Some "Weberized" L^2-Based Methods of Signal/Image Approximation

Ilona A. Kowalik-Urbaniak[1], Davide La Torre[2,3], Edward R. Vrscay[1(✉)], and Zhou Wang[4]

[1] Department of Applied Mathematics, Faculty of Mathematics, University of Waterloo, Waterloo, ON N2L 3G1, Canada
{iakowali,ervrscay}@uwaterloo.ca
[2] Department of Applied Mathematics and Sciences, Khalifa University, Abu Dhabi, United Arab Emirates
[3] Department of Economics, Management and Quantitative Methods, University of Milan, Milan, Italy
davide.latorre@kustar.ac.ae, davide.latorre@unimi.it
[4] Department of Electrical and Computer Engineering, Faculty of Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
zhouwang@ieee.org

Abstract. We examine two approaches of modifying L^2-based approximations so that they conform to Weber's model of perception, i.e., higher/lower tolerance of deviation for higher/lower intensity levels. The first approach involves the idea of intensity-weighted L^2 distances. We arrive at a natural weighting function that is shown to conform to Weber's model. The resulting "Weberized L^2 distance" involves a ratio of functions. The importance of ratios in such distance functions leads to a consideration of the well-known logarithmic L^2 distance which is also shown to conform to Weber's model.

In fact, we show that the imposition of a condition of perceptual invariance in greyscale space $\mathbb{R}_g \subset \mathbb{R}$ according to Weber's model leads to the unique (unnormalized) measure in $\mathbb{R}_g$ with density function $\rho(t) = 1/t$. This result implies that the logarithmic L^1 distance is the most natural "Weberized" image metric. From this result, all other logarithmic L^p distances may be viewed as generalizations.

1 Introduction

In this paper we examine some methods of modifying, or "Weberizing," L^2-based approximations so that they conform as much as possible to Weber's model of perception. The term "Weberized" has been used in recent papers which have incorporated Weber's model into classical image processing methods, namely, total variation (TV) restoration [5] and Mumford-Shah segmentation [6].

For a long time, it has been recognized that the well known and very commonly used mean squared error (MSE) and PSNR – examples of L^2-based measures – perform poorly in terms of perceptual image quality [2,8]. Nevertheless,

A. Campilho and M. Kamel (Eds.): ICIAR 2014, Part I, LNCS 8814, pp. 20–29, 2014.
DOI: 10.1007/978-3-319-11758-4_3

L^2-based methods are still employed to a large degree, most probably due to their relative simplicity of computation. Other perceptually more meaningful image quality measures are generally more difficult to optimize. The Weberized L^2 methods examined in this paper are quite straightforward to compute.

That being said, the structural similarity (SSIM) image quality measure [8,9], which has demonstrated a superior performance in comparison with traditional quality measures such as MSE and PSNR, already has a "Weberized" component, namely, the luminance term, denoted as $S_1(\mathbf{x}, \mathbf{y})$, which characterizes the similarity between mean values, $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$, of image patches/blocks $\mathbf{x}$ and $\mathbf{y}$, respectively. The fact that $S_1(\mathbf{x}, \mathbf{y})$ may be expressed as a function of the ratio $\mathbf{x}/\mathbf{y}$ (or $\mathbf{y}/\mathbf{x}$) accounts for its "Weberized" form.

Let us first recall Weber's model of perception which, for simplicity of treatment, will be restricted to the case of greyscale images: Given a greyscale background intensity $I > 0$, the minimum change in intensity ΔI perceived by the human visual system (HVS) is related to I as follows,

$$\frac{\Delta I}{I} = C, \tag{1}$$

where C is constant, or at least roughly constant over a significant range of intensities I [7]. Eq. (1) suggests that the HVS will be less/more sensitive to a given change in intensity ΔI in regions of an image at which the local image intensity $I(x)$ is high/low. As such, a Weberized L^2 distance between two functions u and v should tolerate greater/lesser differences over regions in which they assume higher/lower intensity values.

The basic mathematical ingredients of our formalism are as follows:

1. The **base (or pixel) space** $X \subset \mathbb{R}$ on which our signals/images are supported. Here, we assume, without generality, that $X = [0, 1]$. For images, $X = [0, 1]^2$. In the case of digital images, X can be the set of pixel locations (i, j), $1 \le i \le n_1$, $1 \le j \le n_2$.
2. The **greyscale range** $\mathbb{R}_g = [A, B] \subset (0, \infty)$.
3. The **signal/image function space** $\mathcal{F} = \{u : X \to \mathbb{R}_g\}$. Note that from our definition of the greyscale range $\mathbb{R}_g$, $u \in \mathcal{F}$ is positive and bounded, i.e., $0 < A \le u(x) \le B < \infty$ for all $x \in X$. A consequence of this boundedness is that $\mathcal{F} \subset L^p(X)$ for all $p \ge 1$, where the $L^p(X)$ function spaces are defined in the usual way. For any $p \ge 1$, the L^p norm can be used to define a metric d_p on $\mathcal{F}$: For $u, v \in \mathcal{F}$, $d_p(u, v) = \|u - v\|_p$. Our primary concern is the approximation of functions in the case $p = 2$, i.e., the Hilbert space, $L^2(X)$. In this case, the distance between two functions $u, v \in L^2(X)$ is given by

$$d_2(u, v) = \|u - v\|_2 = \left[\int_X [u(x) - v(x)]^2 \, dx \right]^{1/2}. \tag{2}$$

2 The Use of Intensity-Dependent Weighting Functions

The approximation of signals and images – and functions in general – must involve some mesaurement of "distance," as determined by an appropriate

metric. In the usual L^2-based methods of approximation employed in signal and image processing, the L^2 metric in Eq. (2) is used. This metric, and indeed all other L^p- based metrics, $p \geq 1$, are not adapted to Weber's model of perception since they involve integrations over appropriate powers of intensity differences, $|u(x) - v(x)|$, with no consideration of the magnitudes of $u(x)$ or $v(x)$.

One way to "Weberize" this metric is to insert a weighting function in the integrand of Eq. (2). The use of weighting functions in metrics is, of course, not a new idea. In mathematics, they have generally been functions of the independent variable – in this case, the spatial variable x. In image processing applications, they have been employed for spatial weighting, for example, in foveated or region-of-interest image processing and coding [3] or frequency weighting in perceptual image quality assessment [10]. In our application, the weighting function should be dependent upon one or both of the intensities of the image functions $u(x)$ and $v(x)$. As such, the weighted L^2 metric may be written in the generic form,

$$d_{2W}(u,v) = \left[\int_X g(u(x),v(x))[u(x)-v(x)]^2\,dx\right]^{1/2}, \tag{3}$$

where $g : \mathbb{R}_g \times \mathbb{R}_g \to \mathbb{R}_+$ denotes the intensity-dependent weighting function.

This leads to an interesting set of questions regarding the properties that must be satisfied by the weighting function g as well as the possible functional forms that it may assume, keeping in mind two important requirements:

1. $d_{2W}(u,v)$ should, if possible, satisfy the mathematical properties of a metric,
2. $d_{2W}(u,v)$ should, in some way, conform to Weber's model of perception.

A detailed discussion of these questions, many of which represent open problems, is well beyond the scope of this paper.

Perhaps one of the most fundamental properites that must be satisfied in order that Weber's model of perception can be accommodated is that $g(u,v)$ be decreasing in both of its arguments. This requirement is satisfied, for example, by the symmetric family of functions, $g(u(x),v(x)) = |u(x)v(x)|^{-q}$, where $q > 0$.

A simplification is achieved if we consider g to be a function of only one intensity function. Furthermore, if we assume that $g(u(x),v(x)) = g(u(x)) = u(x)^{-2}$, then the weighted L^2 distance in Eq. (3) becomes

$$d_{2W}(u,v) = \left[\int_X \left[1-\frac{v(x)}{u(x)}\right]^2 dx\right]^{1/2} =: \Delta(u,v). \tag{4}$$

In this case, we consider the function u, which defines the weighting function g, to be the *reference function*. If we then consider v to be an approximation to u, then $\Delta(u,v)$ in Eq. (4) is the approximation error.

If we assume a weighting function of the form $g(u(x),v(x)) = g(v(x)) = v(x)^{-2}$, the weighted L^2 distance in Eq. (4) becomes

$$d_{2W}(u,v) = \left[\int_X \left[1-\frac{u(x)}{v(x)}\right]^2 dx\right]^{1/2} =: \Delta(v,u). \tag{5}$$

Note that in general, $\Delta(u,v) \neq \Delta(v,u)$, which implies that Δ is not a metric in the strict mathematical sense of the term. This is the price paid for employing weighting functions $g(x)$ which are not symmetric in the functions $u(x)$ and $v(x)$. This complication, however, is not a serious limitation because of the following results that apply to our space $\mathcal{F}$ of image functions.

Theorem 1: Let $u, v \in \mathcal{F}$, with the assumption that the greyscale range $[A,B]$ is bounded away from zero, i.e., $A > 0$. Then

$$\frac{1}{B} d_2(u,v) \leq \left\{ \begin{matrix} \Delta(u,v) \\ \Delta(v,u) \end{matrix} \right\} \leq \frac{1}{A} d_2(u,v)\,, \tag{6}$$

where d_2 denotes the L^2 metric in Eq. (2) from which it follows that

$$\left[2 - \frac{B}{A}\right] \Delta(u,v) \leq \Delta(v,u) \leq \frac{B}{A} \Delta(u,v)\,. \tag{7}$$

The proofs are rather straightforward and will be omitted.

A consequence of the above Theorem is that it it is sufficient to consider only one of these two distance functions, which will be the approach adopted for the remainder of this paper. Unless otherwise stated, the function u will be the *reference function* and v an approximation to it, in which case the approximation error will be given by $\Delta(u,v)$ in Eq. (4).

From Eq. (4), we see that for $\Delta(u,v)$ to be small, the ratio $v(x)/u(x)$ must be close to 1 for all $x \in X$. This already suggests that Weber's model of perception is being followed: Larger values of $u(x)$ will tolerate larger deviations between $v(x)$ and $u(x)$ so that the ratio $v(x)/u(x)$ is kept within a specified distance from 1. The following simple example illustrates this.

Example 1: Consider the "flat" reference image $u(x) = I$, where $I \in \mathbb{R}_g$. Now let $v(x) = I + \Delta I$, with $\Delta I > 0$, be the constant approximation to $u(x)$, where $\Delta I = CI$ is the minimum perceived change in intensity corresponding to I, according to Weber's model in Eq. (1). The L^2 distance between u and v is

$$d_2(u,v) = K \cdot \Delta I = KCI\,, \text{ where } K = \left[\int_X dx\right]^{1/2}. \tag{8}$$

A simple computation shows that the weighted L^2 distance in Eq. (4) is

$$\Delta(u,v) = K \frac{\Delta I}{I} = KC\,. \tag{9}$$

The L^2 distance in Eq. (8) increases with the intensity level I. This is expected since ΔI increases with I. However, the weighted L^2 distance in Eq. (9) remains constant. As such, we claim that $\Delta(u,v)$ can better accommodate Weber's model of perception: Perturbations ΔI of image intensities I according to Eq. (1) yield the same distance measure.

3 Best Approximation in Terms of $\Delta(u, v)$

Firstly, let $\{\phi_k\}_{k=1}^{\infty}$ denote a set of real-valued functions that form a complete orthonormal basis of $L^2(X)$, i.e., $\langle \phi_i, \phi_j \rangle = \delta_{ij}$, where δ_{ij} denotes the usual Kronecker delta. Now let $u \in \mathcal{F} \subset L^2(X)$ denote the reference signal/image function to be approximated. Given an $N > 0$, we are interested in best approximations of the form

$$u \approx u_N = \sum_{k=1}^{N} c_k \phi_k \,. \tag{10}$$

As is well known, the best L^2 approximation to u, which is the *minimizer* of the L^2 distance $\|u - u_N\|_2$, is yielded by the *Fourier coefficients* of u in the $\{\phi_k\}$ basis, i.e.,

$$c_k = \langle u, \phi_k \rangle = \int_X u(x)\phi_k(x)\, dx, \quad 1 \le k \le N \,. \tag{11}$$

We now wish to determine the "best Weberized" approximation, i.e., the expansion in Eq. (10) that minimizes the weighted L^2 distance $\Delta(u, u_N)$. For simplicity, we consider the squared distance $\Delta^2(u, u_N)$,

$$\Delta^2(u, u_N) = \int_X g(x) \left[u(x) - \sum_{k=1}^{N} c_k \phi_k(x) \right]^2 dx =: f(c_1, c_2, \cdots, c_N) \,. \tag{12}$$

Here, the weighting function is $g(x) = 1/u(x)^2$ but the algebraic expressions presented below apply to any weighting function $g(x)$.

Imposition of the stationarity constraints $\frac{\partial f}{\partial c_k} = 0$, $1 \le k \le N$, yields a linear system of equations in the unknowns c_k of the form,

$$\mathbf{A}\mathbf{c} = \mathbf{b}, \tag{13}$$

where $\mathbf{c} = (c_1, c_2, \cdots, c_N)$,

$$a_{ij} = \int_X g(x)\phi_i(x)\phi_j(x)\, dx, \qquad b_j = \int_X g(x)u(x)\phi_j(x)\, dx, \quad 1 \le i, j \le N \,. \tag{14}$$

Note that in the special case $g(x) = 1$, the matrix $\mathbf{A} = \mathbf{I}$, the $n \times n$ identity matrix, and the solution reduces to the Fourier coefficients in Eq. (11).

Note: In the examples that follow, we shall denote the "Weberized approximation" yielded by the solution of Eq. (13) as u_N^W in order to distinguish it from the best L^2 approximation, u_N, yielded by the Fourier coefficients Eq. (11).

Example 2: Consider the following step function on $X = [0, 1]$,

$$u(x) = \begin{cases} 1, & 0 \le x \le 1/2, \\ 3, & 1/2 < x \le 1. \end{cases} \tag{15}$$

The following set of $L^2[0, 1]$ basis functions was used: $\phi_1(x) = 1$, $\phi_k(x) = \sqrt{2}\cos(k\pi x)$, $k \ge 2$. In Figure 1 are presented plots of the best L^2 and best

weighted/Weberized L^2 approximations to $u(x)$ using $N = 5$ (left) and $N = 10$ (right) basis functions. As expected, the Weberized L^2 approximations, u_N^W, yield a higher L^2 errors than their best L^2 counterparts, u_N. Also as expected, the approximations u_N^W yield better approximations of $u(x)$ than u_N over $[0, 0.5]$ and a poorer approximations over $[0.5, 1]$. The Logarithmic L^2 approximations u_N^L shown in the figure will be discussed in Section 5.

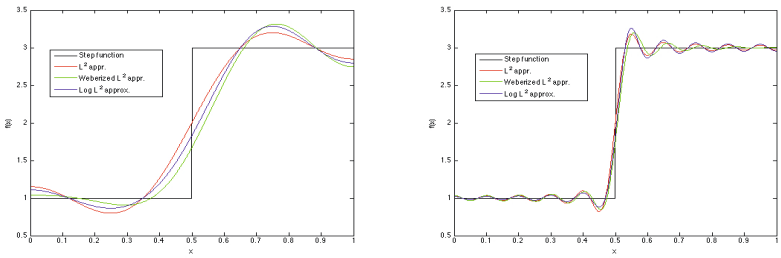

Fig. 1. Best L^2 (u_N, dotted), Weighted L^2 (u_N^W) and Logarithmic L^2 (u_N^L) approximations to step function in Eq. (15) using cosine basis functions. **Left:** $N = 5$. **Right:** $N = 20$. Approximation errors:

N	$\|u - u_N\|_2$	$\|u - u_N^W\|_2$	$\|u - u_N^L\|_2$
5	0.315	0.399	0.345
20	0.142	0.194	0.156

Example 3: The 512×512-pixel, 8 bits-per-pixel *Lena* image, partitioned into nonoverlapping 32×32-pixel blocks, with the first $N = 70$ standard 2D DCT basis functions used over each block (i.e., starting at $(0, 0)$, then $\{(1, 0), (0, 1)\}$, etc.. In Figure 2 are shown the best L^2 (left), Weberized L^2 (center) and Logarithmic L^2 (right) approximations to the shoulder region. The rather small value of N was chosen in order to demonstrate the significant differences as well as similarities between the L^2 and Weberized approximations in this region. The most significant differences occur in blocks containing edges that are formed between regions of low and high greyscale intensities, e.g., the edge defining Lena's shoulder. In each case, as expected, there is a ringing effect due to the low number of DCT basis functions employed ($N = 70$ out of a total of $32^2 = 1024$ functions). In the L^2 case (left), the error due to the ringing appears to be of similar magnitude in both light (shoulder) and dark (background) regions. In the Weberized L^2 cases, however, the ringing error appears to be larger over the lighter region (shoulder) than over the darker background (hair), which is consistent with the Weberized approximation method – a kind of two-dimensional analogy to the 1D step function in Example 2 above. As expected, blocks with little greyscale variation, e.g., the shoulder region without edges, are approximated equally well by the three methods.

Fig. 2. Best L^2 (left), Weberized L^2 (center) and Logarithmic L^2 (right) approximations to *Lena* image using $N = 70$ 2D DCT basis functions over 32×32-pixel blocks comprising the shoulder region of *Lena* image

4 Logarithmic L^2 Metric

Looking back at Eqs. (4) and (5) for the weighted L^2 metrics $\Delta(u, v)$ and $\Delta(v, u)$, we see that their accommodation of Weber's model of perception comes from the fact that their integrands involve ratios of the signals/images u and v. Indeed, a ratio between signals/images can also be obtained if we consider their logarithms. This, of course, is the basis of homomorphic filtering [4] and, indeed, this portion of our paper may be viewed from such a perspective. In this study, however, logarithms of image functions are used for the purpose of image approximation as opposed to image enhancement.

Our choice of logarithms may appear *ad hoc* but can actually be justified mathematically. Only a brief account can be presented here. As introduced in [1], we consider a measure ν defined over the greyscale space $\mathbb{R}_g$. Then define the following intensity-weighted distance between two functions u and v:

$$D(u, v; \nu) = \int_{X_u} \nu(u(x), v(x)]\, dx \; + \; \int_{X_v} \nu(v(x), u(x)]\, dx\,, \tag{16}$$

where $X_u = \{x \in X \mid u(x) < v(x)\} \subset X$ and $X_v = \{x \in X \mid u(x) \geq v(x)\} \subset X$. This distance involves an integration of the sizes of the greyscale intervals $(u(x), v(x)]$ or $(v(x), u(x)]$ over X. Note that in the special case, $\nu = m_g$, uniform Lebesgue measure on $\mathbb{R}_g$, the distance $D(u, v; \nu)$ in Eq. (16) becomes the L^1 distance between u and v [1].

Theorem 2: The unique measure ν on $\mathbb{R}_g$ which accommodates Weber's model of perception over the greyscale space $\mathbb{R}_g \subset \mathbb{R}_+$ is (up to a normalization constant) defined by the continuous density function $\rho(t) = \frac{1}{t}$.

Sketch of Proof: For any two greyscale intensities $I_1, I_2 \in \mathbb{R}_g$,

$$\int_{I_1}^{I_1+\Delta I_1} \frac{1}{t}\, dt = \int_{I_2}^{I_2+\Delta I_2} \frac{1}{t}\, dt \implies \nu(I_1, I_1 + \Delta I_1) = \nu(I_2, I_2 + \Delta I_2)\,, \tag{17}$$

where $\Delta I_1 = CI_1$ and $\Delta I_2 = CI_2$, are the minimum changes in perceived intensity at backgrounds I_1 and I_2, respectively, according to Weber's model in Eq. (1). Eq. (17) is a kind of invariance result with respect to perception.

Using this measure ν, the distance between u and v in Eq. (16) becomes

$$\begin{aligned} D(u,v;\nu) &= \int_{X_u} \left[\int_{u(x)}^{v(x)} \frac{1}{t}\, dt \right] dx \;+\; \int_{X_v} \left[\int_{v(x)}^{u(x)} \frac{1}{t}\, dt \right] dx \\ &= \int_X |\ln u(x) - \ln v(x)|\, dx\,, \end{aligned} \tag{18}$$

the logarithmic L^1 distance between u and v. All other logarithmic L^p distances, $p > 1$, may be viewed as generalizations of this result. This brief treatment hopefully shows why logarithms provide a natural representation for Weber's model.

We now outline the mathematical formalism for a logarithmic L^2-based approximation method. First define the space of functions $\mathcal{G}$ composed of the logarithms of all functions $u \in \mathcal{F}$, i.e.,

$$\mathcal{G} = \{U : X \to [\log A, \log B]\ ,\ U(x) = \log u(x),\ \forall x \in X\}\,. \tag{19}$$

Now consider the $L^2(X)$ distance between two elements $U, V \in \mathcal{G}$,

$$d_2(U,V) = \left[\int_X [U(x) - V(x)]^2\, dx \right]^2 \quad < \infty\,. \tag{20}$$

Use this distance to define the following "logarithmic L^2 distance" on $\mathcal{F}$,

$$d_{\log}(u,v) = d_2(U,V) = d_2(\log u, \log v)\,, \quad u, v \in \mathcal{F}\,. \tag{21}$$

Since $U = \log u$ implies that $u = e^U$ for all $U \in \mathcal{G}$, it can be shown that $d_{\log}$ is a metric on $\mathcal{F}$, i.e., it satisfies all of the properties of a metric, including the triangle inequality. From Eq. (21),

$$\begin{aligned} d_{\log}(u,v) &= \left[\int_X [\log u(x) - \log v(x)]^2\, dx \right]^{1/2} \\ &= \left[\int_X \left[\log \frac{u(x)}{v(x)} \right]^2 dx \right]^{1/2} = \left[\int_X \left[\log \frac{v(x)}{u(x)} \right]^2 dx \right]^{1/2}\,. \end{aligned} \tag{22}$$

The appearance of both ratios is a consequence of the symmetry of the metric.

Example 1 Revisited: The reference image $u(x) = I$ and constant approximation $v(x) = I + \Delta I$ as before. A quick calculation yields

$$d_{\log}(u,v) = \log\left(1 + \frac{\Delta I}{I}\right) = K \log(1 + C)\,, \tag{23}$$

where K is given in Eq. (8). As in the case of the weighted L^2 metric, $\Delta(u,v)$, the logarithmic L^2 distance is independent of the intensity level I.

As an interesting side note, in the case that the Weber constant C in Eq. (1) is small, then $\log(1+C) \approx C$, so that, from Eq. (9),

$$d_{\log}(u,v) \approx KC = \Delta(u,v)\,. \tag{24}$$

Experimentally, $C \approx 0.02$ [7] which justifies the above approximation.

5 Best Approximation in Terms of Logarithmic L^2 Metric

We shall now use the logarithmic L^2 distance to approximate a function $u \in \mathcal{F}$. As before, we consider, for an $N > 0$, an approximation u_N of the form in Eq. (10). The best approximation will minimize the squared $d_{\log}$ distance,

$$d^2_{\log}(u,u_N) = \int_X \left[\log u(x) - \log\left(\sum_{k=1}^{N} c_k\phi_k(x)\right)\right]^2 dx =: h(c_1,\cdots,c_N). \tag{25}$$

Unfortunately, application of the stationarity conditions $\frac{\partial h}{\partial c_k} = 0$, $1 \le k \le N$, yields an extremely complicated set of nonlinear equations in the unknown coefficients c_k. A huge simplification is accomplished if we consider the L^2 approximation of the logarithmic function $U(x) = \log u(x)$. The goal is then to approximate $U \in \mathcal{G} \subset L^2(X)$ as follows,

$$U \approx U_N = \sum_{k=1}^{N} a_k\phi_k\,. \tag{26}$$

The minimization of the squared L^2 distance, $d_2^2(U,U_N)$, is provided by the Fourier coefficients a_k of U in the ϕ_k basis, i.e.,

$$a_k = \langle U,\phi_k\rangle = \int_X U(x)\phi_k(x)\,dx\,. \tag{27}$$

We bypass some technical mathematical details and simply state that the logarithmic L^2-based approximations to u, which we shall denote by u_N^L, are given by

$$u_N^L(x) = \exp\left(U_N(x)\right) = \exp\left(\sum_{k=1}^{N} a_k\phi_k(x)\right)\,. \tag{28}$$

In summary, the logarithmic L^2 approximation method is seen to be much simpler than the weighted/Weberized L^2 method. One finds the Fourier coefficients of the logarithm U of the signal and then exponentiates to recover the approximation u_N. There is no system of equations to be solved.

Example 2 Revisited: We again consider the step function $u(x)$ in Eq. (15) and employ the same orthonormal cosine basis on $[0,1]$. The best logarithmic

L^2 approximations, u_N^L, to $u(x)$ using $N = 5$ basis functions (left) and $N = 10$ basis functions (right) are plotted in Figure 1 along with their L^2 and Weberized L^2 counterparts. As expected, the logarithmic L^2 approximations are seen to behave in a "Weberized" way. Note that the L^2 approximation errors associated with the logarithmic approximations are significantly lower than those of the weighted L^2 method.

Example 3 Revisited: The *Lena* image, approximated over 32×32 pixel blocks with $N = 70$ 2D DCT basis functions. The approximations afforded by the Logarithmic L^2 method are virtually identical to their Weberized L^2 counterparts. As such, they display the same kind of "Weberized ringing" over regions with edges separating high and low greyscale intensities, with lesser ringing error over the latter regions.

Acknowledgments. We gratefully acknowledge that this research has been supported in part by the Natural Sciences and Engineering Research Council of Canada (ERV,ZW) and the Ontario Centre of Excellence (ZW).

References

1. Forte, B., Vrscay, E.R.: Solving the inverse problem for function and image approximation using iterated function systems. Dynamics of Continuous, Discrete and Impulsive Systems **1**, 177–231 (1995)
2. Girod, B.: What's wrong with mean squared error? In: Watson, A.B. (ed.) Digital Images and Human Vision. MIT Press, Cambridge (1993)
3. Lee, S., Pattichis, M.S., Bovik, A.C.: Foveated video quality assessment. IEEE Trans. Multimedia **4**(1), 129–132 (2002)
4. Oppenheim, A.V., Schafer, R.W., Stockham Jr, T.G.: Nonlinear filtering of multiplied and convolved signals. Proc. IEEE **56**(8), 1264–1291 (1968)
5. Shen, J.: On the foundations of vision modeling I. Weber's law and Weberized TV restoration. Physica D **175**, 241–251 (2003)
6. Shen, J., Jung, Y.-M.: Weberized Mumford-Shah model with Bose-Einstein photon noise. Appl. Math. Optim. **53**, 331–358 (2006)
7. Wandell, B.A.: Foundations of Vision. Sinauer Publishers, Sunderland (1995)
8. Wang, Z., Bovik, A.C.: Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Sig. Proc. Mag. **26**, 98–117 (2009)
9. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Proc. **13**(4), 600–612 (2004)
10. Wang, Z., Li, Q.: Information content weighting for perceptual image quality assessment. IEEE Trans. Image Proc. **20**(5), 1185–1198 (2011)