Applied Mathematics Letters 25 (2012) 1921–1925

Contents lists available at SciVerse ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

Geodesics of the Structural Similarity index

Dominique Brunet[a,*], József Vass[a], Edward R. Vrscay[a], Zhou Wang[b]

[a] Department of Applied Mathematics, University of Waterloo, Canada

[b] Department of Electrical and Computer Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 30 September 2011
Received in revised form 1 March 2012
Accepted 1 March 2012

Keywords:
Structural Similarity index
Quasi-hyperbolic metric
M-relative distance
Image distance metrics
Geodesics

ABSTRACT

We construct metrics from the geodesics of the Structural Similarity index, an image quality assessment measure. An analytical solution is given for the simple case of zero stability constants, and the general solution involving the numerical solution of a nonlinear equation is also found.

1. Introduction

In image quality assessment, one seeks for an objective measure that models the perceptual distance between two images. This measure can then be used to assess, design, or optimize image processing algorithms for compression or restoration of images.

Simple metrics such as L^1-distance and L^2-distance are often chosen due to their simplicity and nice mathematical properties, but they do not predict the human perception of image distortions well [1]. Recently, the Structural Similarity (SSIM) index [1] has been proposed as a simple alternative that yields better results in predicting perceived image quality with a low computational cost and at the same time sharing some of the properties of the L^2-norm [2].

In particular, the SSIM index can be expressed as the product of two components, each of which can be modified to form a metric [2]. But the SSIM index itself is not a metric, and thus it cannot be used to study the convergence of algorithms or to bound the distance between several images.

In [3], Richter addresses this problem from the viewpoint of differential geometry theory. He derives a system of second-order ordinary differential equations (ODEs) to be solved in order to describe the minimal path between two images (seen as a vector in $\mathbf{R}^N$). A solution was presented for a few very restricted particular cases.

In this paper, we analytically find the minimal path according to the SSIM index in the case when the stability constants are zeros. A numerical solution of the geodesics for positive stability constants is also given.

2. Background

Given two images $x, y \in \mathbf{R}^N_+$, the positive orthant, the Structural Similarity index combines information on the luminance (mean), contrast (variance), and structural distortion (linear correlation) as follows:

$$S(x, y) = S_{\epsilon_1}(\bar{x}, \bar{y}) S_{\epsilon_2}(x - \bar{x}, y - \bar{y}) = \frac{2\bar{x}\bar{y} + \epsilon_1}{\bar{x}^2 + \bar{y}^2 + \epsilon_1} \frac{2s_{x,y} + \epsilon_2}{s_x^2 + s_y^2 + \epsilon_2}. \quad (1)$$

* Corresponding author. Tel.: +1 519 888 4567x35469; fax: +1 519 746 4319.
E-mail addresses: dbrunet@uwaterloo.ca (D. Brunet), jvass@uwaterloo.ca (J. Vass), ervrscay@uwaterloo.ca (E.R. Vrscay), zhouwang@ieee.org (Z. Wang).

0893-9659/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aml.2012.03.001

Here, $\bar{x}, \bar{y}, s_x^2, s_y^2$, and $s_{x,y}$ represent, respectively, the sample mean of x and of y, the sample variance of x and of y, and the sample covariance between x and y. The constants ϵ_1 and ϵ_2 are small positive constants, set to ensure numerical stability of the division and chosen in order to model the saturation effect of the human visual system. Generally, the constants are chosen to be very small relative to the mean and variance of the images, so the case with $\epsilon_1 = \epsilon_2 = 0$ gives a good approximation of the SSIM index when the image signals have significant power. In fact, this is called the Universal Quality index, the previous version of the SSIM index (see [1]).

Since $S(x, y) = 1$ if and only if $x = y$, the Structural Similarity index is obviously not a metric. It leads us to consider either $1 - S(x, y)$ or $\sqrt{1 - S(x, y)}$ as possible candidates for a metric. Neither of these expressions satisfies the triangular inequality, but it was shown in [2] that

$$\sqrt{1 - S_\epsilon(x, y)} = \frac{\|x - y\|}{\|x\|^2 + \|y\|^2 + \epsilon} \tag{2}$$

is indeed a metric for all $\epsilon \geq 0$. In fact, it belongs to a class of metrics called normalized metrics (or M-relative distances) of the form (see [4])

$$d_n(x, y) = \frac{\|x - y\|}{M(\|x\|, \|y\|)}, \tag{3}$$

where $M : \mathbf{R}_+^2 \to \mathbf{R}_+$ and $\| \cdot \|$ is the Euclidian norm.

3. Theory

The goal is to find the intrinsic metrics induced by the minimal path of the Structural Similarity index or of other normalized metrics. Note that Martin and Osgood [5] describe the solution for $M(r, r) = r$ and that Hästö [4] found a solution from a similar technique for the case $M(r, r) = r^q$ with $0 < q < 1$. Given a curve parameterization $\gamma : [0, 1] \to \mathbf{R}^N$ with $\gamma(0) = x$ and $\gamma(1) = y$, the arc length formula according to a normalized metric is then, by definition,

$$L(\gamma) = \int_0^1 \frac{\|\gamma'(t)\|}{M(\|\gamma(t)\|, \|\gamma(t)\|)} dt. \tag{4}$$

Thus according to this metric, the element of distance is normalized with the magnitude of the curve.

We first study the geodesics for general M and then find an explicit solution for two particular cases of interest for the study of the SSIM index.

3.1. General case

It is clear that the arc length of a normalized metric is invariant under rotation. The problem thus reduces to two dimensions, and the solution can be expressed in polar coordinates:

$$\gamma(t) = (\rho(t) \cos(\phi(t)), \rho(t) \sin(\phi(t))), \tag{5}$$

where $\gamma(t)$ is on the plane generated by 0, x, and y, and where $\rho(t) = \|\gamma(t)\|$ and $\phi(t)$ is the angle between $\gamma(t)$ and x. Note that, if 0, x, and y are colinear, there might be either an infinite number of minimal paths or the minimal path is trivially the straight line. Let us now exclude these cases.

After the change of variables $\theta = \phi(t)$ and $r(\theta) = \rho(t)$, we can express the arc length in an expression of a single variable:

$$L(\gamma) = \int_\alpha^\beta \frac{\sqrt{\dot{r}^2 + r^2}}{M(r, r)} d\theta, \tag{6}$$

with $r(\alpha) = \rho(0)$ and $r(\beta) = \rho(1)$. Here, $\dot{r}$ denotes $dr/d\theta$, and we have assumed that $0 < |\beta - \alpha| < \pi$.

By the Euler–Lagrange equations, the stationary solution $r(\theta)$ must satisfy

$$\frac{\sqrt{\dot{r}^2 + r^2}}{M(r, r)} - \frac{\dot{r}^2}{\sqrt{\dot{r}^2 + r^2}} \frac{1}{M(r, r)} = c_1. \tag{7}$$

Necessarily $c_1 \neq 0$, since otherwise $r \equiv 0$. Solving (7) leads to

$$|\dot{r}| = r\sqrt{\frac{r^2}{c_1^2 M^2(r, r)} - 1}. \tag{8}$$

3.2. *Case 1: zero stability constant, $\epsilon = 0$*

In this case, M is homogeneous, i.e., $M(r, r) = r$. This includes the cases when $M(|x|, |y|) = (\frac{|x|^p+|y|^p}{2})^{1/p}$, $p > 0$, or $M(|x|, |y|) = \sqrt{|x||y|}$. It was found in [5] that the geodesic is a logarithmic spiral, and the length of this geodesic leads to the quasi-hyperbolic metric.

3.3. *Case 2: non-zero stability constant, $\epsilon > 0$*

The SSIM index with non-zero stability constants corresponds to the case $M(r, r) = \sqrt{r^2 + \epsilon^2}$ in (8). Without loss of generality, assume that $\alpha < \beta$ and $\|x\| \le \|y\|$. We denote

$$h(r) := \frac{r}{M(r, r)} = \frac{r}{\sqrt{r^2+\epsilon^2}}. \tag{9}$$

When $\dot{r} = 0$, we find from (8) that $h(r) = c_1$ and $r = \frac{c_1\epsilon}{\sqrt{1-c_1^2}}$; hence the differential equation has at most one critical point. It cannot be a local maximum, as it is clear that the shortest path must satisfy $r \le \|y\|$. Let (r^*, θ^*) be the coordinate for which the radius of the geodesic is minimal. If $\dot{r} \neq 0$, then, after separation of variables,

$$f(r, c_1) := \int \frac{c_1}{r\sqrt{h^2(r) - c_1^2}} dr = \frac{c_1}{\sqrt{1-c_1^2}} \operatorname{artanh}(v(r)) - \arcsin(u(r)), \tag{10}$$

$$\text{where } u(r) = \frac{c_1}{h(r)} \quad \text{and} \quad v(r) = \frac{\sqrt{1-c_1^2}}{\sqrt{h^2(r)-c_1^2}}. \tag{11}$$

Note that, since $c_1 < h(r) < 1$, $v(r) > 1$, and $f(r, c_1)$ always has $\frac{c_1}{\sqrt{1-c_1^2}} i\pi/2$ as an imaginary part. Also, for each r fixed,

$$f(r, h(r)) := \lim_{c_1 \to h(r)} f(r, c_1) = -\pi/2 + \frac{h(r)}{\sqrt{1-(h(r))^2}} i\pi/2, \tag{12}$$

$$f(r, 0) := \lim_{c_1 \to 0} f(r, c_1) = 0. \tag{13}$$

Since $h(r)$ is increasing in r, while $u(r)$ and $v(r)$ are decreasing in r, we see that the real part of $f(r, c_1)$ is increasing in r.

From the initial conditions,

$$f(\|x\|, c_1) - f(r^*, c_1) = \theta^* - \alpha, \tag{14}$$

$$f(\|y\|, c_1) - f(r^*, c_1) = \beta - \theta^*. \tag{15}$$

Adding both equations leads to the single nonlinear equation,

$$\beta - \alpha = f(\|x\|, c_1) + f(\|y\|, c_1) - 2f(r^*, c_1). \tag{16}$$

From (8), either (i) $r^*(\alpha) = \|x\|$, (ii) $\dot{r}(\theta^*) = 0$, or (iii) both. Note that (ii) implies that $h(r^*) = c_1$. In case (i), we have

$$f(\|y\|, c_1) - f(\|x\|, c_1) = \beta - \alpha, \tag{17}$$

whereas, in case (ii), we have

$$f(\|y\|, h(r^*)) + f(\|x\|, h(r^*)) - 2f(r^*, h(r^*)) = \beta - \alpha. \tag{18}$$

Depending on the sign of

$$f(\|y\|, h(\|x\|)) - f(\|x\|, h(\|x\|)) - (\beta - \alpha), \tag{19}$$

we are either in case (i), case (ii), or case (iii).

In case (i),

$$0 < \beta - \alpha < f(\|y\|, h(\|x\|)) - f(\|x\|, h(\|x\|)), \tag{20}$$

since, from (13),

$$\lim_{c_1 \to 0} f(\|y\|, c_1) - f(\|x\|, c_1) = 0. \tag{21}$$

By the intermediate value theorem, there exists a c_1 such that $0 < c_1 < h(\|x\|)$ for which (17) is satisfied.

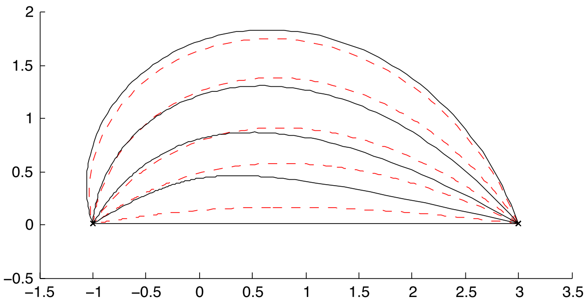

Fig. 1. Geodesics between (3, 0.01) and (−1, 0.01) for $M(r, r) = r^q$ with $q = 0, 1/4, 1/2, 3/4$, and 1 (full black) and for $M(r, r) = \sqrt{r^2 + \epsilon^2}$ with $\epsilon = \sqrt{0.1}, \sqrt{0.5}, \sqrt{0.9}, \sqrt{1.1}$, and $\sqrt{1.3}$ (dashed red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In case (ii),

$$\pi > \beta - \alpha > f(\|y\|, h(\|x\|)) - f(\|x\|, h(\|x\|)), \tag{22}$$

since, from (12) and (13),

$$\lim_{r^*\to 0}(f(\|y\|, h(r^*)) + f(\|x\|, h(r^*)) - 2f(r^*, h(r^*))) = \pi. \tag{23}$$

By the intermediate value theorem, there exists an r^* such that $0 < r^* < \|x\|$ for which (18) is satisfied.

For case (iii), we simply have

$$\beta - \alpha = f(\|y\|, h(\|x\|)) - f(\|x\|, h(\|x\|)). \tag{24}$$

In this last case, both (17) and (18) have a solution for $c_1 = h(\|x\|)$ and $r^* = \|x\|$.

The trajectory can then be drawn by writing θ as a function of r:

$$\theta = \alpha + f(\|x\|, c_1) - f(r, c_1) \quad \text{for } \|x\| \geq r \geq r^* \tag{25}$$

and

$$\theta = \beta - (f(\|y\|, c_1) - f(r, c_1)) \quad \text{for } r^* \leq r \leq \|y\|. \tag{26}$$

4. Results

Examples of geodesics in $\mathbf{R}^2$ are shown in Fig. 1 for various values of ϵ.

Following [2], any element x in $\mathbf{R}^N$ can be decomposed into two orthogonal components containing, respectively, the mean of x and a zero-mean component obtained by orthogonal projection. Thus $x = \bar{x}\mathbf{1} \oplus (x - \bar{x}\mathbf{1})$. According to Richter [3], one can solve for $\bar{x}$ and $x - \bar{x}$ independently to find the minimal path according to the SSIM index.

Denote $\omega = \arccos(s_{x,y}/s_x s_y)$. For the case $\epsilon_1 = \epsilon_2 = 0$, we apply the solution of the minimal path for $\bar{x}$ and $x - \bar{x}$:

$$\gamma(t) = \bar{x}^{1-t}\bar{y}^t + s_x^{1-t}s_y^t\left(\frac{\sin(\omega(1-t))}{\sin(\omega)}\frac{x-\bar{x}}{s_x} + \frac{\sin(\omega t)}{\sin(\omega)}\frac{y-\bar{y}}{s_y}\right). \tag{27}$$

As a direct application, given two still images x and y, one may wish to create a transition (cross-fading) from x to y that preserves most structural information from both images. For example, it could be used in slide transition in a picture slideshow or in scene transition in a movie. The goal is to make a transition that is perceptually linear. So, instead of applying simply a linear interpolation $z(t) = tx+(1-t)y$ that would lead to a nonlinear perception change, one may apply the formula for the SSIM geodesic. Fig. 2 shows an example of the SSIM geodesic. Some SSIM-interpolated images are compared with their linear interpolation counterparts.

5. Conclusions

The geodesic of the Universal Quality index was derived analytically and the geodesic of the Structural Similarity index was solved by a combination of analytical and numerical methods. This solves a problem posed by Richter in [3] and shows another way to derive a metric from the SSIM index.

Acknowledgments

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and in part by the Province of Ontario Ministry of Research and Innovation.

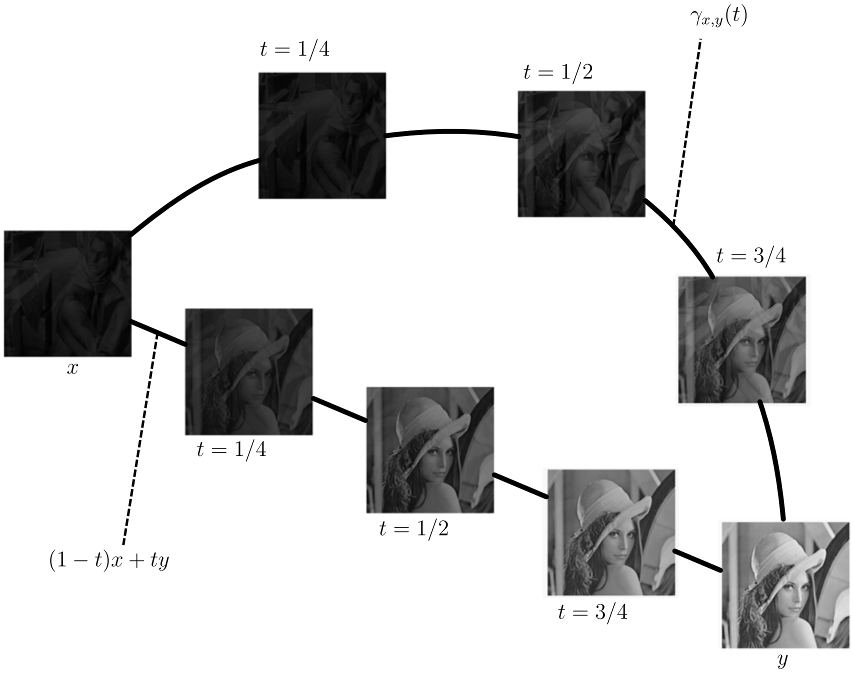

Fig. 2. Image geodesics between Barbara (low luminance and contrast) and Lena according to the SSIM index and the Euclidian metric (linear path).

References

[1] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Processing 13 (2004) 600–612.
[2] D. Brunet, E.R. Vrscay, Z. Wang, On the mathematical properties of the structural similarity index, IEEE Trans. Image Processing 21 (4) (2012) 1488–1499.
[3] T. Richter, SSIM as global quality metric: a differential geometry view, in: Third International Workshop on Quality of Multimedia Experience, QoMEX 2011, IEEE, Mechelen, Belgium, 2011, pp. 189–194. September 7–9.
[4] P.A. Hästö, A new weighted metric: the relative metric II, J. Math. Anal. Appl. 301 (2005) 336–353.
[5] G. Martin, B.G. Osgood, The quasihyperbolic metric and the associated estimate on the hyperbolic metric, J. Anal. Math. 47 (1986) 37–53.