Manifold and Data Filtering on Graphs

Olivier Lezoray, Vinh Thong Ta, Abderrahim Elmoataz

Manifold and Data Filtering on Graphs

Olivier Lézoray, Vinh-Thong Ta, Abderrahim Elmoataz

Université de Caen Basse-Normandie, GREYC UMR CNRS 6072, ENSICAEN,
Équipe Image, 6 Bd. Maréchal Juin, F-14050 Caen, France

Abstract. High-dimensional feature spaces are often corrupted by noise. This is problematic for the processing of manifold and data since most of reference methods are sensitive to noise. This paper presents pre-processing methods for manifold denoising and simplification based on discrete analogues of continuous regularization and mathematical morphology. Both approaches enable to project the data onto a submanifold with a graph generated by the data.

1 Introduction

In manifold and data processing, dimensionality reduction [1], clustering [2] and classification [3] are key processes. In particular, methods based on differences on graphs (e.g. the graph Laplacian) have became increasingly popular in machine learning to perform any of the above-mentioned key processes. All these methods are based on the assumption that the data lies on a submanifold. However, sampled data lies almost never exactly on the submanifold due to the noise scattered around it. Graph based methods being sensitive to noise, it is essential to denoise the manifold data [4]. This enables to project the initial noisy data, that are scattered around the true manifold, onto a submanifold. In this paper, we propose a family of weighted difference operators that lead to the formulation of discrete analogues of continuous regularization and mathematical morphology on graphs. These proposals can be used within the KDD process as a pre-processing method that can ease the next steps (classification, visualization, ...).

2 Operators on weighted graphs

In this Section, we recall some basic definitions on graphs, and define difference, divergence, gradients and p-Laplacian operators.

2.1 Preliminary notations and definitions

A *weighted graph* $G = (V, E, w)$ consists in a finite set $V = \{v_1, \dots, v_N\}$ of N vertices and a finite set $E \subset V \times V$ of weighted edges. We assume G to be undirected, with no self-loops and no multiple edges. Let (u, v) be the edge of E that connects vertices u and v of V. Its weight, denoted by $w(u, v)$, represents the similarity between its vertices. Similarities are usually computed by using a positive symmetric function $w : V \times V \to \mathbb{R}^+$ satisfying $w(u,v) = 0$ if $(u,v) \notin E$. The notation $u \sim v$ is also used to denote two adjacent vertices. We say that G is connected whenever, for any pair of vertices (u,v) there is a finite sequence $u = u_0, u_1, \ldots, u_n = v$ such that u_{i-1} is a neighbor of u_i for every $i \in \{1, \ldots, n\}$. Let $\mathcal{H}(V)$ be the Hilbert space of real-valued functions defined on the vertices of a graph $G = (V, E, w)$. A function $f : V \to \mathbb{R}$ of $\mathcal{H}(V)$ assigns a real value $f(v)$ to each vertex $v \in V$. Clearly, the function f can be represented by a column vector $f = [f(v_1), \ldots, f(v_N)]^T$. By analogy with functional analysis on continuous spaces, the integral of a function $f \in \mathcal{H}(V)$, over the set of vertices V, is defined as $\int_V f = \sum_V f$. The space $\mathcal{H}(V)$ is endowed with the usual inner product $\langle f, h \rangle_{\mathcal{H}(V)} = \sum_{v \in V} f(v)h(v)$, where $f, h : V \to \mathbb{R}$. Similarly, let $\mathcal{H}(E)$ be the space of real-valued functions defined on the edges of G. It is endowed with the inner product $\langle F, H \rangle_{\mathcal{H}(E)} = \sum_{(u,v) \in E} F(u,v)H(u,v) = \sum_{u \in V} \sum_{v \sim u} F(u,v)H(u,v)$, where $F, H : E \to \mathbb{R}$ are two functions of $\mathcal{H}(E)$. Let $\mathcal{A}$ be a set of connected vertices with $\mathcal{A} \subset V$ such that for all $u \in \mathcal{A}$, there exists a vertex $v \in \mathcal{A}$ with $(u,v) \in E$. We denote by $\partial^+\mathcal{A}$ and $\partial^-\mathcal{A}$, the *external* and *internal* boundary sets of $\mathcal{A}$, respectively. For a given vertex $u \in V$, $\partial^+\mathcal{A} = \{u \in \mathcal{A}^c : \exists v \in \mathcal{A} \text{ with } (u,v) \in E\}$ and $\partial^-\mathcal{A} = \{u \in \mathcal{A} : \exists v \in \mathcal{A}^c \text{ with } (u,v) \in E\}$, where $\mathcal{A}^c = V \setminus \mathcal{A}$ is the complement of $\mathcal{A}$.

2.2 Difference operators

Let $G = (V, E, w)$ be a weighted graph, and let $f : V \to \mathbb{R}$ be a function of $\mathcal{H}(V)$. The *difference operator* [5] of f, noted $d : \mathcal{H}(V) \to \mathcal{H}(E)$, is defined on an edge $(u,v) \in E$ by: $(df)(u,v) = w(u,v)^{1/2}(f(v) - f(u))$. The *directional derivative* (or *edge derivative*) of f, at a vertex $v \in V$, along an edge $e = (u,v)$, is defined as: $\left.\frac{\partial f}{\partial e}\right|_u = \partial_v f(u) = (df)(u,v)$. This definition is consistent with the continuous definition of the derivative of a function: $\partial_v f(u) = -\partial_u f(v)$, $\partial_v f(v) = 0$, and if $f(v) = f(u)$ then $\partial_v f(u) = 0$. The *adjoint* of the difference operator, noted $d^* : \mathcal{H}(E) \to \mathcal{H}(V)$, is a linear operator defined by $\langle df, H \rangle_{\mathcal{H}(E)} = \langle f, d^*H \rangle_{\mathcal{H}(V)}$, for all $f \in \mathcal{H}(V)$ and all $H \in \mathcal{H}(E)$. The adjoint operator d^*, of a function $H \in \mathcal{H}(E)$, can by expressed at a vertex $u \in V$ by the following expression: $(d^*H)(u) = \sum_{v \sim u} w(u,v)^{1/2}(H(v,u) - H(u,v))$. The *divergence operator*, defined by $-d^*$, measures the net outflow of a function of $\mathcal{H}(E)$ at each vertex of the graph. Each function $H \in \mathcal{H}(E)$ has a null divergence over the entire set of vertices: $\sum_{u \in V}(d^*H)(v) = 0$. Other general definitions of the difference operator have been proposed (see e.g. in [6]). However, the latter operator is not null when the function f is locally constant and its adjoint is not null over the entire set of vertices [7]. Based on the difference operator, we define two new *weighted morphological directional difference operators* [8]. The weighted morphological *external* and *internal* difference operator are respectively: $(d_w^+ f)(u,v) = w(u,v)^{1/2}\big(\max\big(f(u), f(v)\big) - f(u)\big) = \max\big(0, (d_w f)(u,v)\big)$ and $(d_w^- f)(u,v) = w(u,v)^{1/2}\big(f(u) - \min\big(f(u), f(v)\big)\big) = -\min\big(0, (d_w f)(u,v)\big)$ with $(d_w^- f)(u,v) = (d_w^+ f)(v,u)$. The corresponding external and internal partial derivatives are $\partial_v^+ f(u) = (d_w^+ f)(u,v)$ and $\partial_v^- f(u) = (d_w^- f)(u,v)$.

2.3 Gradients and norms

The *weighted gradient operator* of a function $f \in \mathcal{H}(V)$, at a vertex $u \in V$, is the vector operator defined by $(\nabla_w f)(u) = [\partial_v f(u) : v \sim u]^T = [\partial_{v_1} f(u), \dots, \partial_{v_k} f(u)]^T$, $\forall (u, v_i) \in E$. The $\mathcal{L}_p$-norm of this vector represents the *local variation* of the function f at a vertex of the graph. It is defined by:

$$\|(\nabla_w f)(u)\|_p = \Big[\sum_{v\sim u} w(u,v)^{p/2} \big|f(v) - f(u)\big|^p\Big]^{1/p} \ . \tag{1}$$

The local variation is a semi-norm which measures the regularity of a function around a vertex of the graph. Following this definition, we introduce two *new weighted morphological (internal and external) gradients* based on the external and internal partial derivatives such as $(\nabla^{\pm}_w f)(u) = \big(\partial^{\pm}_v f(u)\big)_{(u,v)\in E}$. The external gradient of a function is a directional difference operator defined as the difference between an extensive operator and the function (a typical one being the max), and similarly for the internal gradient [9]. In the sequel, we use the $\mathcal{L}_p$-norm of the weighted discrete morphological gradients for a given vertex $u \in V$, we have, with $M^+ = \max$ and $M^- = \min$,

$$\|(\nabla^{\pm}_w f)(u)\|_p = \Big[\sum_{v\sim u} w(u,v)^{p/2} \big|M^{\pm}\big(0, f(v) - f(u)\big)\big|^p\Big]^{1/p} \ . \tag{2}$$

For the $\mathcal{L}_\infty$-norm, we have

$$\|(\nabla^{\pm}_w f)(u)\|_\infty = \max_{v\sim u}\Big(w(u,v)^{1/2} \big|M^{\pm}\big(0, f(v) - f(u)\big)\big|\Big) \ . \tag{3}$$

2.4 *p*-Laplace operator

Let $p \in (0, +\infty)$ be a real number. The *weighted p-Laplace operator* of a function $f \in \mathcal{H}(V)$, noted $\Delta^p_w : \mathcal{H}(V) \to \mathcal{H}(V)$, is defined by:

$$\Delta^p_w f(u) = \tfrac{1}{2} d^*(\|(\nabla_w f)(u)\|_2^{p-2} (df)(u,v)). \tag{4}$$

The p-Laplace operator of $f \in \mathcal{H}(V)$, at a vertex $u \in V$, can be computed by:

$$\Delta^p_w f(u) = \tfrac{1}{2} \sum_{v\sim u} \gamma^f_w(u,v)(f(u) - f(v)), \tag{5}$$

with

$$\gamma^f_w(u,v) = w(u,v)(\|(\nabla_w f)(v)\|_2^{p-2} + \|(\nabla_w f)(u)\|_2^{p-2}). \tag{6}$$

The p-Laplace operator is nonlinear, with the exception of $p = 2$. In this latter case, it corresponds to the *combinatorial graph Laplacian*, which is one of the classical second order operators defined in the context of spectral graph theory. Another particular case of the p-Laplace operator is obtained with $p = 1$. In this case, it is the *weighted curvature* of the function f on the graph. Finally, the p-Laplacian matrix L^p satisfies the following properties. (1) its expression is:

$$L^p(v_i, v_j) = \begin{cases} \frac{1}{2} \sum\limits_{v_k \sim v_i} \gamma_w^f(v_i, v_k) \text{ if } v_i = v_j \\ -\frac{1}{2}\gamma_w^f(v_i, v_j) \text{ if } v_i \neq v_j \text{ and } (v_i, v_j) \in E \\ 0 \text{ if } (v_i, v_j) \notin E \end{cases}$$

(2) For every vector $f(u)$, $f^T(u)L^p f(u) = \Delta_w^p f(u)$. (3) L^p is symmetric and positive semi-definite. (4) L^p has non-negative, real-valued eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$. The p-Laplacian matrix can be used for manifold learning and generalizes approaches based on the 2-Laplacian [1]. Therefore, a dimensionality reduction can be obtained with an eigen-decomposition of the matrix L^p.

3 p-Laplacian regularization on graphs

Let $f^0 : V \to \mathbb{R}$ be a given function defined on the vertices of a weighted graph $G = (V, E, w)$. In a given context, the function f^0 represents an observation of a clean function $g : V \to \mathbb{R}$ corrupted by a given noise n such that $f^0 = g+n$. Such noise is assumed to have zero mean and variance σ^2, which usually corresponds to observation errors. To recover the uncorrupted function g, a commonly used method is to seek for a function $f : V \to \mathbb{R}$ which is regular enough on G, and also close enough to f^0. This inverse problem can be formalized by the minimization of an energy functional, which typically involves a regularization term plus an approximation term (also called fitting term). In this paper, we consider the following variational problem [5]:

$$g \approx \min_{f:V\to\mathbb{R}} \left\{ E_w^p(f, f^0, \lambda) = R_w^p(f) + \frac{\lambda}{2}\|f - f^0\|_2^2 \right\}, \tag{7}$$

where the regularization functional $R_w^p : \mathcal{H}(V) \to \mathbb{R}$ is the discrete p-Dirichlet form of the function $f \in \mathcal{H}(V)$: $R_w^p(f) = \frac{1}{p}\sum_{u\in V} \|(\nabla_w f)(u)\|_2^p$. The trade-off between the two competing terms in the functional E_w^p, is specified by the fidelity parameter $\lambda \geq 0$. By varying the value of λ, the variational problem (7) allows one to describe the function f^0 at different scales, each scale corresponding to a value of λ. The degree of regularity, which has to be preserved, is controlled by the value of $p > 0$.

3.1 Diffusion processes

When $p \geq 1$, the energy E_w^p is a convex functional of functions of $\mathcal{H}(V)$. To get the solution of the minimizer (7), we consider the following system of equations $\frac{\partial E_w^p(f,f^0,\lambda)}{\partial f(u)} = 0, \forall u \in V$ which is rewritten as: $\frac{\partial R_w^p(f)}{\partial f(u)} + \lambda(f(u) - f^0(u)) = 0, \quad \forall u \in V$. The solution of the latter system is computed by using the following property. Let f be a function in $\mathcal{H}(V)$, one can easily prove that $\frac{\partial R_w^p(f)}{\partial f(u)} = 2\Delta_w^p f(u)$, and the system is rewritten as $2\Delta_w^p f(u) + \lambda(f(u) -$

$f^0(u)) = 0, \quad \forall u \in V$, which is equivalent to the following system of equations: $\left(\lambda + \sum_{v\sim u} \gamma_w^f(u,v)\right) f(u) - \sum_{v\sim u} \gamma_w^f(u,v) f(v) = \lambda f^0(u)$. We use the linearized Gauss-Jacobi iterative method to solve the previous system. Let t be an iteration step, and let $f^{(t)}$ be the solution at the step t. Then, the method is given by the following algorithm:

$$\begin{cases} f^{(0)} = f^0 \\ f^{(t+1)}(u) = \dfrac{\lambda f^0(u) + \sum_{v\sim u} \gamma_w^{f^{(t)}}(u,v) f^{(t)}(v)}{\lambda + \sum_{v\sim u} \gamma_w^{f^{(t)}}(u,v)}, \ \forall u \in V. \end{cases} \tag{8}$$

It describes a family of discrete diffusion processes [5], which is parameterized by the structure of the graph (topology and weight function), the parameter p, and the parameter λ. At each iteration of the algorithm (8), the new value $f^{(t+1)}(u)$ depends on two quantities: the original value $f^0(u)$, and a weighted average of the filtered values of $f^{(t)}$ in a neighborhood of u. When $\lambda = 0$ and $p \geq 0$, an iteration of the diffusion process given by $f^{(t+1)}(u) = \sum_{v\sim u} \phi_{uv}\left(f^{(t)}\right) f^{(t)}(v)$, $\forall u \in V$ with $\phi_{uv}(f) = \frac{\gamma_w^f(u,v) f(v)}{\sum_{v\sim u} \gamma_w^f(u,v)}$, $\forall (u,v) \in E$. Let Q be the Markov matrix defined by $Q(u,v) = \phi(u,v)$ if $(u,v) \in E$ and $Q(u,v) = 0$ otherwise. Then, an iteration of the diffusion process (8) with $\lambda = 0$ can be written in matrix form as $f^{(t+1)} = Qf^{(t)} = Q^{(t)} f^0$, where Q is a stochastic matrix (nonnegative, symmetric, unit row sum). An equivalent way to look at the power of Q in the diffusion process is to decompose each value of f on the first eigenvectors of Q, therefore the diffusion process can be interpreted as a filtering in the spectral domain [10].

4 Mathematical morphology on graphs

Mathematical morphology (MM) offers a wide range of operators to address various image processing problems. These operators can be defined in terms of algebraic (discrete) sets or as partial differential equations (PDEs). Morphological algebraic flat dilation δ and erosion ε of a function $f^0{:}\mathbb{R}^n{\rightarrow}\mathbb{R}$ are usually formulated by: $\delta_B(f^0)(x) = \sup\{f^0(x{+}y) : y{\in}B\}$ and $\varepsilon_B(f^0)(x) = \inf\{f^0(x{+}y) : y{\in}B\}$ with B a compact convex symmetric set (called structuring element). By using structuring sets $B{=}\{x : \|x\|_p{\leq}1\}$, the general PDEs generating these flat dilations and erosions are as follows [11]:

$$\frac{\partial\delta(f)}{\partial t}{=}\partial_t f{=}{+}\|\nabla f\|_p \text{ and } \frac{\partial\varepsilon(f)}{\partial t}{=}\partial_t f{=}{-}\|\nabla f\|_p \tag{9}$$

where f is a modified version of f^0, ∇ is the gradient operator, $\|.\|_p$ corresponds to the $\mathcal{L}_p$-norm, and one has the initial condition $\partial_{t=0} f{=}f^0$. With different values of p, one obtains different structuring elements: a rhombus for $p{=}\infty$, a disc with $p{=}2$ and a square with $p{=}1$. Whatever the chosen formulation (algebraic or PDEs), if MM is well defined for binary and gray scale images, there exists no general extension for the processing of multivariate unorganized data sets. We propose to consider a discrete version of continuous MM over weighted graphs that naturally enables processing of multivariate data living on any domains [8].

4.1 Morphological gradients and algebraic operators

The previously defined external and internal gradients operate on any graph structure. In the sequel, we show that in the particular case of a unweighted (w=g_0) graph and with p=∞, our gradient formulations recover algebraic morphological operators where the structuring element is provided by the graph topology. The $\mathcal{L}_\infty$-norms (3) of the proposed external and internal gradients ∇_w^+ and ∇_w^- recover the classical definition of algebraic morphological external ($\delta(f)(u)-f(u)$) and internal ($f(u)-\varepsilon(f)(u)$) gradients. Indeed, for the external gradient $\|(\nabla_w^+ f)(u)\|_\infty = \max\limits_{v\sim u}\big(\max\big(0, f(v)-f(u)\big)\big) = \delta(f)(u)-f(u)$, and similarly for the internal one $\|(\nabla_w^- f)(u)\|_\infty = \max\limits_{v\sim u}\big(|\min\big(0, f(v)-f(u)\big)|\big) = f(u)-\epsilon(f)(u)$. One can also prove the relation between graph boundary sets and the weighted discrete morphological gradient norms (2) of the level set function of f at vertex $u\in V$ [8]. Moreover, let the decomposition of f into its level sets be denoted by f^l=$\chi(f-l)$ where χ is the Heaviside function (a step function). Then, one can prove the following relation [8], with $f^l = \chi(\mathcal{A}^l)$:

$$\|(\nabla_w f^l)(u)\|_p = \begin{cases} \|(\nabla_w^+ f^l)(u)\|_p \text{ if } u \in \partial^+\mathcal{A}^l, \\ \|(\nabla_w^- f^l)(u)\|_p \text{ if } u \in \partial^-\mathcal{A}^l. \end{cases} \tag{10}$$

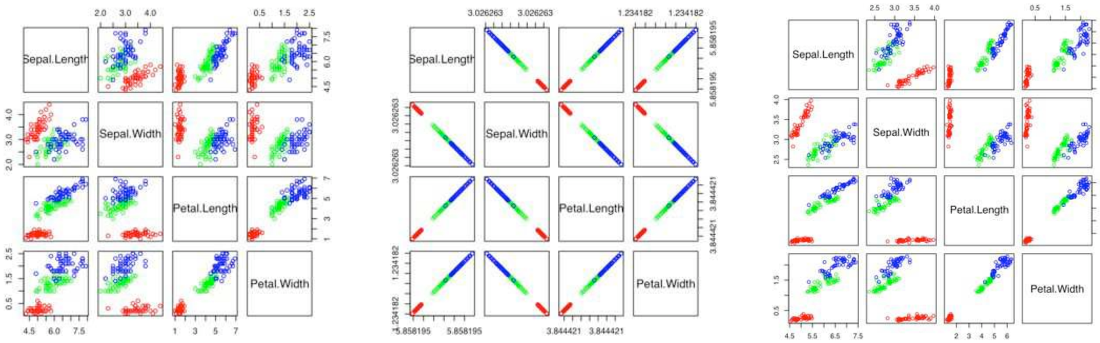
4.2 Dilation and erosion processes

Intuitively from definition (10), dilation over $\mathcal{A}$ can be interpreted as a growth process that adds vertices from $\partial^+\mathcal{A}$ to $\mathcal{A}$. By duality, erosion over $\mathcal{A}$ can be interpreted as a contraction process that removes vertices from $\partial^-\mathcal{A}$. We define the discrete analogue of PDEs-based dilation and erosion formulations and obtain the following expressions over graphs. For a given initial function $f^0\in\mathcal{H}(V)$: $\frac{\partial\delta(f)(u)}{\partial t}$=$\partial_t f(u)$=$+\|(\nabla_w^+ f)(u)\|_p$ and $\frac{\partial\varepsilon(f)(u)}{\partial t}$=$\partial_t f(u)$=$-\|(\nabla_w^- f)(u)\|_p$ $\quad\forall u\in V$, with the initial condition $\partial_{t=0} f$=f^0 (f is a modified version of f^0) and ∇_w^+ and ∇_w^- are the weighted discrete morphological gradients. To solve these dilation and erosion processes, on the contrary to the PDEs case, no spatial discretization is needed thanks to derivatives that are directly expressed in a discrete form. Then, by using discretization in time, and with the usual notation $f(u,n)\approx f(u,n\Delta t)$, the general iterative scheme for dilation and erosion, can be defined at time n+1, for all $u \in V$, as [8] $f^{n+1}(u)$=$f^n(u)$+$\Delta t\|(\nabla_w^+ f^n)(u)\|_p$ and $f^{n+1}(u)$=$f^n(u)$−$\Delta t\|(\nabla_w^- f^n)(u)\|_p$. The initial condition is $f^{(0)}$=f^0 where $f^0\in\mathcal{H}(V)$ is the initial function defined on the graph vertices. If dilation is considered and with the corresponding gradient norms, the iterative scheme becomes for $0<p<+\infty$, $f^{n+1}(u) \overset{(2)}{=} f^n(u)+\Delta t\Big(\sum\limits_{v\sim u} w(u,v)^{p/2}\big|\max\big(0,(f^n(v)-f^n(u))\big)\big|^p\Big)^{\frac{1}{p}}$, and for p=∞, $f^{n+1}(u) \overset{(3)}{=} f^n(u)+\Delta t\max\limits_{v\sim u}\Big(w(u,v)^{1/2}\big|\max\big(0,(f^n(v)-f^n(u))\big)\big|\Big)$.

At each step of the algorithms, the new value at vertex u only depends on its value at step n and the existing values in its neighborhood. The proposed dilation expression also recovers the classical algebraic flat morphological dilation formulation over graphs. Indeed, in the case where $p{=}\infty$ with a constant discretization time $\Delta t{=}1$ and a constant weight function g_0, the iterative scheme becomes for $u{\in}V$: $f^{n+1}(u){=}\max\limits_{v\sim u}\big(f^n(v), f^n(u)\big)$. In that case, the structuring element is provided by the graph topology.

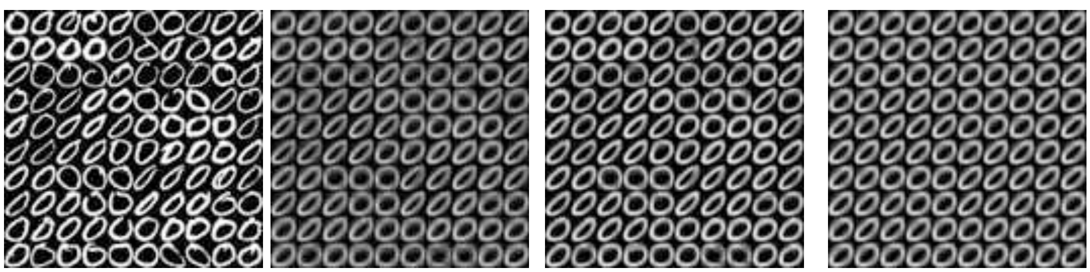
5 Results

In this Section, we illustrate the abilities of regularization and mathematical morphology processes on graphs for denoising and simplification of any function defined on a finite set $V = \{v_1, \cdots, v_N\}$ of discrete data $v_i \in \mathbb{R}^m$. This is achieved by constructing a weighted graph $G = (V, E, w)$ and by considering the function to be simplified as a function $f^0 : V \to \mathbb{R}^m$, defined on the vertices of G. The simplification processes operate on vector-valued function with one process per vector component. If not specified, all graphs are weighted with a Gaussian kernel. Moreover, an ASF (Alternate Sequential Filter) is an iterative morphological filter that performs openings ($\delta\epsilon$) and closings ($\epsilon\delta$)of increasing sizes. Figure 1(a) presents a pairwise feature projection of the original Iris database ($f^0 : V \to \mathbb{R}^4$). With such real-world databases, some noise is present and data smoothing is therefore of interest. A complete graph is considered. One can see in Figure 1 the benefits of both regularization (Figure 1(b)) and morphological (Figure 1(c)) processing: input points that belong to the same class tend to be closer than in the original database: the submanifold where the data lies has been recovered. This effect is illustrated by the results obtained with a standard k-means classification on the original and the simplified versions: the denoising pre-processing enables to increase the recognition rate. Typical manifolds being image libraries, we also consider the USPS handwritten digit database for illustration. Each digit is a 16×16 image which is considered as a vector of 256 dimensions. Let $f^0 : V \to \mathbb{R}^{16\times16}$ be a mapping from the vertices of G to the elements of the manifold. We consider a 10-NN graph constructed over the manifold. Figure 2 presents results on a subset of digits from the USPS database. For denoising (second row of Fig. 2), a Gaussian kernel is considered and for simplification (first row of Fig. 2), $w(u, v) = 1 - \frac{\|f(u)-f(v)\|_2}{d_{max}}$ where d_{max} is the maximum distance between two images. First row shows simplification results on an original set of digits (Fig. 2(a)) with Morphological (Fig. 2(b)) and regularization (Fig. 2(c)-(d)) processes. Second row shows regularization results for denoising : an original set of digits (Fig. 2(e)) is corrupted with Gaussian noise (Fig. 2(f) with $\sigma = 40$) and denoised with regularization. Again, such manifold denoising and simplification processes can be useful for classification purposes on a noiseless submanifold extracted from a noisy manifold. Finally, we show results of manifold denoising in conjunction with manifold learning. The original data is a toroidal helix corrupted with Gaussian noise ($\sigma = 10$). Figure 3 presents a regularization result ($p = 1$, $\lambda = 0$, $t = 500$) and a morphological ASF processing

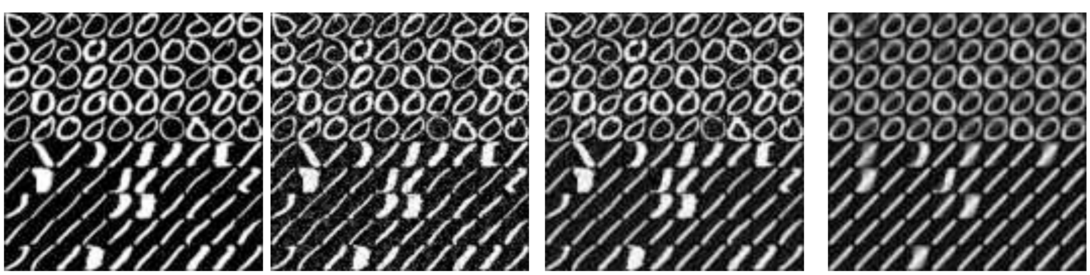


(a) Original Iris database. (b) Regularized Iris database. (c) ASF Iris database.

Fig. 1. Data base regularization ($p = 1$, $\lambda = 0.01$, $t = 10$) and morphological processing with an Alternate Sequential Filter ($p = 1$, $\Delta t = 0.005$, $t = 5$). Classification accuracies with k-means: 89.3% (a), 93.3% (b), and 90.% (c).



(a) Original set (b) Opening ($\delta\epsilon$) (c) $p = 1$, $\lambda = 0.01$ (d) $p = 2$, $\lambda = 0.01$



(e) Original set (f) Corrupted set (g) $p = 1$, $\lambda = 0.01$ (h) $p = 1$, $\lambda = 0$

Fig. 2. Manifold simplification (first row) and denoising (second row) with regularization ($t = 20$, 2(c)-(d) and 2(g)-(h)) and morphological processes (2(b) with $p = 1$, $\Delta t = 0.005$, $t = 5$). See text for details.

result ($p = 1$, $\Delta t = 0.005$, $t = 5$) with a 20-NN graph. For each processing, are provided the simplification result in the original space and a 2D-projection obtained with an eigen-decomposition of the 1-Laplacian from the simplified data. One can easily see the interest of the proposed methods that enable a better manifold learning with the suppression of noise and with compression effects. Once again, the submanifold where the data lies has been recovered whereas this was not possible directly on the noisy data (see the projection obtained directly on the noisy data set in Fig. 3(c)).

6 Conclusion

We proposed a general discrete framework for the filtering of manifolds and data with p-Laplacian regularization and mathematical morphology. The proposed filters can operate on any high dimensional unorganized multivariate data that has been represented with a weighted graph. Both approaches are efficient to denoise manifolds and data to project initial noisy data onto a submanifold to ease dimensionality reduction, clustering and classification.

References

1. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computation **15**(6) (2003) 1373–1396
2. Von Luxburg, U.: A tutorial on spectral clustering. Statistics and Computing **17**(4) (2007) 395–416
3. Belkin, M., Matveeva, I., Niyogi, P.: Regularization and semi-supervised learning on large graphs. In: COLT. (2004) 624–638
4. Hein, M., Maier, M.: Manifold denoising. In: Advances in Neural Information Processing Systems. Volume 20. (2006) 1–8
5. Elmoataz, A., Lezoray, O., Bougleux, S.: Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing. IEEE transactions on Image Processing **17**(7) (2008) 1047–1060
6. Zhou, D., Schölkopf, B.: Regularization on discrete spaces. In: LNCS 3663, Proc. of the 27th DAGM Symp., Springer-Verlag (2005) 361–368
7. Hein, M., Audibert, J.Y., von Luxburg, U.: Convergence of graph Laplacians on random neighborhood graphs. Journal of Machine Learning Research **8** (2007) 1325–1370
8. Ta, V., Elmoataz, A., Lezoray, O.: Partial difference equations over graphs: Morphological processing of arbitrary discrete data. In: European Conference on Computer Vision. (2008) 668–680
9. Rivest, J.F., Soille, P., Beucher, S.: Morphological gradients. Journal of Electronic Imaging **2**(4) (1993)
10. Coifman, R., Lafon, S., Lee, A., Maggioni, M., Nadler, B., Warner, F., Zucker, S.: Geometric diffusions as a tool for harmonic analysis and structure definition of data. Proc. of the National Academy of Sciences **102**(21) (2005)
11. Brockett, R., Maragos, P.: Evolution equations for continuous-scale morphology. In: IEEE International Conference on Acoustics, Speech, and Signal Processing. Volume 3. (1992) 125–128

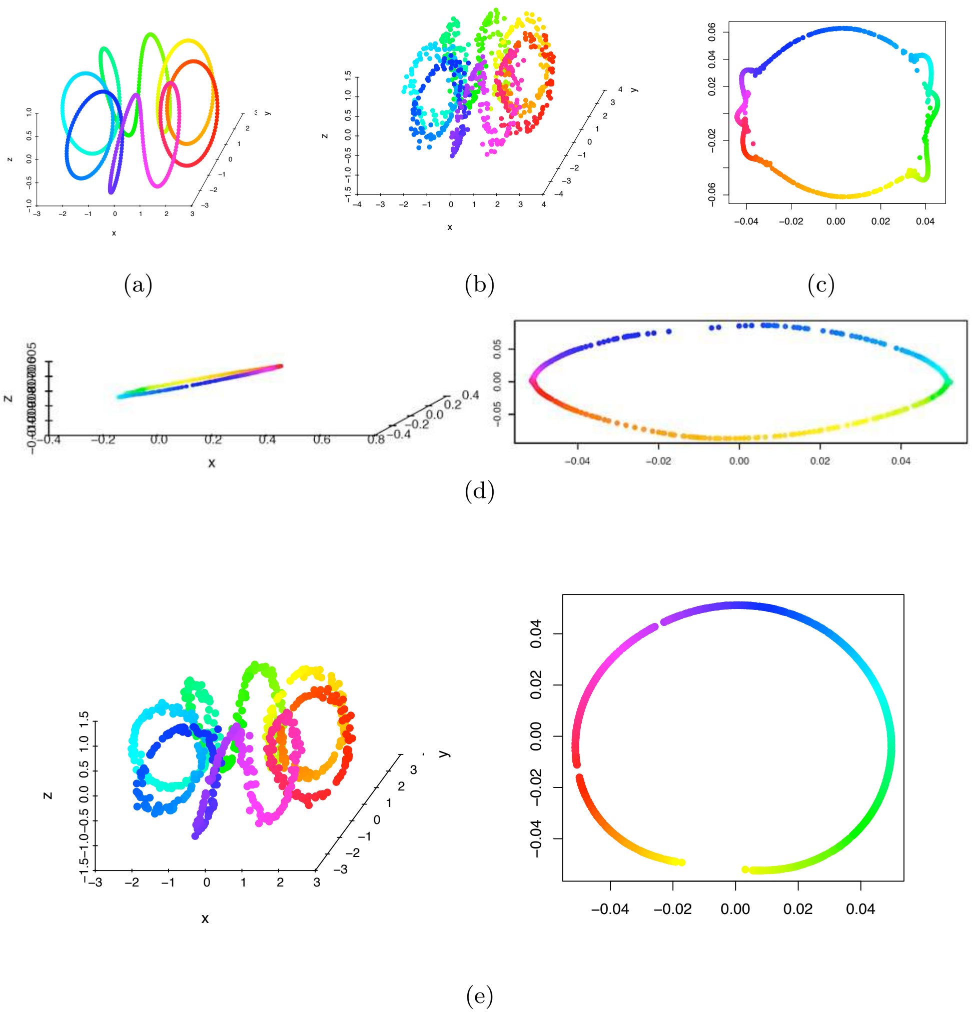


Fig. 3. (a) Original data, (b) Corrupted data with Gaussian noise ($\sigma = 10$), (c) 1-Laplacian dimensionality reduction of (b), (d) regularized and (e) morphological ASF simplification of the data in the original space (left) and their 1-Laplacian dimensionality reduction (right).