LOCALLY MONOTONIC MODELS FOR IMAGE AND VIDEO PROCESSING

Scott T. Acton
School of Electrical & Computer Engineering
Oklahoma State University
Stillwater OK 74078, USA
sacton@okstate.edu

Alfredo Restrepo (Palacios)
Dpt. Ing. Eléctrica y Electrónica
Universidad de los Andes
A.A. 4976, Santafé de Bogotá, Colombia
arestrep@uniandes.edu.co

ABSTRACT

Definitions of locally monotonic images are introduced. The model definitions are complemented with algorithms that compute locally monotonic versions of a given image or video frame input. The property of local monotonicity provides a useful vehicle for image smoothing and denoising. Local monotonicity is also useful for scale space generation, wherein the degree of local monotonicity is the scale parameter. Currently, the property of local monotonicity is well defined for the 1-D case, but is not well defined for images or video. In this paper, models for multidimensional local monotonicity that extend the 1-D definition are rendered. Regression-based and diffusion-based processing methods are prescribed that yield meaningful locally monotonic images. The definitions and associated algorithms are applicable to image enhancement and a variety of multiscale tasks such as image segmentation and video coding.

1. INTRODUCTION

Signal smoothness and scale are fundamental qualities used in digital signal processing, analysis and interpretation. Both the smoothness and scale of a 1-D digital signal can be quantified by the highest *degree* of local monotonicity maintained by the signal [6]. A 1-D LOMO signal is defined as follows:

Definition 1: A 1-D signal is locally monotonic of degree d (or LOMO-d) if every interval of length d is monotonic (non-decreasing or non-increasing).

So, local monotonicity is well defined in 1-D and has been used in digital signal analysis, including, for example, the root properties of the median filter [9]. LOMO signals are desirable because ramp and step edges (gradual and abrupt transitions) are allowed and impulses are not allowed. For multidimensional data, especially for digital images, an extension of the 1-D definition is requisite. Previous work in multidimensional local monotonicity utilized the following definition [3]: In two (or more) dimensions, the signal is LOMO-d if it is LOMO-d in the 1-D sense along connected paths in defined orientations. The enhancement techniques in [3] and the restoration techniques in [2] essentially enforced local monotonicity along the image rows and columns.

Here, we provide alternative, more precise definitions of multidimensional local monotonicity that lead to well-defined sets of images.

First, models for multidimensional local monotonicity that are derived directly from the 1-D definition are discussed. Then, models that are inherently multidimensional are introduced. For both the extended models and the direct M-D models, viable algorithms are provided that generate the LOMO images according to the image models, and image examples are provided.

2. MULTIDIMENSIONAL LOCAL MONOTONICITY

To extend the 1-D definition of local monotonicity to the multidimensional case, we may enforce local monotonicity along prescribed 1-D paths. We first define the *orientation set* for images, which is used in the definition of LOMO images.

Definition 2: The orientation set O is the set of allowed orientations. In the continuous-domain case, this set can include all possible orientations, which we denote O^*. For the common square tessellation (tiling), we define the following three special orientation sets: O^+ denotes the set of horizontal and vertical orientations {N-S, E-W}, while O^x denotes the set of diagonal orientations {NW-SE, SW-NE}. In the discrete case, O^* is the set of horizontal, vertical and diagonal orientations {N-S, E-W, NW-SE, SW-NE}.

2.1 Strong Local Monotonicity

In the M-D LOMO models derived from the 1-D model, we make a distinction between *strong* local monotonicity and *weak* local monotonicity. Strong LOMO

0-7803-5467-2/99/ $10.00 © 1999 IEEE

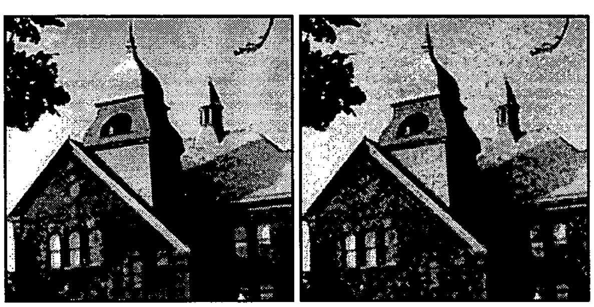
models require the image to be LOMO along *all* linear paths allowed by the orientation set. With weak LOMO models, this definition is relaxed.

Definition 3: An image is locally monotonic in the strong sense (LOMO-*d*) if each 1-D (straight) path allowed by the orientation set is LOMO-*d*.

Further distinctions are given by specific orientation sets. An image is LOMO*-*d* if each straight path through the image is LOMO-*d*. For digital images with square tessellation, we restrict the paths to the horizontal, vertical and diagonal orientations (using the orientation set O^*). Thus, a discrete-domain LOMO*-*d* image is LOMO-*d* along four orientations in the 1-D sense. An image is LOMO$^+$-*d* if the image is LOMO-*d* in two defined orthogonal orientations. In the discrete case, we have LOMO$^+$-*d* and LOMO^x-*d*. The LOMO$^+$-*d* images use the orientation set O^+ and are similar to the images defined in [2] and [3]. Likewise, the LOMO^x-*d* images are LOMO-*d* along the two diagonal orientations (using orientation set O^x).

(a) (b)

Figure 1 (a) Original "Old Central" image; (b) with additive Laplacian-distributed noise (stand. dev. $\sigma = 20$).



(a) (b)

Figure 2 An example of the strong LOMO model using suboptimal LOMO$^+$-3 regression. (a) Result using Fig. 1(a) as input; (b) result using Fig. 1(b) as input.

The Suboptimal Regression Algorithm

Ideally, we would like to produce an optimal LOMO-*d* image – the closest LOMO-*d* image to the input image. This *LOMO regression* is given by

$$\mathbf{I}^* = \arg \min_{\mathbf{H} \in \mathbf{LOMO}-d} \| \mathbf{I} - \mathbf{H} \|_D . \qquad (1)$$

Here, the estimate $\mathbf{I}^*$ is the image closest to the observed image $\mathbf{I}$, among all images that lie *within* the set of LOMO-*d* images (denoted by **LOMO-*d***). The term $\| \mathbf{I} - \mathbf{H} \|_D$ gives the distance between image $\mathbf{H}$ and the observed image $\mathbf{I}$, defined by an appropriate distance norm $\| \cdot \|_D$.

In 1-D, algorithms have been given that produce LOMO regressions [6] and [8]. In 2-D, a Viterbi-based algorithm has been presented for binary images [7], but no practical 2-D LOMO regression algorithm exists currently for grayscale imagery. To compute LOMO images that are close to the original image, we attempt to minimize the distance between the input observation $\mathbf{I}$ and the solution $\mathbf{H}$ while also attempting to minimize the deviation from strong local monotonicity. The problem can be approached by minimizing the following energy functional:

$$E(\mathbf{H}) = \| \mathbf{I} - \mathbf{H} \| + \lambda \| \mathbf{LOMO}(\mathbf{H}) \|. \qquad (2)$$

The minimal energy solution to (2) is

$$\hat{\mathbf{H}} = \arg \min_{\mathbf{H}} \{E(\mathbf{H})\}. \qquad (3)$$

Here λ is a regularization parameter that balances the penalties for deviation from the input image $\mathbf{I}$ and for deviation from the set of LOMO-*d* images. $\mathbf{LOMO}(\mathbf{H})$ gives a penalty for violating local monotonicity on a local basis in the image.

To find acceptable (but not optimal) solutions to the nonconvex energy function of (2), we use the generalized deterministic annealing algorithm (GDA) [4]. GDA is a deterministic approximation of stochastic simulated annealing. In GDA, each pixel intensity is modeled by a "local" Markov chain. The stationary distributions of these Markov chains are computed as the annealing temperature is lowered, freezing the solution in a local minimum. Another difficulty with the energy minimization approach is the selection of the regularization parameter. Given the presence of nonlinear operators in (2), the regularization parameter is determined via cross validation [2]. An example result from the suboptimal LOMO regression procedure is shown in Fig. 2. The output image obtained from the noisy image, shown in Fig. 2(b), demonstrates the tendency of the suboptimal regression method to produce blotching artifacts.

The Strong Diffusion Algorithm

A less expensive method that does not produce blotching artifacts involves a *PDE* (partial differential equation) approach. PDE's can be used to enact anisotropic diffusion of the image intensities – an adaptive smoothing that preserves edges [1]. A typical discrete anisotropic diffusion update is given by:

$$I(x) \leftarrow I(x) + \sum_{p=1}^{\Omega} \nabla I_p(x) c_p(x) \quad (4)$$

where p represents the diffusion direction, either "E", "W", "N" or "S" in the 2-D case, and Ω is the number of diffusion directions allowed by the orientation set. $\nabla I_p(x)$ is the first partial derivative approximation in the p^{th} direction, and $c_p(x)$ is the diffusion coefficient in the p^{th} direction. As the diffusion coefficient approaches zero (near high gradient magnitude), the diffusion process is impeded. In relatively smooth areas, the diffusion coefficient should be near one in value.

To obtain a set of PDE's that generate LOMO signals, we attempt to remove signal features that cause both increases and decreases in signal intensity within a local neighborhood. A PDE that limits the sign changes of pixel differences within a local neighborhood may be designed using the following diffusion coefficient:

$$c_p(x) = \frac{1}{|\nabla I_p(x)|} \quad (5)$$

Substituting this diffusion coefficient into (4), we have the following LOMO diffusion iterate for images:

$$[I(x)]_{t+1} \leftarrow \left(I(x) + (1/4) \left\{ \begin{matrix} \mathrm{sgn}[\nabla I_W(x)] + \mathrm{sgn}[\nabla I_E(x)] \\ + \mathrm{sgn}[\nabla I_N(x)] + \mathrm{sgn}[\nabla I_S(x)] \end{matrix} \right\} \right)_t . \quad (6)$$

(a) (b)

Figure 3 An example of strong LOMO$^+$-3 using the strong diffusion PDE. (a) Result using Fig. 1(a) as input; (b) result using Fig. 1(b) as input.

After 256 iterations of (6), we obtain the diffusion results shown in Fig. 3. Note that features are preserved and noise is removed without the introduction of artifacts. A drawback of this method is over-smoothing – a coarser scale image is generated by strong LOMO diffusion, in comparison to the regression approach.

2.2 Weak Local Monotonicity

A relaxed form of local monotonicity for images is defined as follows:

Definition 4: An image is locally monotonic in the weak sense (weak lomo-*d*) if it is 1-D LOMO-*d* in at least one orientation in the orientation set at each point. This means that there exists a direction at each point in which all *d*-length 1-D intervals containing that point are LOMO-*d*.

An image is lomo*-*d* if the image is weak lomo-*d* using the set of all allowed orientations. With digital images, the property of local monotonicity must hold over at least one of the four defined orientations of O^* at each point. A lomo$^+$-*d* image is LOMO-*d* row-wise or column-wise at each point in the image. The lomo^x-*d* images are LOMO-*d* along one of the two diagonal orientations at each point.

(a) (b)

Figure 4 An example of the weak lomo$^+$-3 diffusion algorithm. (a) Result using Fig. 1(a) as input; (b) result using Fig. 1(b) as input.

The Weak Diffusion Algorithm

A weak lomo image can be generated by utilizing a modified version of the diffusion PDE in (6). Here, 1-D diffusion occurs at each image location in the direction that is "closest" to becoming locally monotonic. For example, given the orientation set O^+, the diffusion PDE is

$$[I(x)]_{t+1} \leftarrow \left\{ \begin{matrix} I(x) + \frac{\beta}{2}[\mathrm{sgn}(\nabla_W) + \mathrm{sgn}(\nabla_E)] \\ + \frac{1-\beta}{2}[\mathrm{sgn}(\nabla_N) + \mathrm{sgn}(\nabla_S)] \end{matrix} \right\}_t . \quad (7)$$

In (7), $\beta = 1_{[\min(|\nabla_W|,|\nabla_E|)<\min(|\nabla_N|,|\nabla_S|)]}$, and $1_{[\ ]}$ is the indicator function, which equals unity if the parenthetical expression is true and zero otherwise. So, if the pixel in question is closer in value to the western or eastern neighbor, diffusion is implemented in the east-west direction. Otherwise, diffusion occurs in the north-south direction. Fig. 4 provides an example of the weak lomo diffusion technique. In this case, denoising and scaling are accomplished without the over-smoothing problems caused by the strong LOMO model.

A further relaxed version of weak local monotonicity may be defined using contiguous subsequences centered at each point in the image, as in Definition 5.

Definition 5: If there exists a d-length contiguous subsequence *centered* at each point in the image, allowed by the orientation set, that is monotonic, then the image is lomo-d^c.

Instead of requiring all d-length intervals containing a given point to be LOMO-d in the 1-D sense, only the interval centered at the particular point must be monotonic when using Definition 5. In this case, the degree d must be an odd integer for discrete-domain imagery.

Closest Neighbor Algorithm

For the case of the lomo-3^c image model, we can use a simple algorithm to transform a given image into a lomo-3^c image. At each point in the image, we evaluate the existence of a length-3 monotonic segment in accordance with Definition 5. At points where the definition is violated, we update the pixel intensity to its closest neighbor, as allowed by the orientation set. For lomo*-3^c, each of the eight neighbors is considered. An example of this simple weak lomo algorithm is given in Fig. 5. Note that this weak lomo model is not sufficient for denoising (see Fig. 5(b)). Furthermore, the closest neighbor approach will not be successful for degrees of local monotonicity greater than three.

3. DIRECT MULTIDIMENSIONAL LOCALLY MONOTONIC MODELS

The previous models given for multidimensional local monotonicity represent attempts to extend the 1-D interval-based definition to higher dimensioned domains. In this section, we introduce two LOMO models for images that are based on monotonicity within image neighborhoods. Using the concepts of structuring elements from morphology and filter windows from nonlinear filtering, direct multidimensional LOMO definitions can be implemented.

3.1 Morphological Local Monotonicity

With 1-D signals, a LOMO signal $\mathbf{I}$ will maintain the following property: $\mathbf{I} \bullet \mathbf{B} = \mathbf{I} \circ \mathbf{B}$ for some $\mathbf{B}$. Here, $\mathbf{I} \bullet \mathbf{B}$ is the morphological closing of $\mathbf{I}$ by structuring element $\mathbf{B}$, and $\mathbf{I} \circ \mathbf{B}$ is the opening of $\mathbf{I}$ by $\mathbf{B}$. We can extend this definition to 2-D images.

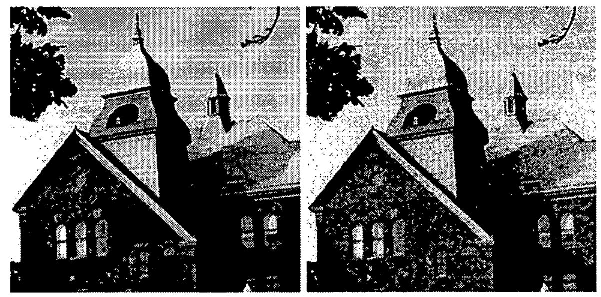
Definition 6: An image $\mathbf{I}$ is morphologically LOMO of degree $\mathbf{B}$ (LOMO^{M}-$\mathbf{B}$) if $\mathbf{I} \bullet \mathbf{B} = \mathbf{I} \circ \mathbf{B}$, where $\mathbf{B}$ is a structuring element.

The Morphological Algorithm

An update scheme based on morphological filters can be used to "force" regions of the image to become morphologically LOMO. We attempt to generate LOMO^{M}-$\mathbf{B}$ images by iterating on

$$\mathbf{I} \leftarrow ((\mathbf{I} \bullet \mathbf{B}) \circ \mathbf{B} + (\mathbf{I} \circ \mathbf{B}) \bullet \mathbf{B})/2 \qquad (8)$$

where the scale of the resultant signal is determined by the size of the structuring element $\mathbf{B}$. Although the average of the opening and closing alone can be used to produce morphologically LOMO signals, an improved rate of convergence is obtained by utilizing the open-close and close-open filters, as in (8).



(a) (b)

Figure 5 An example of the weak lomo*-3^c closest neighbor algorithm. (a) Result using Fig. 1(a) as input; (b) result using Fig. 1(b) as input.

Typically, fewer than ten iterations are required to reach a root image that is a fixed point. The iterate in (8) converges to a root image that is morphologically LOMO except at "saddlepoints." The saddlepoints exist where the signal is simultaneously a ridge (a local maximum in one direction) and a valley (a local minimum in one direction). For more information of the morphological algorithm see [5]. The images shown in Fig. 6 are obtained by iterating (8) ten times on the corresponding images in Fig. 1. One negative aspect of the morphological approach is the influence of the structuring element on the result images. Using a circular structuring element, sharp features and corners are lost as in Fig. 6. The advantage of the morphological approach is the ability to preserve certain shapes in a scaled LOMO image. Moreover, the morphological approach extends easily to video, 3-D and M-D data.

3.2 Window-based Local Monotonicity

Alternatively, we can define windows that are LOMO. Then, a 2-D LOMO image is defined as an image that consists of LOMO windows.

Definition 7: A monotonic window is a subimage that has at least one orientation in the orientation set in which each path is nonincreasing or each path is nondecreasing.

Definition 8: An image $\mathbf{I}$ is LOMO of degree $\mathbf{B}$ (LOMO-$\mathbf{B}$) if each subimage determined by the structuring element $\mathbf{B}$ and fully contained in the image, is monotonic.

For example, if we consider a 3x3 window under O^+, there are four possible conditions that would make the subimage a monotonic window. The three rows could be nonincreasing; the three rows could be nondecreasing; the three columns could be nonincreasing, or the three columns could be nondecreasing. Otherwise, the window is not monotonic. In the same image, some windows may be LOMO in one orientation, such as along the rows, where other windows may be LOMO in another orientation, such as along the columns.

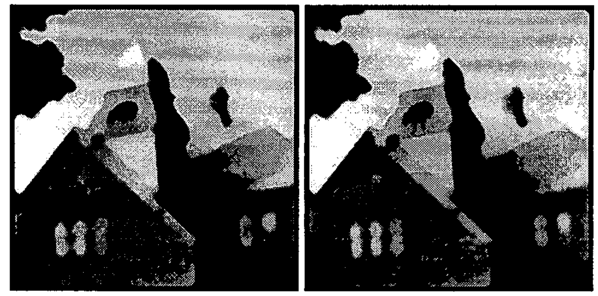
Algorithms for generating window-based LOMO images are under development. The strengths of the window-based definition include edge-preservation and the ability to extend to higher dimensions.

4. CONCLUSIONS AND FUTURE WORK

The LOMO models and associated algorithms should be compared using metrics of quality and computational complexity. One important quality metric is fidelity to the original image for a given scale. We have listed mean squared error (MSE) results associated with the algorithms in the Table for the case of local monotonicity of degree 3. However, a scale generating process cannot be evaluated solely by MSE. The type of features retained, edge preservation, and noise reduction are important qualities for image segmentation, enhancement, and feature extraction for image analysis. For example, the weak closest neighbor algorithm gives the lowest MSE values, but is incapable of removing noise or removing small-scale objects.

The computational cost of each algorithm introduced in this paper has been given in the Table. With the localized, iterative update schemes, the computational cost is driven by the number of updates needed for each pixel, rather than the complexity of the updates. In the Table, K denotes the number of intensity levels used (e.g., $K = 256$ for 8-bit grayscale imagery). So, the regression approach is the most expensive method at $O(10K)$ (order $10K$) updates per pixel. The least expensive method is the simple closest neighbor algorithm that requires only two passes through the image in typical cases. One attractive feature of the algorithms given is that the number of updates does not increase with additional pixels (larger images).

This paper serves as a commencement in the exploration of multidimensional local monotonicity. We have defined several feasible models for locally monotonic images and have designed techniques that generate images in accordance with the models. The properties of these models and the performance of the algorithms need further analysis. The usefulness of the LOMO models will be evaluated in their future application to multidimensional signal processing problems such as image segmentation and object-based image coding.



(a) (b)

Figure 6 An example of the LOMO^M-**B** morphological algorithm. (a) Result using Fig. 1(a) as input; (b) result using Fig. 1(b) as input. Here, **B** is a circular structuring element with diameter = 9.

TABLE: Comparison of Algorithms

Algorithm	MSE from Fig. 1(a)	MSE from Fig. 1(b)	No. of Updates
Sub. Regression	114.3	365.1	$O(10K)$
Strong Diffusion	526.8	858.1	$O(K)$
Weak Diffusion	481.7	808.8	$O(K)$
Closest Neighbor	14.2	70.1	2
Morphological	223.3	504.2	$O(10)$

5. REFERENCES

[1] S.T. Acton, "Anisotropic diffusion and local monotonicity," *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing* (ICASSP-98), Seattle, May 12-15, 1998.

[2] S.T. Acton and A.C. Bovik, "Piecewise and local image models for regularized image restoration using cross validation," *IEEE Transactions on Image Processing*, vol.8, pp. 652-665, 1999.

[3] S.T. Acton and A.C. Bovik, "Nonlinear image estimation using piecewise and local image models," *IEEE Trans. Image Processing*, vol. 7, pp. 979-991, 1998.

[4] S.T. Acton and A.C. Bovik, "Generalized deterministic annealing," *IEEE Trans. Neural Networks*, vol. 7, pp. 686-699, 1996.

[5] J. Bosworth and S.T. Acton, "The morphological lomo filter for multiscale image processing," *Proc. IEEE Int. Conf. on Image Processing*, Kobe, Japan, Oct. 25-29, 1999.

[6] A. Restrepo (Palacios) and A.C. Bovik, "Locally monotonic regression," *IEEE Trans. Signal Process.*, vol. SP-41, pp. 2796-2810, 1993.

[7] A. Restrepo (Palacios) and S.T. Acton, "2-D binary locally monotonic regression," *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-99)*, Phoenix, March 14-19, 1999.

[8] N. Sidiropoulos, "Fast digital locally monotonic regression," *IEEE Trans. on Signal Processing*, vol. 45, pp. 389-395, 1997.

[9] S.G. Tyan, "Median filtering: deterministic properties," in *Two-dimensional Signal Processing: Transforms and Median Filters*, T.S. Huang, ed. New York: Springer-Verlag, 1981.