

Fuzzy Partitions and Multi-Scale Representation of 1D Data

***Irina Perfilieva^a and David Adamczyk^a**

^aCE IT4I - IRAFM, University of Ostrava, 30. dubna 22, Ostrava, Czech Republic

Irina.Perfilieva, David.Adamczyk@osu.cz

Abstract

We analyze various multi-scale representations of a one-dimensional signal in spaces with a closeness relation determined by a symmetric and positive semi-definite kernel. We discuss other kernels than the predominant Gaussian kernel in a scale space representation of a one-dimensional signal. We show that kernels arising from generating functions of fuzzy partitions can be treated with the same success. We also show that the reconstruction from the proposed multi-scale representations of better quality than the reconstruction from MLP with almost double the number of neurons in 4 hidden layers.

Keywords: Multi-scale representation, Fuzzy partition, Fuzzy transform.

1 Introduction

The proposed contribution is focused on a multi-scale representation of image/signal data. This processing closely relates to the data processing by neural networks, extraction of stable features that are invariant under various geometric transformations.

Among the first types of multi-scale representation of image data, the quad-tree methodology [3] should be mentioned. It focuses on recursively dividing an image into smaller areas. A low-pass pyramid representation was proposed in [2], where the added benefit to multi-scaling was that the image size decreased exponentially compared to. scale level.

Koenderink [4] emphasized that the problem of scale must be solved in any visualization situation. Any real image has a limited size, determined by two scales: external and internal. For a digital image, the internal scale is determined by the pixel size, and for a

photographic image, by the size of the grains in the emulsion. We can say that the outer scale of an object corresponds to the (minimum) size of the window that completely contains that object, while (informally) the inner scale corresponds to the scale where the features of the object begin to appear.

To have a successful image analysis we need the ability to scale up and down the internal scope of observation in accordance with the specific task. Koenderink also emphasized that if there is no a priori reason for considering specific image structures, then the successful analysis should be able to handle image structures at all scales. The challenge is to understand the image at all relevant scales at the same time, but not as an unrelated set of derived images at different levels of blur.

The basic idea (in Lindeberg [5]) how to obtain a multi-scale representation of an object is to embed it into a one-parameter family of gradually smoothed ones where fine-scale details are sequentially suppressed. Under fairly general conditions, the author showed that the Gaussian kernel and its derivatives are the only possible smoothing kernels. These conditions are mainly linearity and shift invariance, combined with various ways of formalizing the notion that structures on a coarse scale should correspond to simplifications of corresponding structures on a fine scale.

The main result that we arrive at here is that the Gaussian kernel can be replaced with the same success by a symmetric positive semi-definite kernel with a local support. In particular, we will show that generating function of any uniform fuzzy partition of \mathbb{R} can be used for determining such kernel.

2 Scale-space Representation

A scale-space representation is a special type of multi-scale representation that is determined by a continuous scale parameter and preserves the same spatial sampling at all scales. By the construction in Witkin [8],

a scale-space representation is a one-parameter family of derived signals constructed using convolution with a one-parameter family of Gaussian kernels of increasing width.

Formally, a scale space family of a continuous signal is constructed as follows. For a signal $f : \mathbb{R}^N \rightarrow \mathbb{R}$, the scale-space representation $L : \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$ is defined by:

$$\begin{aligned} L(\cdot, 0) &= f(\cdot), \\ L(\cdot, t) &= g(\cdot, t) \star f, \end{aligned} \quad (1)$$

where $t \in \mathbb{R}_+$ is the scale parameter and $g : \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$ is the Gaussian kernel as follows:

$$g(x, t) = \frac{1}{(2\pi t)^{N/2}} \exp - \sum_{i=1}^N \frac{x_i^2}{2t}.$$

The scale parameter t relates to the standard deviation of the kernel g , and is a natural measure of spatial scale at the level t .

As an important remark, we note that the scale-space family L can be defined as the solution to the diffusion (heat) equation

$$\partial_t L = \frac{1}{2} \nabla^T \nabla L, \quad (2)$$

with initial condition $L(\cdot, 0) = f$. The Laplace operator, $\nabla^T \nabla$ or Δ , the divergence of the gradient, is taken in the spatial variables. The ‘‘diffusivity constant’’ $\frac{1}{2}$ in (2) characterizes the ‘‘thermal diffusivity of the medium’’, which, among other things, depends on its mass density.

The solution to (2) in one-dimension and in the case where the spatial domain is R is known as the convolution of f (initial condition) and the fundamental solution:

$$L(x, t) = g(\cdot, t) \star f, \quad (3)$$

$$g(x, t) = \frac{1}{(\sqrt{2\pi t})} \exp - \frac{x^2}{2t}. \quad (4)$$

The following two questions arise: is this approach the only reasonable way to perform low-level processing, and are Gaussian kernels and their derivatives the only smoothing kernels that can be used? Many authors [5, 8, 4] answer these questions positively, which leads to the default choice of Gaussian kernels in most image processing tasks. In this article, we want to expand on the set of useful kernels suitable for performing scale-space representations. In particular, we propose to use kernels arising from generating functions of fuzzy partitioning.

3 Space with a Fuzzy Partition

In this section, we introduce space that plays an important role in our research. A space with a fuzzy partition is considered as a space with a proximity (closeness) relation, which is a weak version of a metric space. Our goal is to show that the diffusion (heat conduction) equation in (2) can be extended to spaces with closeness, where the concepts of derivatives are adapted to nonlocal cases.

Let us first recall the basic definitions. As we indicated at the beginning, our goal is to extend the Laplace operators to those that take into account the specifics of spaces with fuzzy partitions. For this reason, in the following sections, we recall the basic concepts on this topic.

3.1 Fuzzy partition

Definition 1: Fuzzy sets $A_1, \dots, A_n : [a, b] \rightarrow \mathbb{R}$, establish a *fuzzy partition* of the real interval $[a, b]$ with nodes $a = x_1 < \dots < x_n = b$, if for all $k = 1, \dots, n$, the following conditions are valid (we assume $x_0 = a$, $x_{n+1} = b$):

1. $A_k(x_k) = 1$, $A_k(x) > 0$ if $x \in (x_{k-1}, x_{k+1})$;
2. $A_k(x) = 0$ if $x \notin (x_{k-1}, x_{k+1})$;
3. $A_k(x)$ is continuous,
4. $A_k(x)$, for $k = 2, \dots, n$, strictly increases on $[x_{k-1}, x_k]$ and $A_k(x)$ strictly decreases on $[x_k, x_{k+1}]$ for $k = 1, \dots, n-1$,

The membership functions A_1, \dots, A_n are called *basic functions* [6].

Definition 2: The fuzzy partition A_1, \dots, A_n , where $n \geq 2$, is *h-uniform* if nodes $x_1 < \dots < x_n$ are *h-equidistant*, i.e. for all $k = 1, \dots, n-1$, $x_{k+1} = x_k + h$, where $h = (b-a)/(n-1)$, and the following additional properties are fulfilled [6]:

1. for all $k = 2, \dots, n-1$ and for all $x \in [0, h]$, $A_k(x_k - x) = A_k(x_k + x)$,
2. for all $k = 2, \dots, n-1$, and for all $x \in [x_k, x_{k+1}]$, $A_k(x) = A_{k-1}(x-h)$, and $A_{k+1}(x) = A_k(x-h)$.

Proposition 1: If the fuzzy partition A_1, \dots, A_n of $[a, b]$ is *h-uniform*, then there exists an even function $A_0 : [-1, 1] \rightarrow [0, 1]$, such that for all $k = 1, \dots, n$:

$$A_k(x) = A_0 \left(\frac{x - x_k}{h} \right), \quad x \in [x_{k-1}, x_{k+1}].$$

A_0 is called a *generating function* of uniform fuzzy partition [6].

Remark. We further assume that the generation function A_0 is such that

$$\int_{-1}^1 A_0(x) dx = 1.$$

Under this assumption, the h -rescaled generating function $A_h(x) = A_0(x/h)$ of an h -uniform fuzzy partition produces the corresponding to it kernel $A_h(x-y)$ and the normalized kernel $\frac{1}{h}A_h(x-y)$, so that for all $x \in \mathbb{R}$,

$$\frac{1}{h} \int_{-\infty}^{\infty} A_h(x-y) dy = 1.$$

3.2 Discrete Universe and Its Fuzzy Partition

From the point of view of image / signal processing, we assume that the domain of the corresponding functions is finite, i.e. finitely sampled in \mathbb{R} , and the functions are identified with high-dimensional vectors of their values at the selected samples in the discretized domain. Moreover, we assume that the domain and the range of all considered functions are equipped with the corresponding relations of closeness.

The best formal model of all these assumptions is a weighted graph $G = (V, E, w)$ where $V = \{v_1, \dots, v_\ell\}$ is a finite set of vertices, and E ($E \subset V \times V$) is a set of weighted edges so that $w : E \rightarrow \mathbb{R}_+$. The edge $e = (v_i, v_j)$ connects two vertices v_i and v_j , and then the weight of e is $w(v_i, v_j)$ or just w_{ij} . Weights are set using the function $w : V \times V \rightarrow \mathbb{R}_+$, which is symmetric ($w_{ij} = w_{ji}, \forall 1 \leq i, j \leq \ell$), non-negative ($w_{ij} \geq 0$) and $w_{ij} = 0$ if $(v_i, v_j) \notin E$. The notation $v_i \sim v_j$ denotes two adjacent vertices v_i and v_j with an existing edge connecting them.

Let $H(V)$ denote the Hilbert space of real-valued functions on the set of vertices V of the graph, where if $f, h \in H(V)$ and $f, h : V \rightarrow \mathbb{R}$, then the inner product $\langle f, h \rangle_{H(V)} = \sum_{v \in V} f(v)h(v)$. Similarly, $H(E)$ denotes the space of real-valued functions defined on the set E of edges of a graph G . This space has the inner product $\langle F, H \rangle_{H(E)} = \sum_{(u,v) \in E} F(u,v)H(u,v) = \sum_{u \in V} \sum_{v \sim u} F(u,v)H(u,v)$, where $F, H : E \rightarrow \mathbb{R}$ are two functions on $H(E)$.

We assume that the set of vertices V is identified with the set of indices $V = \{1, \dots, \ell\}$ and that $[1, \ell]$ is h -uniform fuzzy partitioned with normalized basic functions A_1^h, \dots, A_ℓ^h , so that $A_k^h(x) = A_h(x-k)/h$, $k = 1, \dots, \ell$, $A_h(x) = A_0(x/h)$ and A_0 is the generating function.

Definition 3: A weighted graph $G = (V, E, w)$ is fuzzy weighted, if $V = \{1, \dots, \ell\}$, A_1^h, \dots, A_ℓ^h is an h -uniform

fuzzy partition, generated by A_0 , and $w_{ij} = A_i^h(j)$, $i, j = 1, \dots, \ell$. The fuzzy weighted graph $G = (V, E, w)$, corresponding to the h -uniform fuzzy partition, will be denoted $G_h = (V, E, A_h)$.

4 Discrete Laplace operator

In this section, we recall the definition of (non-local) Laplace operator as a differential operator given by the divergence of the gradient of a function (see [1]).

Let $G = (V, E, w)$ be a weighted graph, and let $f : V \rightarrow \mathbb{R}$ be a function in $H(V)$. The difference operator $d : H(V) \rightarrow H(E)$ of f , is defined on $(u, v) \in E$ by

$$(df)(u, v) = \sqrt{w(u, v)}(f(v) - f(u)). \quad (5)$$

The directional derivative of f , at vertex $v \in V$, along the edge $e = (u, v)$, is defined as:

$$\partial_v f(u) = (df)(u, v). \quad (6)$$

The adjoint to the difference operator $d^* : H(E) \rightarrow H(V)$, is a linear operator defined by:

$$\langle df, H \rangle_{H(E)} = \langle f, d^* H \rangle_{H(V)}, \quad (7)$$

for any function $H \in H(E)$ and function $f \in H(V)$.

Proposition 1: The adjoint operator d^* can be expressed at a vertex $u \in V$ by the following formula:

$$(d^* H)(u) = \sum_{v \sim u} \sqrt{w(u, v)}(H(v, u) - H(u, v)). \quad (8)$$

The divergence operator, defined by $-d^*$, measures the network outflow of a function in $H(E)$, at each vertex of the graph.

The weighted gradient operator of $f \in H(V)$, at vertex $u \in V$, $\forall (u, v_i) \in E$, is a column vector:

$$\nabla_w f(u) = (\partial_v f(u) : v \sim u)^T = (\partial_{v_1} f(u), \dots, \partial_{v_k} f(u))^T.$$

The weighted Laplace operator $\Delta_w : H(V) \rightarrow H(V)$, is defined by:

$$\Delta_w f = -\frac{1}{2} d^* (df). \quad (9)$$

Proposition 2 [1]: The weighted Laplace operator Δ_w at $f \in H(V)$ acts as follows:

$$(\Delta_w f)(u) = -\sum_{v \sim u} w(u, v)(f(v) - f(u)).$$

This Laplace operator is linear and corresponds to the graph Laplacian.

Proposition 3 [7]: Let $G_h = (V, E, A_h)$ be a fuzzy weighted graph, corresponding to the h -uniform fuzzy

partition of $V = \{1, \dots, \ell\}$. Then, the weighted Laplace operator Δ_h at $f \in H(V)$ acts as follows:

$$(\Delta_h f)(i) = - \sum_{i \sim j} A_i^h(j)(f(j) - f(i)) = f(i) - F^h[f]_i,$$

where $F^h[f]_i$, $i = 1, \dots, \ell$, is the i -th discrete F-transform component of f , cf. [6].

5 Multi-scale Representation in a Space with a Fuzzy Partition

Taking into account the introduced notation, we propose the following scheme for the multi-scale representation L_{FP} of the signal $f : V \rightarrow \mathbb{R}$, where $V = \{1, \dots, \ell\}$ and subscript ‘‘FP’’ stands for an h -uniform fuzzy partition determined by parameter $h \in \mathbb{N}$, $h \geq 1$:

$$\begin{aligned} L_{FP}(\cdot, 0) &= f(\cdot), \\ L_{FP}(\cdot, t) &= F^{2^t h}[f], \end{aligned} \quad (10)$$

where $t \in \mathbb{N}$ is the scale parameter and $F^{2^t h}[f]$ is the whole vector of F-transform components of f . The scale parameter t relates to the length of the support of the corresponding basic function. As in the case of (1), it is a natural measure of spatial scale at level t . To show the relationship to the diffusion equation, we formulate the following general result.

Proposition 4: Assume that two time continuously differentiable real function $f : [a, b] \rightarrow \mathbb{R}$, and $[a, b]$ is h - and $2h$ -uniform fuzzy partitioned by A_1^h, \dots, A_n^h and $A_1^{2h}, \dots, A_n^{2h}$, where basic functions $A_i^h (A_i^{2h})$, $i = 1, \dots, n$, are generated by $A_0(x) = 1 - |x|$ with the node at $x_i = a + \frac{b-a}{n-1}(i-1)$. Then,

$$F^{2h}[f]_i - F^h[f]_i \approx \frac{h^2}{4} f''(x_i).$$

The semantic meaning of this proposition in relation to the proposed scheme (10) of multi-scale representation L_{FP} of f is as follows: the Laplacian of f can be approximated by the (weighted) differences of two adjacent convolutions determined by the triangular-shaped generating function of a fuzzy partition.

6 Experiments with Time Series

Among other purposes (feature extraction, regularization, etc.), the multi-scale representation is used for reconstruction to ensure that all steps are performed correctly. To demonstrate the effectiveness of the proposed representation, we used several time series with high volatility. Then, with each value of $t = 1, 2, \dots$ we obtain the corresponding Laplacian as the difference

between two adjacent convolutions or vectors with F-transform components, so that we obtain the sequence

$$\{L_{FP}(\cdot, t+1) - L_{FP}(\cdot, t) \mid t = 1, 2, \dots\}$$

The stop criterion is closeness to zero of the current difference. We then compute the reconstruction by summing all the elements in the sequence. Figure 1 shows the step-by-step reconstruction and the final reconstructed time series. The latter is plotted on the bottom image along with the original time series to give confidence in a perfect fit.



Figure 1: *Top.* The sequence of reconstruction steps, where with each value $t = 1, 2, \dots$ we improve the quality of the reconstruction by adding the corresponding Laplacian to the previous one. *Bottom.* The original time series is rendered with a blue line.

For comparison, in Figure 2, we show two MLP reconstructions of the same time series with the following configurations: 4 hidden layers with 4086 neurons in each layer (common setting) and learning rates 0.001 and 0.0001 (different setting). It is obvious that the proposed multi-scale representation and subsequent reconstruction are computationally cheaper and give results with better reconstruction quality. To confirm, we give three estimates of the Euclidean distances between the original time series and its reconstructions: 89.644 (multi-scale in the space with a fuzzy partition), 159.2996 (MLP, first configuration), 127.4569 (MLP, second configuration).

Conclusion

We focused on the multi-scale representation of a 1D signal in the space with a fuzzy partition. We have shown that the predominant statement about the uniqueness of the Gaussian kernel in the scale space representation can be weakened by using kernels that arise from the generating functions of fuzzy partitions. We also showed that the reconstruction from the proposed multiscale representation is of better quality than

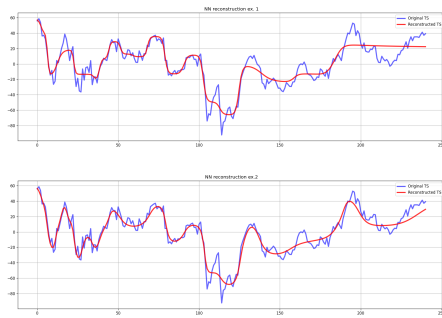


Figure 2: Two MLP reconstructions of the same time series as in Figure 1. The original time series is rendered on each image with a blue thin line.

the reconstruction from MLP with almost twice as many neurons in 4 hidden layers.

Acknowledgement

The support of the grant project SGS18/PrF-MF/2021 (Ostrava University) is kindly announced.

References

- [1] A. Elmoataz, O. Lezoray, S. Bougleux, Discrete regularization on weighted graphs: A framework for image and manifold processing, *IEEE Transactions on Image Processing*, 17 (2008) 1047 - 1060.
- [2] P. J. Burt, Fast filter transform for image processing, *Computer Graphics and Image Processing*, 16 (1981) 20 - 51.
- [3] A. Klinger, (1971) Pattern and search statistics, *Proc. Symposium Held at the Center for Tomorrow, the Ohio State University, June 14–16, 1971*, pp. 303 - 337.
- [4] J. J. Koenderink, The structure of images, *Biological Cybernetics*, 50 (1984) 363 - 370.
- [5] T. Lindeberg, Scale-space theory: a basic tool for analyzing structures at different scales, *Journal of Applied Statistics*, 21 (1994) 225 - 270.
- [6] I. Perfilieva, Fuzzy transforms: Theory and applications, *Fuzzy sets and systems* 157 (8) (2006) 993 - 1023.
- [7] I. Perfilieva, P. Vlasanek, Total variation with nonlocal FT-Laplacian for patch-based inpainting, *Soft Computing* 23 (2019) 1833 - 1841.
- [8] A. P. Witkin, Scale-space filtering, *Proc. 8th International Joint Conference on Artificial Intelligence, IJCAI'83, v.2, 1983*, pp. 1019 - 1022.