here show clearly the relation of the SSIM to traditional masking approaches by Daly and Watson [6], [7], though the masking exponent — the power in the denominator of eqn. 9 — is somewhat different to that of traditional masking strength estimates. This might explain some of the deficiencies of the SSIM index. This equivalence of course not a new result and is very much in line with other works [19] on the same issue.

Performances of metrics, often only expressed in terms of correlation coefficients, become visible by our approach, also in a very literal sense: For the SSIM, one deficiency seems to be that contributions of high frequencies are underestimated, and the effect of masking on razor-edges seems to be overestimated.

We will try to drive this research program further in future works, namely understanding the nature of metrics by employing them as objective functions. One desirable extension would be to have a SSIM-like metric for color images; to our knowledge, existing proposals use a simple weighting matrix in YCbCr color space, ignoring that the contrast sensitivity function for chroma differs from that for luma. These approaches should also be refined to use a perceptually more appropriate color space (e.g. IPT [23], [24]).

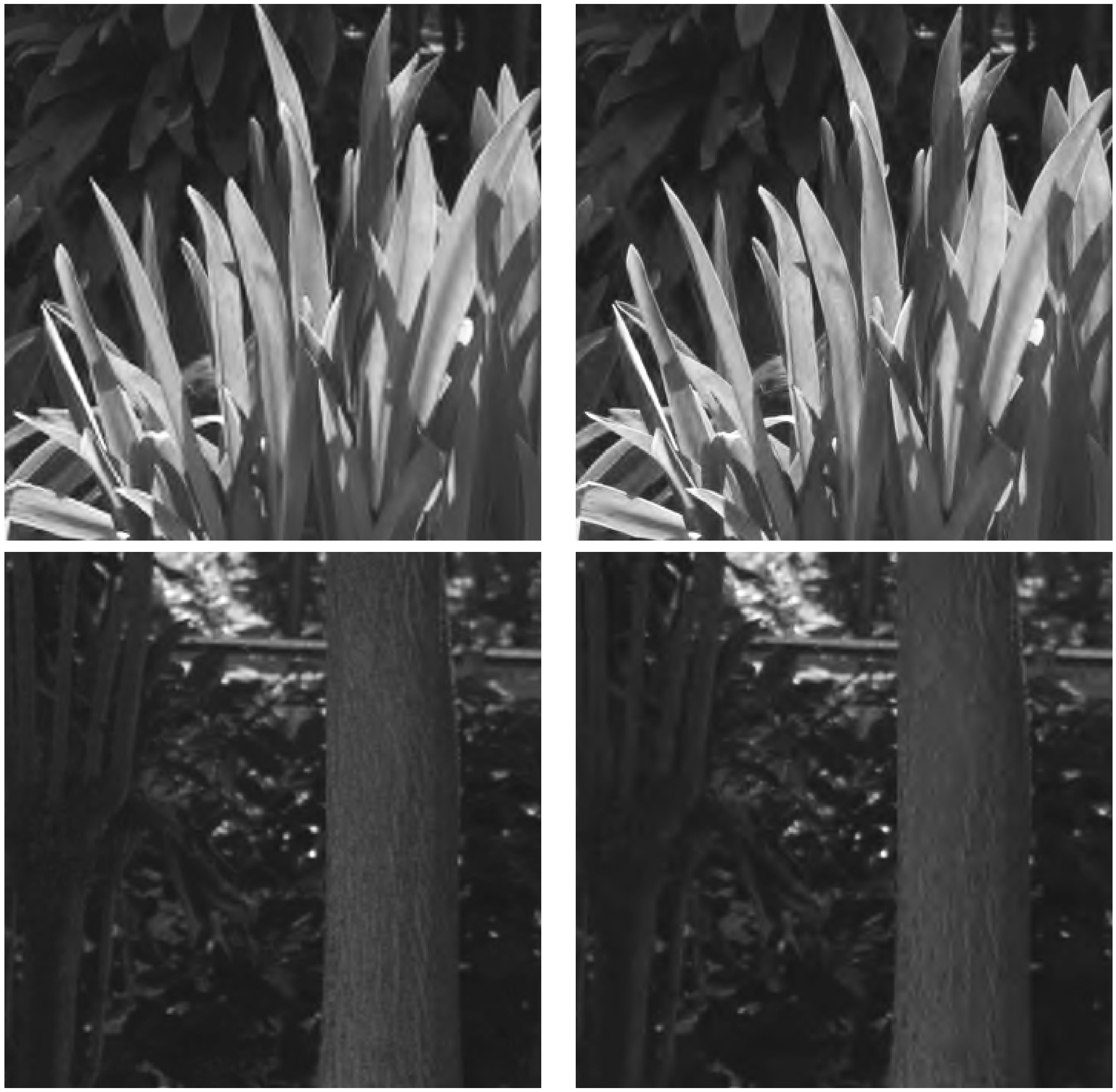

Figure 3. Two image regions from the Zoo image, left side is the SSIM optimal JJ2000 implementation, right side is the visually optimized JPEG2000 code, both compressed to 0.75 bpp. Blurring artifacts on the plant in the top row for the SSIM optimal version, the reverse situation for the tree in the second image pair.

We will also extend this work in the direction of more advanced metrics than SSIM, for example towards the recently proposed VIF index by Bovik et al [20], [21].

9. Acknowledgements

Kil Joong Kim acknowledges the support of the DAAD for funding his stay in Stuttgart during the summer 2008. Thomas Richter thanks Pegasus Imaging for the long term support and cooperation, and for providing their code-base for the purpose of this study. He also acknowledges the support of the Computing Center and the Institute for IT Services, Prof. Sabina Jeschke, at the University of Stuttgart.

10. Bibliography

References

[1] W. Zeng; S. Daly; S. Lei: "Point-wise extended visual masking for JPEG-2000 image compression" *IEEE Proc. of Intl. Conference on Image Processing*, Vol. 1, pp. 657 - 660, (2000)

[2] W. Pennebaker, J. Mitchell: "JPEG Still Image Data Compression Standard" *Chapman & Hall, New York* (1993)

[3] M. Boliek (Ed.): "Information Technology — The JPEG2000 Image Coding System: Part 1", *ISO/IEC IS 15444-1*, (2000)

[4] D.S. Taubman, M.W. Marcellin: "JPEG2000 Image Compression Fundamentals, Standards and Practice", *Springer, New York* (2002)

[5] D. Taubman: " High performance scalable image compression with EBCOT", *IEEE Transactions on Image Processing*, Vol. 9 No. 7, pp. 1151-1170, (2000)

[6] S. Daly: " Application of a noise-adaptive contrast sensitiy function to image data compression", *Optical Engineering*, vol. 29, pp. 977-987, (1990)

[7] A.B. Watson: "Perceptual Optimization of DCT Color Quantization Matrices", *Proc. of 1994 IEEE Intl. Conf. on Image Processing ,ICIP-94*, pp. 100-104, (1994).

[8] R. Mantiuk, S. Daly, K. Myszkowski, H.P. Seidel: "Predicting Visible Differences in High Dynamic Range Images - Model and its Calibration", *Proc. of Human Vision and Electronic Imaging X, IS&T/SPIE's 17th Annual Symposium on Electronic Imaging 2005* pp. 204-214, (2005)

[9] Watson, Yang, Solomon, Vilasenor: "Visibility of wavelet quantization noise", *IEEE Trans. Image Proc.*, Vol. 6, No. 8, pp. 1164-1175, (1997)

[10] A. Mazzari, R. Leonardi: "Perceptual embedded image coding using wavelet transforms", *Proc. IEEE Int. Conf. Image Proc.*, Vol 1, pp. 586-589, (1995)

[11] I. Hontsch, L.J. Karam: "APIC: adaptive perceptual image coding based on subband decomposition with locally adaptive perceptual weighting", *Proc. of Intl. Conf. on Image Proc. 1997*, Vol. 1, pp. 37-40, (1997)

[12] A. Reibman, D. Poole: "Quality assessment for super-resolution image enhancement", *Proc. of 2006 13th IEEE Intl. Conf. on Image Proc, ICIP 2006*, (2006)

[13] Z. Wang, Q. Li, X. Shang: "Perceptual image coding based on a maximum of minimal Structural Similarity Criterion", *Proc. of 2007 14th IEEE Intl. Conf. on Image Proc, ICIP 2007*, (2007)

[14] A.N. Netravali, B.G. Haskel: "Digital Pictures - Representation and Compression", *Plenum Publishing Corp.*, USA, (1988)

[15] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli: "Image quality assessment: From error visibility to structural similarity", *IEEE Trans. on Image Proc.*, Vol. 13, No. 4, pp. 600-612, (2004).

[16] Z. Wang, E.P. Simoncelli, A.C. Bovik: "Multi-scale structural similarity for image quality assessment", *IEEE Asilomar Conf. on Signals, Systems and Computers*, (2003).

[17] S.S. Channappayya, A.C. Bovik, R.W. Heath Jr., C.S. Caramanis: "Rate Bounds on SSIM Index of Quantized Image DCT Coefficients", *Proc. of Data Compression Conference 2008*, (2008)

[18] D.M. Rouse, S.S.Hemami: "Understanding and Simplifying the Structural Similarity Index", *Proc. of 2008 15th IEEE Intl. Conf. on Image Proc, ICIP 2008*, (2008)

[19] K. Seshadrinathan, A. Bovik: "Unifying Analysis of Full Reference Image Quality Assessment", *Proc. of 2008 15th IEEE Intl. Conf. on Image Proc, ICIP 2008*, (2008)

[20] H.R. Sheikh, A.C. Bovik, G. de Veciana: "An information fidelity criterion for image quality assessment using natural scene statistics", *IEEE Trans. Image Process.*, Vol. 13, No. 4, pp. 600-612, (2004)

[21] H.R. Sheikh, A.C. Bovik: "Image information and visual quality", *IEEE Trans. Image Process.*, Vol. 15, No. 2, pp. 430-444, (2006)

[22] R. Grosbois, D. Santa Cruz, J. Askelöf, B. Berthelot, D. Bouchard, F. Henry, G. Mozelle, P. Onno: "JJ2000: Java implementation of JPEG 2000", version 5.1 available for download at http://jpeg2000.epfl.ch/, (2000)

[23] F.Ebner, M.D. Fairchild: "Development and Testing of a color space(IPT) with improved hue uniformity", *IS&T/SID 6th Color Imaging Conference*, pp. 8-13 (1998)

[24] M.D. Fairchild, G.M. Johnson: "The iCAM framework for image appearance, image differences, and image quality", *J. of Electronic Imaging*, (2004)

[25] Th. Richter, C. Larabi: "Towards object image quality metrics: the AIC eval program of the JPEG", in: *Proc. SPIE: Applications of Digital Image Processing XXXI*, A.G. Tescher(Ed.), Vol. 7073, No. 34, (2008)

[26] Th. Richter: "Effective Visual Masking Techniques in JPEG2000", *Proc. of 2008 15th IEEE Intl. Conf. on Image Proc, ICIP 2008*, (2008)