Structural Similarity-Based Approximation over Orthogonal Bases: Investigating the Use of Individual Component Functions $S_k(\mathbf{x}, \mathbf{y})$

Paul Bendevis and Edward R. Vrscay(✉)

Department of Applied Mathematics, Faculty of Mathematics,
University of Waterloo, Waterloo, ON N2L 3G1, Canada
{ptbendev,ervrscay}@uwaterloo.ca

Abstract. We examine the use of individual components of the Structural Similarity image quality measure as criteria for best approximation in terms of orthogonal expansions. We also introduce a family of higher order SSIM-like rational functions.

1 Introduction

In this paper, we wish to examine further the idea of orthogonal expansions in $\mathbb{R}^N$ that are best approximations in terms of the Structural Similarity (SSIM) image quality measure [6]. This study represents a kind of followup of the work presented in an earlier ICIAR conference [3] in which the optimization was done with respect to the SSIM function, $S(\mathbf{x}, \mathbf{y})$, a product of three terms, namely, (i) luminance, (ii) constrast and (iii) structure. The luminance term, $S_1(\mathbf{x}, \mathbf{y})$ is a function only of the means of $\mathbf{x}$ and $\mathbf{y}$ and therefore cannot provide a nontrivial approximation. Here we examine whether either of the contrast or structure terms alone can be used as a criterion for best approximation determination and find that the answer is negative. At least two of the three components, one of which must be the luminance term, are necessary to provide a unique solution.

In the final section of this paper, we introduce a family of higher-order or "generalized" SSIM functions, using a method that is analogous to the construction of the rational function $S_1(\mathbf{x}, \mathbf{y})$.

2 SSIM and SSIM-Based Approximations of Signals/Images

In what follows, we let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ denote two N-dimensional signal/image blocks or local patches, i.e., $\mathbf{x} = (x_1, x_2, \cdots, x_N)$. The SSIM measure between $\mathbf{x}$ and $\mathbf{y}$ was defined originally as follows [6,7],

$$\begin{aligned} S(\mathbf{x}, \mathbf{y}) &= S_1(\mathbf{x}, \mathbf{y}) S_2(\mathbf{x}, \mathbf{y}) S_3(\mathbf{x}, \mathbf{y}) \\ &= \left[\frac{2\bar{\mathbf{x}}\bar{\mathbf{y}} + \epsilon_1}{\bar{\mathbf{x}}^2 + \bar{\mathbf{y}}^2 + \epsilon_1}\right] \left[\frac{2 s_{\mathbf{x}} s_{\mathbf{y}} + \epsilon_2}{s_{\mathbf{x}}^2 + s_{\mathbf{y}}^2 + \epsilon_2}\right] \left[\frac{s_{\mathbf{xy}} + \epsilon_3}{s_{\mathbf{x}} s_{\mathbf{y}} + \epsilon_3}\right], \end{aligned} \tag{1}$$

A. Campilho and M. Kamel (Eds.): ICIAR 2014, Part I, LNCS 8814, pp. 55–64, 2014.
DOI: 10.1007/978-3-319-11758-4_7

where

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad s_{\mathbf{xy}} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{\mathbf{x}})(y_i - \bar{\mathbf{y}}), \quad s_{\mathbf{x}} = \sqrt{s_{\mathbf{xx}}}, \quad \text{etc.}. \tag{2}$$

The small positive constants ϵ_k are added for numerical stability and can be adjusted to accomodate the perception of the human visual system.

The functional form of the component S_1 in Eq. (1), which measures the similarities of local patch luminances or brightness values, was originally chosen in an effort to accomodate Weber's law of perception [7]. The form of S_2, which measures the similarities of local patch contrasts, follows the idea of divisive normalization [5]. In the case that $\epsilon_3 = 0$, the component S_3, which measures the similarities of local patch structures, is precisely the correlation $C(\mathbf{x}, \mathbf{y})$ between $\mathbf{x}$ and $\mathbf{y}$. We shall be taking a closer look at S_2 and S_3 below.

Note that $-1 \leq S(\mathbf{x}, \mathbf{y}) \leq 1$, and $S(\mathbf{x}, \mathbf{y}) = 1$ if and only if $\mathbf{x} = \mathbf{y}$. The component $S_1(\mathbf{x}, \mathbf{y})$ measures the similarity between the means of $\mathbf{x}$ and $\mathbf{y}$: If $\bar{\mathbf{x}} = \bar{\mathbf{y}}$, then $S_1(\mathbf{x}, \mathbf{y}) = 1$, its maximum possible value. We shall return to this idea in a later Section.

It is common practice to set $\epsilon_2 = 2\epsilon_3$, in which case the product of S_2 and S_3 in Eq. (1) collapses to a single term, namely,

$$S_{2'}(\mathbf{x}, \mathbf{y}) = \frac{2s_{\mathbf{xy}} + \epsilon_2}{s_{\mathbf{x}}^2 + s_{\mathbf{y}}^2 + \epsilon_2}\,. \tag{3}$$

It was indeed this form of the SSIM, i.e., $S(\mathbf{x}, \mathbf{y}) = S_1(\mathbf{x}, \mathbf{y})S_{2'}(\mathbf{x}, \mathbf{y})$, that was analysed in [3]. (In that paper, $S_{2'}$ was denoted as S_2.) In this paper, we wish to examine the roles of the individual terms S_2 and S_3, as opposed to their product.

In the discussion that follows, we consider $\mathbf{x}$ to be a given signal and $\mathbf{y} \in A$ to be an approximation to $\mathbf{x}$ where A is an M-dimensional subset of $\mathbb{R}^N$, with $M \leq N$. Of course, we shall be concerned with *best approximations* to $\mathbf{x}$, but not in terms of SSIM, as was done in [3]. Instead, we shall be determining best approximations with respect to each of the SSIM components S_2 and S_3.

As in [3], we work with a set of (complete) orthonormal basis functions $\mathbb{R}^N$, to be denoted as $\{\phi_0, \phi_1, \cdots, \phi_{N-1}\}$. We assume that only the first element has nonzero mean: $\overline{\phi_k} = 0$ for $1 \leq k \leq N-1$. We also assume that ϕ_0 is "flat", i.e., constant: $\phi_0 = N^{-1/2}(1, 1, \cdots, 1)$, which accomodates the discrete cosine transform (DCT) as well as Haar multiresolution system on $\mathbb{R}^N$. (If ϕ_0 were not constant, then the definitions of the mean $\bar{\mathbf{x}}$ in Eq. (2) can be modified accordingly.) The L^2-based expansion of $\mathbf{x}$ in this basis is

$$\mathbf{x} = \sum_{k=0}^{N-1} a_k\phi_k, \qquad a_k = \langle \mathbf{x}, \phi_k \rangle, \quad 0 \leq k \leq N-1, \tag{4}$$

from which it follows that

$$\bar{\mathbf{x}} = a_0 N^{-1/2}. \tag{5}$$

The expansions of the approximation $\mathbf{y} \in A$ to $\mathbf{x}$ will be denoted as follows,

$$\mathbf{y} = \mathbf{y}(\mathbf{c}) = \sum_{k=0}^{N-1} c_k \phi_k, \tag{6}$$

where the notation $\mathbf{y}(\mathbf{c})$ acknowledges the dependence of the approximation on the coefficients c_k. It also follows that

$$\bar{\mathbf{y}} = c_0 N^{-1/2}. \tag{7}$$

Note that at this point, we do not assume a relationship between a_0 and c_0.

For simplicity of discussion, we consider the approximation spaces A to be

$$A_M = \text{span}\{\phi_0, \phi_1, \cdots, \phi_{M-1}\} \tag{8}$$

where $0 \leq M \leq N-1$. The discussion which follows can easily be adapted to conform to the situation studied in [3] in which an arbitrary subset of distinct functions $\{\phi_{\gamma(k)}\}$, $0 \leq k \leq M-1$ was chosen from the complete set of N functions. As is well known, the best L^2-based approximation of x in A_M is

$$\mathbf{y}_{M,L^2} = \mathbf{x}_M := \sum_{k=0}^{M-1} a_k \phi_k, \quad a_k = \langle \mathbf{x}, \phi_k \rangle, \tag{9}$$

a truncation of the exact expansion of $\mathbf{x}$ in Eq. (4). In terms of Eq. (6),

$$c_k = a_k, \quad 0 \leq k \leq M-1 \text{ and } c_k = 0, \quad M \leq k \leq N-1\,. \tag{10}$$

As is also well known,

$$\mathbf{y}_{M,L^2} = \arg \min_{z \in A_m} \|x - z\|_2, \tag{11}$$

the unique element in A_M that that lies closest to $\mathbf{x} \in \mathbb{R}^N$.

In [3], the best SSIM-based approximation $\mathbf{y}_{SSIM} \in A_M$ to $\mathbf{x}$, using the SSIM function $S(\mathbf{x},\mathbf{y}) = S_1(\mathbf{x},\mathbf{y})S_{2'}(\mathbf{x},\mathbf{y})$, was found to be

$$\mathbf{y}_{M,SSIM} = \arg \max_{\mathbf{z} \in A_M} S(\mathbf{x},\mathbf{z}) = \sum_{k=0}^{M-1} c_k \phi_k\,, \tag{12}$$

where $c_0 = a_0$ and

$$c_k = \alpha a_k\,, \quad 1 \leq k \leq M-1\,. \tag{13}$$

The scaling coefficient α is given by

$$\alpha = \frac{-\epsilon_2 + \sqrt{\epsilon_2^2 + \left(\frac{4}{N-1}\sum_{k=1}^{M-1} a_k^2\right)(s_{\mathbf{x}}^2 + \epsilon_2)}}{\frac{2}{N-1}\sum_{k=1}^{M-1} a_k^2}\,. \tag{14}$$

In the special case $\epsilon_2 = 0$,

$$\alpha = \left[\sum_{k=1}^{N-1} a_k^2\right]^{1/2} \left[\sum_{k=1}^{M-1} a_k^2\right]^{-1/2} . \tag{15}$$

When $M < N$, the scaling coefficient $\alpha > 1$, which implies that the $\mathbf{y}_{M,SSIM}$ is a *contrast enhanced* version of $\mathbf{y}_{M,L^2}$.

3 Best S_3/Correlation-Based Approximations

Here we consider the following problem: Given an $\mathbf{x} \in \mathbb{R}^N$, find, if possible, the best S_3-based approximation to $\mathbf{x}$ in A_M, i.e.,

$$\mathbf{y}_{M,S_3} = \arg \max_{\mathbf{z} \in A_M} S_3(\mathbf{x}, \mathbf{z}) . \tag{16}$$

As mentioned in Section 1, in the case $\epsilon_3 = 0$, $S_3(\mathbf{x}, \mathbf{y}) = C(\mathbf{x}, \mathbf{y})$, the correlation between $\mathbf{x}$ and $\mathbf{y}$. It is necessary to express the function $S_3(\mathbf{x}, \mathbf{y})$ in terms of the unknown coefficients c_k of the expansion for $\mathbf{y}(\mathbf{c})$. The following results, which are obtained after some simple algebra, are useful:

$$s_{\mathbf{x}}^2 = \frac{1}{N-1} \sum_{k=1}^{N-1} a_k^2, \quad s_{\mathbf{xy}} = \frac{1}{N-1} \sum_{k=1}^{M-1} a_k c_k, \quad s_{\mathbf{y}}^2 = \frac{1}{N-1} \sum_{k=1}^{M-1} c_k^2 . \tag{17}$$

We then have that

$$S_3(\mathbf{x}, \mathbf{y}(\mathbf{c})) = \frac{\frac{1}{N-1} \sum_{k=1}^{M-1} a_k c_k + \epsilon_3}{\frac{1}{N-1} \left[\sum_{k=1}^{N-1} a_k^2\right]^{1/2} \left[\sum_{k=1}^{M-1} c_k^2\right]^{1/2} + \epsilon_3} . \tag{18}$$

Note that the right side is a function of the coefficients $c_1, c_1, \cdots c_{M-1}$, but *not* of c_0. This is already an indication that the maximizer of S_3 may not be unique.

We now look for stationary points that will be candidates for maximum points of $S_3(\mathbf{x}, \mathbf{y}(\mathbf{c}))$. Imposition of the stationarity conditions $\partial S_3 / \partial c_p = 0$ for $1 \le p \le M-1$ leads to the following set of equations,

$$\frac{1}{s_{\mathbf{xy}} + \epsilon_3} a_p - \frac{1}{s_{\mathbf{x}} s_{\mathbf{y}} + \epsilon_3} \frac{s_{\mathbf{x}}}{s_{\mathbf{y}}} c_p = 0 , \quad 1 \le p \le M-1 . \tag{19}$$

If $a_p = 0$ for any $1 \le p \le M-1$, then $c_p = 0$. Otherwise, we may rewrite the above relations as

$$\frac{c_p}{a_p} = \frac{s_{\mathbf{x}} s_{\mathbf{y}} + \epsilon_3}{s_{\mathbf{xy}} + \epsilon_3} \frac{s_{\mathbf{y}}}{s_x} , \quad 1 \le p \le M-1 . \tag{20}$$

The RHS of each equation is independent of p, implying that

$$c_p = \alpha a_p , \quad 1 \le p \le M-1 , \tag{21}$$

Such a proportionality between the c_p and a_p was also found in [3] for the case of best SSIM-based approximations. We now attempt to determine α by rewriting Eq. (20) and using Eq. (21) to obtain the relation

$$\alpha s_{\mathbf{x}}(s_{\mathbf{xy}} + \epsilon_3) = s_{\mathbf{y}}(s_{\mathbf{x}} s_{\mathbf{y}} + \epsilon_3)\,. \tag{22}$$

Substition of the expansions in (17) followed by simplification yields

$$\alpha^2 \epsilon_3^2 \sum_{k=1}^{N-1} a_k^2 = \alpha^2 \epsilon_3^2 \sum_{k=1}^{M-1} a_k^2\,. \tag{23}$$

In the special case that $M = N$, Eq. (23) is an identity that is satisfied by any $\alpha \in \mathbb{R}$, independent of the value of ϵ_3. This is due to the following result, the proof of which is very straightforward and therefore omitted.

Theorem 1: Let $\mathbf{x} \in \mathbb{R}$ and $\mathbf{y} = a\mathbf{x} + b\mathbf{1}$, where $a, b \in \mathbb{R}$, $a \neq 0$ and $\mathbf{1} = (1, 1, \cdots, 1) \in \mathbb{R}$. Then

$$C(\mathbf{x}, \mathbf{y}) = \frac{s_{\mathbf{xy}}}{s_{\mathbf{x}} s_{\mathbf{y}}} = \operatorname{sgn}(a) = \begin{cases} 1, & a > 0, \\ -1, & a < 0. \end{cases} \tag{24}$$

When $2 \leq M < N$, we must consider two cases in Eq. (23):

1. $\epsilon_3 = 0$**:** Eq. (23) is satisfied for all $\alpha \in \mathbb{R}$. This is again a consequence of Theorem 1 and can be confirmed by substituting Eq. (21) into Eq. (18).
2. $\epsilon_3 \neq 0$**:** In order that Eq. (23) be satisfied for any sequence of Fourier coefficients $\{a_k\}_{k=1}^{N-1}$, it is necessary that $\alpha = 0$, which implies that $\mathbf{y}_{M,S_3} = c_0\phi_0 \in A_1$, i.e., the constant approximation, with c_0 as yet undetermined.

The conclusion is that S_3-based best approximation is possible only in the case $\epsilon_3 = 0$. But even in this case, the result is not unique since α is arbitrary (but positive). What makes matters worse is that the leading coefficient c_0 of the approximation $\mathbf{y}(\mathbf{c})$ also remains undetermined! It appears that two additional conditions are required in order to obtain unique values of c_0 and α.

Eqs. (5) and (7) imply that a unique value of c_0 can be obtained by imposing a relation between the means $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$. It would seem natural to impose the following "equal means" condition,

$$\bar{\mathbf{x}} = \bar{\mathbf{y}}\,, \tag{25}$$

which, as is well known, maximizes the component function $S_1(\mathbf{x}, \mathbf{y})$ in Eq. (1). In this case, $c_0 = a_0$, as was found in [3].

Using this result, we must still determine a unique value of the proportionality coefficient α in Eq. (21). If we impose a condition of "equal norms," i.e.,

$$\|\mathbf{x}\|_2 = \|\mathbf{y}_{M,S_3}\|_2\,, \tag{26}$$

then squaring both sides and substituting the respective expansions along with the property in Eq. (21) and the condition $c_0 = a_0$ yields

$$a_0^2 + \sum_{k=1}^{N-1} a_k^2 = c_0^2 + \sum_{k=1}^{M-1} c_k^2 = a_0^2 + \alpha^2 \sum_{k=1}^{M-1} a_k^2 \,. \tag{27}$$

This result leads to the (positive) value of α in Eq. (15).

Here it is important to mention that if the "equal means" condition of (25) is replaced by another relationship between c_0 and a_0, then the "equal norms" condition of (26) yields a scaling coefficient α different from that in Eq. (15).

The results in this section lead to the following important conclusions:

1. Using only the component function $S_3(\mathbf{x}, \mathbf{y})$ is insufficient to determine a best SSIM-based approximation $y_{M,S_3} \in A_M$ to $\mathbf{x}$.
2. Using the components $S_1(\mathbf{x}, \mathbf{y})$ and $S_3(\mathbf{x}, \mathbf{y})$ is also insufficient to find a best approximation. In the special case, $\epsilon_3 = 0$, however, a unique solution may be obtained by imposing two additional conditions, e.g., equality of means, Eq. (25), and equality of norms, Eq. (26).

4 Best S_2/Contrast-Based Approximations

We now consider the following problem: Given an $\mathbf{x} \in \mathbb{R}^N$, find, if possible, the best S_2-based approximation to $\mathbf{x}$ in A_M, i.e.,

$$\mathbf{y}_{M,S_2} = \arg \max_{\mathbf{z} \in A_M} S_2(\mathbf{x}, \mathbf{z}) \,, \tag{28}$$

where $S_2(\mathbf{x}, \mathbf{y})$ is defined in Eq. (1). From the equations in (17) we may express $S_2(\mathbf{x}, \mathbf{y})$ in terms of the expansion coefficients c_k. It is not absolutely necessary to present this expansion here, but only to note that, as in the case of the $S_3(\mathbf{x}, \mathbf{y})$ function examined in the previous section, $S_2(\mathbf{x}, \mathbf{y})$ is a function of the $M-1$ coefficients $c_1, c_1, \cdots c_{M-1}$ and independent of the coefficient c_0.

Imposition of the stationarity conditions $\partial S_2 / \partial c_p = 0$, for $1 \leq p \leq M-1$ leads to the following set of equations,

$$\left[\frac{s_{\mathbf{x}}}{s_{\mathbf{y}}} (s_{\mathbf{x}}^2 + s_{\mathbf{y}}^2 + \epsilon_2) - (2 s_{\mathbf{x}} s_{\mathbf{y}} + \epsilon_2) \right] c_p = 0, \quad 1 \leq p \leq M-1 \,. \tag{29}$$

The most noteworthy feature of these equations is the absence of of a direct relation between c_p and a_p as was seen in Eq. (19) for the S_3 case. As such, the existence of a proportionality result of the form in Eq. (21) cannot be proved. Since all of the coefficients c_p cannot, in general, be zero, it follows that the term in the square brackets must vanish. We now examine two cases:

1. $\epsilon_2 = 0$**:** The vanishing of the term in the square brackets of Eq. (29) reduces to the result

$$s_{\mathbf{x}}^2 = s_{\mathbf{y}}^2 \quad \Longrightarrow \quad \sum_{k=1}^{N-1} a_k^2 = \sum_{k=1}^{M-1} c_k^2 \,. \tag{30}$$

If we now *assume* a proportionality between the c_p and a_p as in Eq. (21), we arrive at the result for α in Eq. (15). Note that we must still impose the equal-means condition of Eq. (25) to obtain a unique approximation $\mathbf{y}_{M,S_2} \in A_M$ in this case.

2. $\epsilon_2 \neq 0$**:** Assuming that $s_{\mathbf{y}} \neq 0$, the vanishing of the term in square brackets implies that

$$s_{\mathbf{x}}^3 - s_{\mathbf{x}} s_{\mathbf{y}}^2 + \epsilon_2 s_{\mathbf{x}} - \epsilon_2 s_{\mathbf{y}} = 0\,. \tag{31}$$

If we once again *assume* a proportionality between the c_p and the a_p as in Eq. (21), a quadratic equation in α is obtained. It is convenient to rewrite it as an equation in a scaled variable β as follows,

$$s_{\mathbf{x}}\beta^2 + \epsilon_2\beta - s_{\mathbf{x}}^3 - \epsilon_2 s_{\mathbf{x}} = 0\,, \text{ where } \beta = \frac{\alpha}{\sqrt{N-1}}\left[\sum_{k=1}^{M-1} a_k^2\right]^{1/2} = s_{\mathbf{x}_M}\alpha\,. \tag{32}$$

The definition of $s_{\mathbf{x}_M}$ follows from Eqs. (9) and (17). The value of α which results from the positive solution of this quadratic equation is

$$\alpha = \frac{-\epsilon_2 + \sqrt{\epsilon_2^2 + 4s_{\mathbf{x}}^2(s_{\mathbf{x}}^2 + \epsilon_2)}}{2 s_{\mathbf{x}} s_{\mathbf{x}_M}}\,. \tag{33}$$

In the case $\epsilon_2 \to 0$, the result in Eq. (15) is obtained. For $\epsilon_2 \neq 0$, however, a comparision between this equation and Eq. (14) shows that the results are *almost* the same. They do differ, however, except in the case $M = N$. The reason for this difference is the appearance of the term $s_{\mathbf{xy}}$ in the numerator of the SSIM function $S_{2'}(\mathbf{x}, \mathbf{y})$, cf. Eq. (3), as opposed to $s_{\mathbf{x}} s_{\mathbf{y}}$ in the SSIM function $S_2(\mathbf{x}, \mathbf{y})$, cf. Eq. (1).

In summary, the S_2-based best approximation method of this section is possible with a scaling condition of the form Eq. (21). As with the S_3-based case, however, the coefficient c_0 is undetermined. If the equal-means condition of Eq. (25) is employed then we obtain, as before, $c_0 = a_0$.

In this section, we have arrived at the following important conclusions:

1. Using only the component function $S_2(\mathbf{x}, \mathbf{y})$ is insufficient to determine a best SSIM-based approximation $y_{M,S_2} \in A_M$ to $\mathbf{x}$.
2. Using components $S_1(\mathbf{x}, \mathbf{y})$ and $S_2(\mathbf{x}, \mathbf{y})$ is sufficient to find a unique, best SSIM-based approximation $\mathbf{y}_{M,S_2}$ for both zero and nonzero values of the stability constant ϵ_2. For all $M < N$, however, this approximation is different from the best SSIM-based counterpart of Eq. (12), obtained by using the entire SSIM function $S(\mathbf{x}, \mathbf{y})$ in Eq. (1).

5 A Family of Higher-Order Rational SSIM Functions

In this section, we show how a rationalization procedure that can be used to construct the SSIM functions $S_1(\mathbf{x}, \mathbf{y})$ in Eq. (1) and $S_{2'}(\mathbf{x}, \mathbf{y})$ in Eq. (3) may be used to construct higher order SSIM-like rational functions. We then consider the use of these functions for best SSIM-based approximations and present some preliminary results. In the discussion that follows, as before, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$.

With reference to the function $S_1(\mathbf{x}, \mathbf{y})$, note that $\mathbf{x} = \mathbf{y}$ implies $\bar{\mathbf{x}} = \bar{\mathbf{y}}$ and $S_1(\mathbf{x}, \mathbf{y}) = 1$. But $\bar{\mathbf{x}} = \bar{\mathbf{y}}$ also implies $(\bar{\mathbf{x}} - \bar{\mathbf{y}})^2 = 0$, so that

$$\bar{\mathbf{x}}^2 + \bar{\mathbf{y}}^2 = 2\bar{\mathbf{x}}\bar{\mathbf{y}}. \tag{34}$$

Now add a "stability constant" $\epsilon_1 > 0$ to each side and divide to obtain

$$S_1(\mathbf{x}, \mathbf{y}) = \frac{2\bar{\mathbf{x}}\bar{\mathbf{y}} + \epsilon_1}{\bar{\mathbf{x}}^2 + \bar{\mathbf{y}}^2 + \epsilon_1} = 1\,. \tag{35}$$

Now define

$$\mathbf{x}_0 = \mathbf{x} - \bar{\mathbf{x}}\mathbf{1}, \qquad \mathbf{y}_0 = \mathbf{y} - \bar{\mathbf{y}}\mathbf{1}, \tag{36}$$

where $\mathbf{1} = (1, 1, \cdots, 1) \in \mathbb{R}^N$. By construction, $\mathbf{x}_0$ and $\mathbf{y}_0$ have zero-mean. Clearly, $\mathbf{x} = \mathbf{y}$ implies that $\mathbf{x}_0 = \mathbf{y}_0$ which, in turn, implies that

$$\|\mathbf{x}_0 - \mathbf{y}_0\|_n^n = 0, \quad n = 2, 3, 4, \cdots\,. \tag{37}$$

We consider only the case that n is even, so that Eq. (37) becomes

$$\sum_{k=1}^{N} \left[(x_k - \bar{\mathbf{x}}) - (y_k - \bar{\mathbf{y}})\right]^n = 0, \quad n = 2, 4, 6, \cdots\,. \tag{38}$$

Now use the binomial theorem and then rearrange to produce the result,

$$\sum_{l=0}^{n} (-1)^l B_{nl} \sum_{k=1}^{N} (x_k - \bar{\mathbf{x}})^{n-l} (y_k - \bar{\mathbf{y}})^l = 0, \quad \text{where} \quad B_{nl} = \binom{n}{l}\,. \tag{39}$$

Dividing by $N - 1$, this relation becomes

$$\sum_{l=0}^{n} (-1)^l B_{nl} s_{n-l,l} = 0, \tag{40}$$

where we have defined

$$s_{p,q} = \frac{1}{N-1} \sum_{k=1}^{N} (x_k - \bar{\mathbf{x}})^p (y_k - \bar{\mathbf{y}})^q, \quad p, q \geq 0\,. \tag{41}$$

Now rewrite Eq. (40) as follows,

$$\sum_{l \text{ even}} B_{nl} s_{n-l,l} = \sum_{l \text{ odd}} B_{nl} s_{n-l,l}\,. \tag{42}$$

Once again we add a "stability constant" $\epsilon_n \geq 0$ to both sides and divide by the left-hand term to obtain the result,

$$\Sigma_n(\mathbf{x}, \mathbf{y}) := \frac{\sum_{l \text{ odd}} B_{nl} s_{n-l,l} + \epsilon_n}{\sum_{l \text{ even}} B_{nl} s_{n-l,l} + \epsilon_n} = 1. \tag{43}$$

The rational functions $\Sigma_n(\mathbf{x}, \mathbf{y})$, $n \in \{2, 4, 6, \cdots\}$ define a set of *generalized SSIM functions* between vectors $\mathbf{x}$ and $\mathbf{y}$. When $n = 2$, $\Sigma_2(\mathbf{x}, \mathbf{y}) = S_{2'}(\mathbf{x}, \mathbf{y})$, the SSIM function of Eq. (3). For $n = 4$, we have the function

$$\Sigma_4(\mathbf{x}, \mathbf{y}) = \frac{4s_{\mathbf{xxxy}} + 4s_{\mathbf{xyyy}} + \epsilon_4}{s_{\mathbf{xxxx}} + 6s_{\mathbf{xxyy}} + s_{\mathbf{yyyy}} + \epsilon_4}\,. \tag{44}$$

At this point it is helpful to recall that the construction of the function $\Sigma_n(\mathbf{x}, \mathbf{y})$ is based on the L^n norm in Eq. (37). The closeness of $\Sigma_n(\mathbf{x}, \mathbf{y})$ to 1 is related to the closeness of $\|\mathbf{x}_0 - \mathbf{y}_0\|_n^n$ to 0.

We now define the following family of associated distance functions,

$$\begin{aligned} T_n(\mathbf{x}, \mathbf{y}) &:= 1 - \Sigma_n(\mathbf{x}, \mathbf{y}) \\ &= \frac{\sum_{l \text{ even}} B_{nl} s_{n-l,l} - \sum_{l \text{ odd}} B_{nl} s_{n-l,l}}{\sum_{l \text{ even}} B_{nl} s_{n-l,l} + \epsilon_n}\,. \end{aligned} \tag{45}$$

From (37) and (40),

$$T_n(\mathbf{x}, \mathbf{y}) = \frac{\|\mathbf{x}_0 - \mathbf{y}_0\|_n^n}{\sum_{l \text{ even}} B_{nl} s_{n-l,l} + \epsilon_n}\,. \tag{46}$$

If $\mathbf{x} = \mathbf{y}$, then $T_n(\mathbf{x}, \mathbf{y}) = 0$. The $T_n(\mathbf{x}, \mathbf{y})$ functions are weighted L^n distances – to the nth power – between $\mathbf{x}_0$ and $\mathbf{y}_0$. The case $n = 2$ in Eq. (46) has been examined in the past [3,4]:

$$T_2(\mathbf{x}, \mathbf{y}) = 1 - S_{2'}(\mathbf{x}, \mathbf{y}) = \frac{\|\mathbf{x}_0 - \mathbf{y}_0\|_2^2}{s_{\mathbf{xx}} + s_{\mathbf{yy}} + C_2} = \frac{\|\mathbf{x}_0 - \mathbf{y}_0\|_2^2}{\|\mathbf{x}_0\|_2^2 + \|\mathbf{y}_0\|_2^2 + C_2}\,. \tag{47}$$

It is an example of a *normalized metric* [1,2].

One may now wish to consider the problem of finding best approximations, using these rational SSIM-like functions as the objective functions: Given an $\mathbf{x} \in \mathbb{R}^N$, an $M < N$ and and $n \geq 1$, find, if possible,

$$\mathbf{y}_{M,\Sigma_{2n}} = \arg \max_{\mathbf{z} \in A_M} \Sigma_{2n}(\mathbf{x}, \mathbf{z}) = \arg \min_{\mathbf{z} \in A_M} T_{2n}(\mathbf{x}, \mathbf{z})\,. \tag{48}$$

The case $n = 1$, which corresponds to the function $S_{2'}(\mathbf{x}, \mathbf{y})$, along with the equal-means condition coming from the function $S_1(\mathbf{x}, \mathbf{y})$, was analyzed in [2].

The cases $n \geq 2$ present a major challenge, however, since we may no longer exploit the orthogonality properties of the ϕ_k basis in the computation of the functions $s_{p,q}$ in Eq. (41). At this time, only a little progress has been made on this problem. For example, in the case $n = 2$, the generalized SSIM function Σ_4 is a complicated rational function of the unknown expansion coefficients c_k:

$$\Sigma_4(\mathbf{c}) = \frac{\sum_{n,m,p,q}[4a_n a_m a_p c_q + 4a_n c_m c_p c_q] P_{nmpq}}{\sum_{n,m,p,q}[a_n a_m a_p a_q + 6a_n a_m c_p c_q + c_n c_m c_p c_q] P_{nmpq}}\,, \tag{49}$$

where the summation limits are $1 \le m, n \le N-1$ and $1 \le p, q \le M-1$ and

$$P_{nmpq} = \sum_{k=1}^{N} \phi_{nk}\phi_{mk}\phi_{pk}\phi_{qk} \,. \tag{50}$$

These coefficients are clearly basis-dependent. Moreover, $P_{nmpq} = P_{\mathcal{P}[nmpq]}$, where $\mathcal{P}$ denotes any permutation. Numerical experiments indicate that many of these coefficients are zero.

The stationarity conditions $\partial\Sigma_4/\partial c_p = 0$, $1 \le p \le M-1$, yield an enormously complicated set of coupled nonlinear equations in the c_p. In the special case that $M = N$, the solution of this system is $c_p = a_p$, as expected. When $M < N$, there exist solutions of the form $c_p = \alpha a_p$, as was found in in Eq. (21). The scaling coefficient α, however, satisfies a cubic equation in α^2. The coefficients of this equation are complicated functions of the Fourier coefficients a_k and the P_{nmpq}.

The fact that solutions of the form $c_p = \alpha a_p$ exist is interesting. They will, however, probably turn out to be of little use since they have the same basic form as those obtained from the S_2 and S_3 SSIM functions (along with S_1), which represents nothing more than another adjustment in the contrast.

In summary, we have introduced, and hopefully motivated, the idea of constructing higher-order rational SSIM-like functions. The method outlined above is the most straightforward one. It may well be useful to consider other methods which are based on higher-order statistics of images.

Acknowledgments. We gratefully acknowledge that this research has been supported in part by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (ERV).

References

1. Brunet, D.: A Study of the Structural Similarity Image Quality Measure with Applications to Image Processing, Ph.D. Thesis, University of Waterloo (2010)
2. Brunet, D., Vrscay, E.R., Wang, Z.: On the mathematical properties of the structural similarity index. IEEE Trans. Image Proc. **21**(4), 1488–1499 (2012)
3. Brunet, D., Vrscay, E.R., Wang, Z.: Structural similarity-based approximation of signals and images using orthogonal bases. In: Campilho, A., Kamel, M. (eds.) ICIAR 2010. LNCS, vol. 6111, pp. 11–22. Springer, Heidelberg (2010)
4. Richter, T., Kim, K.J.: A MS-SSIM optimal JPEG 2000 encoder. In: Data Compression Conference, Snowbird, Utah, pp. 401–410 (March 2009)
5. Wainwright, M.J., Schwartz, O., Simoncelli, E.P.: Natural image statistics and divisive normalization: Modeling nonlinearity and adaptation in cortical neurons. In: Rao, R., Olshausen, B., Lewicki, M. (eds.) Probabilistic Models of the Brain: Perception and Neural Function, pp. 203–222. MIT Press (2002)
6. Wang, Z., Bovik, A.C.: Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Signal Proc. Mag. **26**(1), 98–117 (2009)
7. 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)
8. Wandell, B.A.: Foundations of Vision. Sinauer Publishers, Sunderland (1995)