# Systematic approach to nonlinear filtering associated with aggregation operators. Part 2. Fréchet MIMO-filters

V. Labunets[a], E. Osthaimer[b]

[a] Ural State Forest Engineering University, 620100, Ekaterinburg, Russia
[b] Capricat LLC, Pompano Beach, Florida, USA

**Abstract**

Median filtering has been widely used in scalar-valued image processing as an edge preserving operation. The basic idea is that the pixel value is replaced by the median of the pixels contained in a window around it. In this work, this idea is extended onto vector-valued images. It is based on the fact that the median is also the value that minimizes the sum of distances between all grey-level pixels in the window. The Fréchet median of a discrete set of vector-valued pixels in a metric space with a metric is the point minimizing the sum of metric distances to the all sample pixels. In this paper, we extend the notion of the Fréchet median to the general Fréchet median, which minimizes the Fréchet cost function (FCF) in the form of aggregation function of metric distances, instead of the ordinary sum. Moreover, we propose use an aggregation distance instead of classical metric distance. We use generalized Fréchet median for constructing new nonlinear Fréchet MIMO-filters for multispectral image processing.

*Keywords:* Nonlinear MIMO-filters; Fréchet point; median hyperspectral image processing; generalized aggregation means

## 1. Introduction

The basic idea behind this paper is the estimation of the uncorrupted image from the distorted or noisy image, and is also referred to as image “denoising”. To denoise images is to filter out the noise. The challenge is to preserve and enhance important features during the denoising process. For images, for example, an edge is one of the most universal and crucial features. There are various methods to help restore an image from noisy distortions [1-3]. Each technique has its advantages and disadvantages. Selecting the appropriate method plays a major role in getting the desired image. Noise removal or noise reduction can be done on an image by linear or nonlinear filtering. The more popular linear technique is based on average (on mean) linear operators. Denoising via linear filters normally does not perform satisfactorily since both noise and edges contain high frequencies. Therefore, any practical denoising model has to be nonlinear. In this paper, we propose a new type of nonlinear data-dependent denoising filter called the *aggregation digital* MIMO-*filter*.

Almost 2500 years ago, the ancient Greeks defined a list of ten (actually eleven) distinct “means” [4-5]. All these means are constructed using geometric proportions. Among these, are the well-known arithmetic, geometric, and harmonic means. These three principal means, which are used particularly in the works of Nicomachus of Gerasa and Pappus, are the only ones that survived in common usage. In fact, for a set of $N$ positive numbers $x^1, x^2, \ldots, x^N \in \mathbf{R}^+$, the arithmetic mean is the positive number $c = \frac{1}{N}\sum_{k=1}^{N} x^k$. The arithmetic mean has a variational property; it minimizes the sum of the squared distances to the given points $x^1, x^2, \ldots, x^N$ : $c_{opt} \equiv \mathbf{arg}\min_{c\in\mathbf{R}}\left(\sum_{i=1}^{N}\rho_1^2\left(c, x^i\right)\right)$, where $\rho_1\left(c, x^i\right) = \left|c - x^i\right|$ represents the usual Euclidean distance in $\mathbf{R}$. The geometric mean which is given by $c = \sqrt[N]{x^1 x^2 \ldots x^N}$ also has a variational property; it minimizes the sum of the squared hyperbolic distances to the given points $x^1, x^2, \ldots, x^N$ : $c_{opt} \equiv \mathbf{arg}\min_{c\in\mathbf{R}}\left(\sum_{i=1}^{N}\rho_h^2\left(c, x^i\right)\right)$, where $\rho_h\left(c, x^i\right) = \left|\log c - \log x^i\right|$ is the hyperbolic distance between $c$ and $x^i$. The harmonic mean is simply given by the inverse of the arithmetic mean of their inverses, i.e., $c = \left(\frac{1}{N}\sum_{k=1}^{N}\left(x^k\right)^{-1}\right)^{-1}$ and thus it has a variational characterization as well.

There is similar situation for vector-valued data. For a given set of $N$ points $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N \in \mathbf{R}^K$, the arithmetic vector-valued mean is given by the barycenter $\mathbf{c} = \frac{1}{N}\sum_{k=1}^{N}\mathbf{x}^k$ of the $N$ points $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$. The arithmetic vector-valued mean has a variational property; it minimizes the sum of the squared distances to the given points $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$ : $\mathbf{c}_{opt} \equiv \mathbf{arg}\min_{c\in\mathbf{R}^K}\left(\sum_{i=1}^{N}\rho_2^2\left(\mathbf{c}, \mathbf{x}^i\right)\right)$, where $\rho_2\left(\mathbf{c}, \mathbf{x}^i\right) = \sqrt{\sum_{k=1}^{K}\left|c_k - x_k^i\right|^2}$ represents the usual Euclidean distance in $\mathbf{R}^K$. The most common distance metrics in continuous space are those known as the class of $\ell_p$ distance metrics:

$$\rho_p\left(\mathbf{c}, \mathbf{x}^i\right) = \sqrt[p]{\sum_{k=1}^{K}\left|c_k - x_k^i\right|^p}. \quad (1)$$

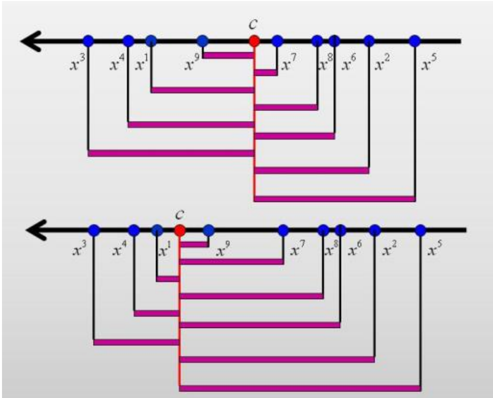


**Fig. 1.** Distances from an arbitrary point $c$ to each point $x^1, x^2, \ldots, x^N \in \mathbf{R}$ from the 9-cellular window.

Note for $p=1$, $\ell_1$ represents the rectilinear, or Manhattan, or city distance metric, for $p=2$, $\ell_2$ is the Euclidean, or straight-line, distance metric, and for $p=\infty$, $\ell_\infty$ is known as the Chebyshev distance metric. The Chebyshev distance in $K$ dimensions can be written as: $\rho_\infty\left(\mathbf{c},\mathbf{x}^i\right)=\max\left(\left|c_1-x_1^i\right|,\left|c_2-x_2^i\right|,\ldots,\left|c_K-x_K^i\right|\right)$. In this paper, we extend the notion of centrality of empirical data, using aggregation distance: $\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^i\right)=\mathbf{Aggreg}\left(\left|c_1-x_1^i\right|,\left|c_2-x_2^i\right|,\ldots,\left|c_K-x_K^i\right|\right)$ instead of (1), where **Aggreg** is an aggregation operator (function) [6-10] and used $\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^i\right)$ for designing of new MIMO-filters. We develop a conceptual framework and design methodologies for multichannel image median filtering systems with assessment capability.

## 2. The object of the study. Optimal Fréchet point, mean and median

The term multichannel (multicomponent, multispectral, hyperspectral) image is used for an image with more than one component. They are composed of a series of images in different optical bands at wavelengths $\lambda_1,\lambda_2,\ldots,\lambda_K$, called the spectral channels: $\vec{\mathbf{f}}(x,y)=\left(f_{\lambda_1}(x,y), f_{\lambda_2}(x,y),\ldots, f_{\lambda_K}(x,y)\right)$, where $K$ is the number of different optical channels, *i.e.*, $\mathbf{f}(x,y):\mathbf{R}^2\to\mathbf{R}^K$, where $\mathbf{R}^K$ is the multicolor space. Let us introduce the observation model and notion used throughout the paper. We consider noise signals or images of the form $\mathbf{f}(\mathbf{x})=\mathbf{s}(\mathbf{x})+\boldsymbol{\eta}(\mathbf{x})$, where $\mathbf{s}(\mathbf{x})=\left(s_1(\mathbf{x}),s_2(\mathbf{x}),\ldots,s_K(\mathbf{x})\right)$ is the original multichannel signal, $\boldsymbol{\eta}(\mathbf{x})=\left(\eta_1(\mathbf{x}),\eta_2(\mathbf{x}),\ldots,\eta_K(\mathbf{x})\right)$ denotes the multichannel noise introduced into the signal $\mathbf{s}(\mathbf{x})$ to produce the corrupted signal $\mathbf{f}(\mathbf{x})=\left(f_1(\mathbf{x}),f_2(\mathbf{x}),\ldots,f_K(\mathbf{x})\right)$. Here $\mathbf{x}=i\in\mathbf{Z}$, $\mathbf{x}=(i,j)\in\mathbf{Z}^2$, or $\mathbf{x}=(i,j,k)\in\mathbf{Z}^3$ are a 1D, 2D, or 3D coordinates, respectively, that belong to the signal (image) domain and represent the pixel location. If $\mathbf{x}\in\mathbf{Z},\mathbf{Z}^2,\mathbf{Z}^3$ then $\mathbf{f}(\mathbf{x}),\mathbf{s}(\mathbf{x})$, $\boldsymbol{\eta}(\mathbf{x})$ are 1D, 2D and 3D multichannel signals, respectively. The aim of image enhancement is to reduce the noise as much as possible or to find a method which, given $\mathbf{s}(\mathbf{x})$, derives an image $\mathbf{y}(\mathbf{x})=\hat{\mathbf{s}}(\mathbf{x})$ as close as possible to the original $\mathbf{s}(\mathbf{x})$, subject to a suitable optimality criterion.

In 2D standard linear and median SISO-filters with a square window $\left[\mathrm{M}_{(i,j)}(m,n)\right]_{m=-r,n=-r}^{m=+r,n=+r}$ of size $(2r+1)\times(2r+1)$ is located at $(i,j)$ the arithmetic mean and median replace the central grey-level (scalar-valued) pixel

$$\hat{s}(i,j)=\underset{(m,n)\in M_{(i,j)}}{\mathbf{Arithm}}\left\{f(m,n)\right\},\quad \hat{s}(i,j)=\underset{(m,n)\in M_{(i,j)}}{\mathbf{Median}}\left\{f(m,n)\right\},\tag{2}$$

where $\hat{s}(i,j)$ is the filtered image, $\left\{f(m,n)\right\}_{(m,n)\in M_{(i,j)}}$ is image block of the fixed size $N=\left|M_{(i,j)}\right|=M\times M=(2r+1)\times(2r+1)$ extracted from $\vec{\mathbf{f}}$ by moving window $\mathrm{M}_{(i,j)}$ at the position $(i,j)$. Symbols **Arithm** and **Median** are the arithmetic mean (average) and median operators, respectively. In the multichannel case (for hyperspectral images), we need to define vector-valued arithmetic mean (average) and median. Median filtering has been widely used in image processing as an edge-preserving filter. The basic idea is that the pixel value is replaced by the median of the pixels contained in the window around it. In this work, this idea is extended to vector-valued images, because the median is also the value that minimizes the $\ell_1$ distance in $\mathbf{R}$ (according to (1)) between all the gray-level pixels in the $N$-cellular window (see Fig. 1). In the multichannel case, we need to define a distance $\rho$ between pairs of objects on the domain $\mathbf{R}^K$.

**Definition 1** [11-12]. The optimal weighted Fréchet median and mean associated with the metric $\rho(\mathbf{x},\mathbf{y})$ are the points $\mathbf{m}_{opt},\mathbf{c}_{opt}\in\mathbf{R}^K$ that minimize the Fréchet cost functions $\mathbf{FCF}_1(\mathbf{c})=\sum_{i=1}^N w_i\rho\left(\mathbf{c},\mathbf{x}^i\right)$ and $\mathbf{FCF}_2(\mathbf{c})=\sum_{i=1}^N w_i\rho^2\left(\mathbf{c},\mathbf{x}^i\right)$ (the weighted sum distances from an arbitrary point $\mathbf{c}$ to each point $\mathbf{x}^1,\mathbf{x}^2,\ldots,\mathbf{x}^N$). They are formally defined as

$$\mathbf{m}_{opt} \equiv \mathbf{FrechMed}\left(\rho \mid \mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N\right) = \mathbf{arg}\min_{\mathbf{m}\in\mathbf{R}^K} \mathbf{FCF}_1(\mathbf{m}) = \mathbf{arg}\min_{\mathbf{m}\in\mathbf{R}^K}\left(\sum_{i=1}^{N} w_i \rho\left(\mathbf{m}, \mathbf{x}^i\right)\right), \quad (3)$$

$$\mathbf{c}_{opt} \equiv \mathbf{FrechMean}\left(\rho \mid \mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N\right) = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K} \mathbf{FCF}_2(\mathbf{c}) = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left(\sum_{i=1}^{N} w_i \rho^2\left(\mathbf{c}, \mathbf{x}^i\right)\right). \quad (4)$$

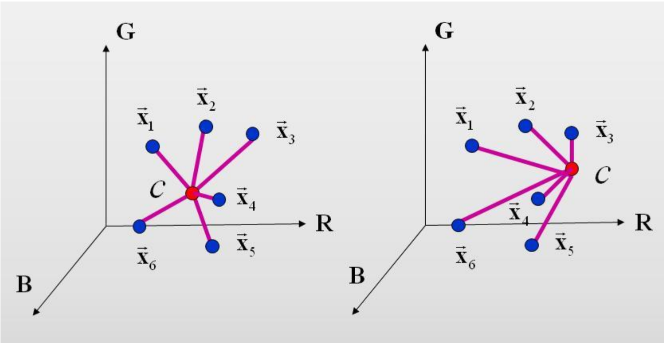



**Fig. 2.** Distances (red lines) from an arbitrary point $\mathbf{c}$ to each point $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N \subset \mathbf{R}^K$ from the 9-cellular window for two probe locations.

When all the weights are equal ( $w_i = 1/N$ ) we call $\mathbf{m}_{opt}, \mathbf{c}_{opt}$ simply the geometric median and mean. Note that $\mathbf{arg}\min$ means the argument, for which the sum is minimized. In this case, it is the point $\mathbf{c}_{opt}$ from $\mathbf{R}^K$, for which the sum of all distances to the $\mathbf{x}^i$ 's is minimum. So, the optimal Fréchet *median and mean* of a discrete set of the observations ( $N$ pixels) in the metric space $\left\langle \mathbf{R}^K, \rho \right\rangle$ are points minimizing the sum of distances and the sum-of-squared distances to the $N$ pixels, respectively (see Fig. 2).

This generalizes the ordinary median, which has the property of minimizing the sum of distances for one-dimensional data. The properties of these points have been extensively studied since the time of Fermat (this points are often called the *Fréchet points* or *Fermat-Weber points* [12]). When filters (3) are modified as follows:

$$\hat{\mathbf{s}}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{FrechMed}}\left[\rho, w(m,n) \mid \mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left(\sum_{(m,n)\in M_{(i,j)}} w(m,n)\rho\left(\mathbf{s}, \mathbf{f}(m,n)\right)\right),$$

$$\hat{\mathbf{s}}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{FrechMean}}\left[\rho^2, w(m,n) \mid \mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left(\sum_{(m,n)\in M_{(i,j)}} w(m,n)\rho^2\left(\mathbf{s}, \mathbf{f}(m,n)\right)\right). \quad (5)$$

it becomes the Fréchet median and mean MIMO-filters (vector-valued filters). Note, that the Fréchet median and mean MIMO-filters are not equivalent to classical vector-median and vector-mean filters (see [13-14]), where, in the first, $\mathbf{s}$ runs among observed $N$ data $\left\{\mathbf{f}(m,n)\right\}_{(m,n)\in M_{(i,j)}} \subset \mathbf{R}^K$ and, in the second, $\rho\left(\mathbf{s}, \mathbf{f}(m,n)\right) \equiv \rho_1\left(\mathbf{s}, \mathbf{f}(m,n)\right) = \left\|\mathbf{s} - \mathbf{f}(m,n)\right\|_1$. In our case $\mathbf{s}$ runs among whole space $\mathbf{R}^K$ and $\rho\left(\mathbf{s}, \mathbf{f}(m,n)\right)$ is an arbitrary distance.

In this paper, we extend the notion of the Fréchet median and mean (3)-(4) to generalized Fréchet point, which minimizes an arbitrary positive convex function on $N+1$ variables - generalized Fréchet cost function (GFCF) - $\mathbf{GFCF}(\rho^{(1)}, \rho^{(2)}, ..., \rho^{(N)})$

$$\begin{aligned}\mathbf{p}_{opt} &\equiv \mathbf{FrechPt}\left(\mathbf{GFCF}, \rho \mid \mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N\right) = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\mathbf{GFCF}(\rho^{(1)}, \rho^{(2)}, ..., \rho^{(N)})\right] \\ &= \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\mathbf{GFCF}\left(\rho(\mathbf{c}, \mathbf{x}^1), \rho(\mathbf{c}, \mathbf{x}^2), ..., \rho(\mathbf{c}, \mathbf{x}^N)\right)\right],\end{aligned} \quad (6)$$

instead of the ordinary sum, where $\rho^1 = \rho(\mathbf{c}, \mathbf{x}^1),\ \rho^2 = \rho(\mathbf{c}, \mathbf{x}^2),\ ...,\ \rho^N = \rho(\mathbf{c}, \mathbf{x}^N)$. In particular, important case we are going to use *aggregation Fréchet cost function* in the form of an aggregation function $\mathbf{GFCF} = \mathbf{Agg}_{\mathbf{CF}}$ :

$$\mathbf{GFCF}(\rho^1, \rho^2, ..., \rho^N) \equiv \mathbf{Agg}_{\mathbf{CF}}\left\{w_1\rho\left(\mathbf{c}, \mathbf{x}^1\right), w_2\rho\left(\mathbf{c}, \mathbf{x}^2\right), ..., w_N\rho\left(\mathbf{c}, \mathbf{x}^N\right)\right\} = \mathbf{Agg}_{\mathbf{CF}}\left\{w_i\rho\left(\mathbf{c}, \mathbf{x}^i\right)\right\}_{i=1}^{N} \quad (7)$$

In this case

$$\mathbf{p}_{opt} \equiv \mathbf{FrechPt}\left(\mathbf{Agg}_{\mathbf{CF}}, \rho \mid \mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N\right) = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\mathbf{Agg}_{\mathbf{CF}}\left\{w_1\rho\left(\mathbf{c}, \mathbf{x}^1\right), w_2\rho\left(\mathbf{c}, \mathbf{x}^2\right), ..., w_N\rho\left(\mathbf{c}, \mathbf{x}^N\right)\right\}\right]. \quad (8)$$

Moreover, we propose to use an aggregation distance $\rho_{\mathbf{Agg}}\left(\mathbf{c}, \mathbf{x}\right)$ instead of the classical distance $\rho$. It gives a new cost function $\mathbf{Agg}_{\mathbf{CF}}\left[w_1\rho_{\mathbf{Agg}}\left(\mathbf{c}, \mathbf{x}^1\right), w_2\rho_{\mathbf{Agg}}\left(\mathbf{c}, \mathbf{x}^2\right), ..., w_N\rho_{\mathbf{Agg}}\left(\mathbf{c}, \mathbf{x}^N\right)\right]$ and new optimal Fréchet point associated with the aggregation distance $\rho_{\mathbf{Agg}}\left(\mathbf{c}, \mathbf{x}\right)$ and $\mathbf{Agg}_{\mathbf{CF}}$

$$\mathbf{p}_{opt} \equiv \mathbf{AgFrechPt}\left(\mathbf{Agg_{CF}}, \rho_{\mathbf{Agg}}; w^1, w^2, ..., w^N \mid \mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N\right) = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\mathbf{Agg_{CF}}\left\{w_1\rho^1_{\mathbf{Agg}}, w_2\rho^2_{\mathbf{Agg}}, ..., w_N\rho^N_{\mathbf{Agg}}\right\}\right] =$$
$$= \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\mathbf{Agg_{CF}}\left\{w_1\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^1\right), w_2\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^2\right), ..., w_N\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^N\right)\right\}\right] = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\mathbf{Agg_{CF}}\left\{w_i\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^i\right)\right\}_{i=1}^N\right]. \quad (9)$$

We use the generalized Fréchet point for constructing new nonlinear filters. When filters (3) are modified as follows:

$$\hat{\mathbf{s}}(i,j) = \mathbf{AgFrechPt}_{(m,n)\in M_{(i,j)}}\left[\mathbf{Agg_{CF}}, \rho_{\mathbf{Agg}}, w(m,n) \mid \mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left[\mathbf{Agg_{CF}}\left\{w_1\rho_{\mathbf{Agg}}\left(\mathbf{s},\mathbf{x}^1\right), ..., w_N\rho_{\mathbf{Agg}}\left(\mathbf{s},\mathbf{x}^N\right)\right\}\right], \quad (10)$$

it becomes the Fréchet aggregation MIMO-filters. They are based on an arbitrary pair of aggregation operators $\mathbf{Agg_{CF}}$ and $\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}\right)$, which could be changed independently of one another.

In the first part [6], the notion of digital nonlinear SISO-filters (single-input single-output) associated with aggregation operators of averaging types was defined. In this part, we are going to consider a general theory of nonlinear MIMO-filters (multi-input multi-output). They are based on the generalized Fréchet point and on an arbitrary pair of aggregation operators $\mathbf{Agg_{CF}}, \rho_{\mathbf{Agg}}$, which could be changed independently of one another. For each pair of aggregation operators, we get the unique class of new nonlinear filters. We show that a large body of non-linear filters proposed to date constitute a proper subset of aggregation digital MIMO-filters.

## 3. Methods

### *3.1. Aggregation operators*

The aggregation problem consist in aggregating $N$-tuples of objects $\left(x_1, x_2, ..., x_N\right)$ all belonging to a given set $D$, into a single object of the same set $D$, *i.e.*, $\mathbf{Agg}: D^N \to D$. In fuzzy logic theory, the set $D$ is an interval of the real $D = [0,1] \subset \mathbf{R}$. In image processing theory $D = [0,255] \subset \mathbf{Z}$. In this setting, an aggregation operator is simply a function, which assigns a number $y$ to any $N$-tuple $\left(x_1, x_2, ..., x_N\right)$ of numbers that satisfies [15]:

1) $\mathbf{Agg}_N(x_1, x_2, ..., x_N)$ is continuous and monotone in each variable; to be definite, we assume that $\mathbf{Agg}$ is increasing in each variable.
2) The aggregation of identical numbers is equal to their common value: $\mathbf{Aggreg}_N(x, x, ..., x) = x$.
3) $\mathbf{Min}\left(x_1, ..., x_N\right) \le \mathbf{Agg}\left(x_1, ..., x_N\right) \le \mathbf{Max}\left(x_1, ..., x_N\right)$. Here $\mathbf{Min}\left(x_1, x_2, ..., x_N\right)$ and $\mathbf{Max}\left(x_1, x_2, ..., x_N\right)$ are the *minimum* and the *maximum* values among the elements of $\left(x_1, x_2, ..., x_N\right)$.
4) $\mathbf{Agg}\left(x_1, x_2, ..., x_N\right)$ is a symmetric function: $\mathbf{Agg}\left(x_{\sigma(1)}, x_{\sigma(2)}, ..., x_{\sigma(N)}\right) = \mathbf{Agg}\left(x_1, x_2, ..., x_N\right)$, $\forall\sigma \in \mathbf{S}_N$ of $\left\{1, 2, ..., N\right\}$, where $\mathbf{S}_N$ is the set of all permutations of $1, 2, ..., N$. In this case $\mathbf{Agg}\left(x_1, ..., x_N\right)$ is invariant (symmetric) with respect to the permutations of the elements of $\left(x_1, x_2, ..., x_N\right)$.

We list below a few particular cases of aggregation means:

1) Arithmetic and weighted means ( $K(x) = x$ ): $\mathbf{Arithm}\left(x_1, x_2, ..., x_N\right) = N^{-1}\sum_{i=1}^N x_i$, $\mathbf{WArithm}\left(x_1, x_2, ..., x_N\right) = \sum_{i=1}^N \bar{w}_i x_i$,

where $\sum_{i=1}^N \bar{w}_i = 1$. Classical operator $\mathbf{Arithm}\left(x_1, x_2, ..., x_N\right)$ is interesting because it gives an aggregated value that is smaller than the greatest argument and bigger than the smallest one. Therefore, the resulting aggregation is "a middle value". This property is known as the compensation property that is described mathematically by: $\mathbf{min}\left(x_1, x_2, ..., x_N\right) \le \mathbf{Arith}\left(x_1, x_2, ..., x_N\right) \le$ $\le \mathbf{max}\left(x_1, x_2, ..., x_N\right)$, where $\mathbf{min}\left(x_1, x_2, ..., x_N\right)$) and $\mathbf{max}\left(x_1, x_2, ..., x_N\right)$ are the algebraic minimum and maximum functions, respectively. The mappings **min** and **max** both satisfy the defining conditions and therefore are aggregations (means), even though they are rarely mentioned - or even perceived - as such. It is often used since it is simple and satisfies the properties of monotonicity, continuity, symmetry, associativity, idempotence and stability for linear transformations.

2) Another operator that follows the idea obtaining "a middle value" is the $k$-order statistic. It consists in ordering the arguments from the smallest one to the biggest one $\left(x_1, x_2, ..., x_m, ..., x_N\right) \to \left(x_{(1)}, x_{(2)}, ..., x_{(m)}, ..., x_{(N)}\right)$ (from the smallest to the biggest element, where $N = 2m+1$). The $k$-order statistic chooses the element on the $k$ th position on the ordered list: $\mathbf{OS}_k\left(x_1, x_2, ..., x_k, ..., x_N\right) = \mathbf{OS}_k\left(x_{(1)}, x_{(2)}, ..., x_{(k)}, ..., x_{(N)}\right) = x_{(k)}$. This aggregation operator satisfies the boundary conditions, the monotonicity, the symmetry, the idempotence and evidently the compensation behavior.

3) Three remarkable particular cases of the $k$-order statistic are the minimum, median and maximum: $\mathbf{min}\left(x_1, x_2, ..., x_N\right) = x_{(1)}$, $\mathbf{med}\left(x_1, x_2, ..., x_N\right) = x_{(m)}$, $\mathbf{max}\left(x_1, x_2, ..., x_N\right) = x_{(N)}$. The minimum gives the smallest value of a set, while the maximum gives

the greatest one. They are aggregation operators since they satisfy the axioms of the definition. The main properties of these operators are monotonicity, symmetry, associativity, idempotence. Mathematically speaking they have a compensation behavior, but these are the limit cases. Using these operators, we will never obtain an aggregated value "in the middle". For this reason, we do not consider that we can talk about compensation behavior in this case.

4) Very notable particular case corresponds to the function $K(x)=x^p$. We obtain then Hölder mean: $\mathbf{Hol}_p(x_1,x_2,...,x_N)=$ $=\mathbf{Power}_p(x_1,x_2,...,x_N)=\left(\frac{1}{N}\sum_{i=1}^{N}x_i^p\right)^{1/p}$. It is easy to see, that distances are particular cases of aggregation operator. We can use an arbitrary $\mathbf{Agg}$ as $\rho_{\mathbf{Agg}}(\mathbf{c},\mathbf{x})$. For example,

1) The Kolmogorov aggregation distances $\rho_{\mathbf{Agg}}(\mathbf{c},\mathbf{x}^i)=\rho_{Kol}(\mathbf{c},\mathbf{x}^i)=K^{-1}\left(\sum_{k=1}^{K}K\left(\left|c_k-x_k^i\right|\right)\right)$.

2) In particular, the Hölder aggregation distances $\rho_{\mathbf{Agg}}(\mathbf{c},\mathbf{x}^i)=\rho_p(\mathbf{c},\mathbf{x}^i)=\sqrt[p]{\sum_{k=1}^{K}\left|c_k-x_k^i\right|^p}$.

3) The $k$-order statistic distance $\rho_{\mathbf{Agg}}(\mathbf{c},\mathbf{x}^i)=\rho_{OS_k}(\mathbf{c},\mathbf{x}^i)=\mathbf{OS}_k\left(\left|c_1-x_1^i\right|,\left|c_2-x_2^i\right|,...,\left|c_K-x_K^i\right|\right)$.

Three remarkable particular cases are maximum, median and minimum distances $\rho_{\mathbf{Agg}}(\mathbf{c},\mathbf{x}^i)=\rho_{\max}(\mathbf{c},\mathbf{x}^i)=\mathbf{max}\left(\left|c_1-x_1^i\right|,...,\left|c_K-x_K^i\right|\right)$,
$\rho_{\mathbf{Agg}}(\mathbf{c},\mathbf{x}^i)=\rho_{\text{med}}(\mathbf{c},\mathbf{x}^i)=\mathbf{med}\left(\left|c_1-x_1^i\right|,...,\left|c_K-x_K^i\right|\right)$, $\rho_{\mathbf{Agg}}(\mathbf{c},\mathbf{x}^i)=\rho_{\min}(\mathbf{c},\mathbf{x}^i)=\mathbf{min}\left(\left|c_1-x_1^i\right|,...,\left|c_K-x_K^i\right|\right)$.

The same situation is true for the aggregation Fréchet cost function $\mathbf{Agg_{CF}}$. Using different aggregation operators, we can obtain different aggregation Fréchet cost functions. For example,

1) The Kolmogorov-Fréchet cost functions $\mathbf{Agg_{CF}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{Kol_{CF}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=K^{-1}\left(\sum_{i=1}^{K}K\left(\rho_{\mathbf{Agg}}^{(i)}\right)\right)$.

2) In particular, the Hölder-Fréchet cost functions $\mathbf{Agg_{CF}}\left\{\rho_{\mathbf{Agg}}^{(1)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{Hol}_p\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\sqrt[p]{\sum_{k=1}^{K}\left(\rho_{\mathbf{Agg}}^{(k)}\right)^p}$.

3) The $k$-order statistic-Fréchet cost functions $\mathbf{Agg_{CF}}\left\{\rho_{\mathbf{Agg}}^{(1)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{Agg}_{\mathbf{OS}}^{k}\left\{\rho_{\mathbf{Agg}}^{(1)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{OS}_k\left\{\rho_{\mathbf{Agg}}^{(1)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}$.

Three remarkable particular cases are maximum-, median- and minimum-Fréchet cost functions

$$\mathbf{Agg_{CF}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{Agg_{max}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{max}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\},$$
$$\mathbf{Agg_{CF}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{Agg_{med}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{med}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\},$$
$$\mathbf{Agg_{CF}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{Agg_{min}}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}=\mathbf{min}\left\{\rho_{\mathbf{Agg}}^{(1)},\rho_{\mathbf{Agg}}^{(2)},...,\rho_{\mathbf{Agg}}^{(N)}\right\}$$

and so on. Every pair $\mathbf{Agg_{CF}},\rho_{\mathbf{Agg}}$ gives us an exotic Fréchet aggregation MIMO-filters (17). For example,

$$\hat{\mathbf{s}}(i,j)=\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg}_{Hol},\rho_{\max}\,|\,\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left[\mathbf{Agg}_{Hol}\left\{\mathbf{max}\left(\mathbf{s}-\mathbf{x}^1\right),\mathbf{max}\left(\mathbf{s}-\mathbf{x}^2\right),...,\mathbf{max}\left(\mathbf{s}-\mathbf{x}^N\right)\right\}\right]=$$
$$=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left[\sqrt[p]{\left[\mathbf{max}\left(\mathbf{s}-\mathbf{x}^1\right)\right]^p+\left[\mathbf{max}\left(\mathbf{s}-\mathbf{x}^2\right)\right]^p+...+\left[\mathbf{max}\left(\mathbf{s}-\mathbf{x}^N\right)\right]^p}\right],$$

$$\hat{\mathbf{s}}(i,j)=\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg}_{Kol},\rho_{\max}\,|\,\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left[\mathbf{Agg}_{Hol}\left\{\mathbf{max}\left(\mathbf{s}-\mathbf{x}^1\right),\mathbf{max}\left(\mathbf{s}-\mathbf{x}^2\right),...,\mathbf{max}\left(\mathbf{s}-\mathbf{x}^N\right)\right\}\right]=$$
$$=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}K^{-1}\left[K\left(\mathbf{max}\left(\mathbf{s}-\mathbf{x}^1\right)\right)+K\left(\mathbf{max}\left(\mathbf{s}-\mathbf{x}^2\right)\right)+...+K\left(\left(\mathbf{s}-\mathbf{x}^N\right)\right)\right].$$

### *3.2. Suboptimal 2D Fréchet MIMO-filters*

In computation point view, it is better to restrict the infinite search domain from $\mathbf{R}^K$ to a finite subset $\mathbf{D}\in\mathbf{R}^K$. We are going to use the following finite subsets:

- The set of observed data $\mathbf{D}_{ob}=\left\{\mathbf{x}^1,\mathbf{x}^2,...,\mathbf{x}^N\right\}\subset\mathbf{R}^K$.
- The hyperspectral hypercube $\mathbf{D}_{dig}:=[0,255]^K$. For example, if $K=3$, then $\mathbf{D}_{dig}=[0,255]^3$ is the RGB-color cube.

In this case, we obtain definition of $\mathbf{D}$-optimal Fréchet points.

**Definition 2**. The suboptimal classical Fréchet mean and median (or $\mathbf{D}$-optimal Fréchet points) associated with the classical metric $\rho(\mathbf{x},\mathbf{y})$ are the points $\mathbf{c}_{subopt}^{dig}\in\mathbf{D}_{dig}=[0,255]^K\subset\mathbf{R}^K$ and $\mathbf{c}_{subopt}^{ob}\in\mathbf{D}_{ob}=\{\mathbf{x}^1,\mathbf{x}^2,...,\mathbf{x}^N\}\subset\mathbf{R}^K$ that minimizes the classical FCF over restricted search domains $\mathbf{D}_{dig}=[0,255]^K$ and $\mathbf{D}_{ex}=\{\mathbf{x}^1,\mathbf{x}^2,...,\mathbf{x}^N\}$

$$\mathbf{c}_{subopt}^{dig}=\mathbf{D}_{dig}\text{-}\mathbf{FrechPt}\left(\rho;w^1,w^2,...,w^N\mid\mathbf{x}^1,\mathbf{x}^2,...,\mathbf{x}^N\right)=\mathbf{arg}\min_{\mathbf{c}\in\mathbf{D}_{dig}}\left(\sum_{i=1}^N w_i\rho\left(\mathbf{c},\mathbf{x}^i\right)\right),$$

$$\mathbf{c}_{subopt}^{ob}=\mathbf{D}_{ob}\text{-}\mathbf{FrechPt}\left(\rho;w^1,w^2,...,w^N\mid\mathbf{x}^1,\mathbf{x}^2,...,\mathbf{x}^N\right)=\mathbf{arg}\min_{\mathbf{c}\in\mathbf{D}_{ob}}\left(\sum_{i=1}^N w_i\rho\left(\mathbf{c},\mathbf{x}^i\right)\right).$$

We use $\mathbf{D}$-Fréchet points for constructing the following nonlinear digital MIMO-filters

$$\hat{\mathbf{s}}_{opt}(i,j)=\mathbf{R}^K\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{FrechPt}}\left[\rho,w(m,n)\mid\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left[\sum_{(m,n)\in M_{(i,j)}}w(m,n)\rho\left(\mathbf{s},\mathbf{f}(m,n)\right)\right],\quad(11)$$

$$\hat{\mathbf{s}}_{subopt}^{dig}(i,j)=\mathbf{D}_{dig}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{FrechPt}}\left[\rho,w(m,n)\mid\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{D}_{dig}}\left[\sum_{(m,n)\in M_{(i,j)}}w(m,n)\rho\left(\mathbf{s},\mathbf{f}(m,n)\right)\right],\quad(12)$$

$$\hat{\mathbf{s}}_{subopt}^{ob}(i,j)=\mathbf{D}_{ob}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{FrechPt}}\left[\rho,w(m,n)\mid\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{D}_{ob}}\left[\sum_{(m,n)\in M_{(i,j)}}w(m,n)\rho\left(\mathbf{s},\mathbf{f}(m,n)\right)\right].\quad(13)$$

The next generalization of Fréchet MIMO-filters is based on the following suboptimal Fréchet points.

**Definition 3**. The suboptimal generalized Fréchet points associated with an aggregation metric $\rho_{\mathbf{Agg}}(\mathbf{x},\mathbf{y})$ are the points $\mathbf{c}_{subopt}^{dig}\in\mathbf{D}_{dig}=[0,255]^K\subset\mathbf{R}^K$ and $\mathbf{c}_{subopt}^{ob}\in\mathbf{D}_{ob}=\{\mathbf{x}^1,\mathbf{x}^2,...,\mathbf{x}^N\}\subset\mathbf{R}^K$ that minimizes the AFCF over restricted search domains $\mathbf{D}_{dig}=[0,255]^K$ *and* $\mathbf{D}_{ob}=\{\mathbf{x}^1,\mathbf{x}^2,...,\mathbf{x}^N\}$

$$\mathbf{c}_{subopt}^{dig}=\mathbf{D}_{dig}\text{-}\mathbf{AgFrechPt}\left(\mathbf{Agg}_{\mathbf{CF}},\rho_{\mathbf{Agg}};w^1,...,w^N\mid\mathbf{x}^1,...,\mathbf{x}^N\right)=\mathbf{arg}\min_{\mathbf{c}\in\mathbf{D}_{dig}}\left[\mathbf{Agg}_{\mathbf{CF}}\left\{w_1\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^1\right),...,w_N\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^N\right)\right\}\right],$$

$$\mathbf{c}_{subopt}^{ob}=\mathbf{D}_{ob}\text{-}\mathbf{AgFrechPt}\left(\mathbf{Agg}_{\mathbf{CF}},\rho_{\mathbf{Agg}};w^1,...,w^N\mid\mathbf{x}^1,...,\mathbf{x}^N\right)=\mathbf{arg}\min_{\mathbf{c}\in\mathbf{D}_{ob}}\left[\mathbf{Agg}_{\mathbf{CF}}\left\{w_1\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^1\right),...,w_N\rho_{\mathbf{Agg}}\left(\mathbf{c},\mathbf{x}^N\right)\right\}\right].$$

We use these points for constructing the following nonlinear digital MIMO-filters

$$\hat{\mathbf{s}}_{opt}(i,j)=\mathbf{R}^K\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg}_{\mathbf{CF}},\rho_{\mathbf{Agg}},w(m,n)\mid\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left[\underset{(m,n)\in M_{(i,j)}}{\mathbf{Agg}_{\mathbf{CF}}}\left\{w(m,n)\cdot\rho_{\mathbf{Agg}}\left(\mathbf{s},\mathbf{f}(m,n)\right)\right\}\right],\quad(14)$$

$$\hat{\mathbf{s}}_{subopt}^{dig}(i,j)=\mathbf{D}_{dig}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg}_{\mathbf{CF}},\rho_{\mathbf{Agg}},w(m,n)\mid\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{D}_{dig}}\left[\underset{(m,n)\in M_{(i,j)}}{\mathbf{Agg}_{\mathbf{CF}}}\left\{w(m,n)\cdot\rho_{\mathbf{Agg}}\left(\mathbf{s},\mathbf{f}(m,n)\right)\right\}\right],\quad(15)$$

$$\hat{\mathbf{s}}_{subopt}^{ob}(i,j)=\mathbf{D}_{ob}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg}_{\mathbf{CF}},\rho_{\mathbf{Agg}},w(m,n)\mid\mathbf{f}(m,n)\right]=\mathbf{arg}\min_{\mathbf{s}\in\mathbf{D}_{ob}}\left[\underset{(m,n)\in M_{(i,j)}}{\mathbf{Agg}_{\mathbf{CF}}}\left\{w(m,n)\cdot\rho_{\mathbf{Agg}}\left(\mathbf{s},\mathbf{f}(m,n)\right)\right\}\right].\quad(16)$$

### *3.3. Examples of Fréchet MIMO-filters*

**Example 1.** If observation data are real numbers, *i.e.*, $x^1,x^2,...,x^N\in\mathbf{R}$, and the distance function is the city distance $\rho(x,y)=\rho_1(x,y)=|x-y|$, AFCF is $L_p$-distance, then the optimal and suboptimal Fréchet points for data $x^1,x^2,...,x^N\in\mathbf{R}$ to be

$$c_{opt}=\mathbf{R}^K\text{-}\mathbf{AgFrechPt}\left(\mathbf{Hol}_p,\rho_{\mathbf{Agg}}=\rho_1\mid x^1,x^2,...,x^N\right)=\mathbf{arg}\min_{c\in\mathbf{R}}\left(\sqrt[p]{\sum_{i=1}^N\left|c-x^i\right|^p}\right),$$

$$c_{subopt}^{dig}=\mathbf{D}\text{-}\mathbf{FrechPt}\left(\mathbf{Hol}_p,\rho_{\mathbf{Agg}}=\rho_1\mid x^1,x^2,...,x^N\right)=\mathbf{arg}\min_{c\in\mathbf{D}_{dig}}\left(\sqrt[p]{\sum_{i=1}^N\left|c-x^i\right|^p}\right),$$

$$c_{subopt}^{ob}=\mathbf{D}\text{-}\mathbf{FrechPt}\left(\mathbf{Hol}_p,\rho_{\mathbf{Agg}}=\rho_1\mid x^1,x^2,...,x^N\right)=\mathbf{arg}\min_{c\in\mathbf{D}_{ob}}\left(\sqrt[p]{\sum_{i=1}^N\left|c-x^i\right|^p}\right).$$

In particular, if $p=1,2,\infty$ then we obtain the Fréchet point (**FrechPt**), arithmetic mean (**ArithMean**) and midrange (**MidPt**) of a set of observations $x^1,x^2,...,x^N\in\mathbf{R}$, respectively:

$$c_{opt} = \mathbf{R}\text{-}\mathbf{AgFrechPt}\left(\mathbf{Hol}_1, \rho_{\mathbf{Agg}} = \rho_1 \mid x^1, x^2, ..., x^N\right) = \mathbf{arg}\min_{c\in\mathbf{R}}\left(\sum_{i=1}^{N}\left|c - x^i\right|\right) \equiv \mathbf{FrechPt}(x^1, x^2, ..., x^N),$$

$$c_{opt} = \mathbf{R}\text{-}\mathbf{AgFrechPt}\left(\mathbf{Hol}_2, \rho_{\mathbf{Agg}} = \rho_1 \mid x^1, x^2, ..., x^N\right) = \mathbf{arg}\min_{c\in\mathbf{R}}\left(\sqrt{\sum_{i=1}^{N}\left|c - x^i\right|^2}\right) \equiv \mathbf{ArithMean}(x^1, x^2, ..., x^N) = \frac{1}{N}\sum_{i=1}^{N} x^i,$$

$$c_{opt} = \mathbf{R}\text{-}\mathbf{AgFrechPt}\left(\mathbf{Hol}_\infty, \rho_{\mathbf{Agg}} = \rho_1 \mid x^1, x^2, ..., x^N\right) = \mathbf{arg}\min_{c\in\mathbf{R}}\left(\max_{i=1,2,..N}\left|c - x^i\right|\right) \equiv \mathbf{MidPt}(x^1, x^2, ..., x^N) =$$

$$= \left[\mathbf{Max}(x^1, x^2, ..., x^N) + \mathbf{Min}(x^1, x^2, ..., x^N)\right]/2.$$

In this case, filters

$$\hat{s}_{opt}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{R}\text{-}\mathbf{AgFrechPt}}\left[\mathbf{Hol}_p, \rho_{\mathbf{Agg}} = \rho_1 \mid f(m,n)\right] = \mathbf{arg}\min_{s\in R}\left[\sqrt[p]{\sum_{(m,n)\in M_{(i,j)}}\left|s - f(m,n)\right|^p}\right]. \quad (17)$$

are the optimal maximum likelihood SISO-filter for Laplacian ( $p = 1$), Gaussian ( $p = 2$ ) and Uniform (on [-0.5,+0.5]) PDF ( $p = \infty$) of noises, respectively.

If $\mathbf{D} = \mathbf{D}_{dig} = [0, 255]$ then we obtain the following suboptimal estimates (for the same values $p = 1, 2, \infty$):

$$c_{subopt}^{dig} = \mathbf{D}_{dig}\text{-}\mathbf{AgFrechPt}\left(\mathbf{Hol}_p, \rho_{\mathbf{Agg}} = \rho_1 \mid x^1, x^2, ..., x^N\right) = \mathbf{arg}\min_{c\in[0,255]}\left(\sqrt[p]{\sum_{i=1}^{N}\left|c - x^i\right|^p}\right).$$

In this case, filters

$$\hat{s}_{subopt}^{dig}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{D}_{dig}\text{-}\mathbf{AgFrechPt}}\left[\mathbf{Hol}_p, \rho_1) \mid f(m,n)\right] = \mathbf{arg}\min_{s\in[0,255]}\left[\sqrt[p]{\sum_{(m,n)\in M_{(i,j)}}\left|s - f(m,n)\right|^p}\right]. \quad (18)$$

are the suboptimal maximum likelihood SISO-filter for the same noises.

If $\mathbf{D} = \mathbf{D}_{ob} = \left\{x^1, x^2, ..., x^N\right\}$ then we obtain the next suboptimal estimates (for the same values $p = 1, 2, \infty$): the suboptimal Fréchet point is the classical median (**Med**), arithmetic mean (**ArithMean**) and midrange (**MidPt**) of a set of observations $x^1, x^2, ..., x^N \in \mathbf{R}$, respectively:

$$c_{subopt}^{ob} = \mathbf{D}_{ob}\text{-}\mathbf{AgFrechPt}\left(\mathbf{Hol}_p, \rho_{\mathbf{Agg}} = \rho_1 \mid x^1, x^2, ..., x^N\right) = \mathbf{arg}\min_{c\in D_{ob}}\left(\sqrt[p]{\sum_{i=1}^{N}\left|c - x^i\right|^p}\right).$$

In this case, filters

$$\hat{s}_{subopt}^{ob}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{D}_{ob}\text{-}\mathbf{AgFrechPt}}\left[\mathbf{Hol}_p, \rho_1 \mid f(m,n)\right] = \mathbf{arg}\min_{s\in D_{ob}}\left[\sqrt[p]{\sum_{(m,n)\in M_{(i,j)}}\left|s - f(m,n)\right|^p}\right]. \quad (19)$$

are the suboptimal maximum likelihood SISO-filter for the same noises. It is interesting that only in the first case we have classical median filter $\underset{(m,n)\in M_{(i,j)}}{\mathbf{Med}}\left[f(m,n)\right] \equiv \underset{(m,n)\in M_{(i,j)}}{\mathbf{D}_{ob}\text{-}\mathbf{AgFrechPt}}\left[\mathbf{Hol}_1, \rho_1, \mid f(m,n)\right] = \mathbf{arg}\min_{s\in D_{ob}}\left[\sum_{(m,n)\in M_{(i,j)}}\left|s - f(m,n)\right|\right].$

**Example 2.** If observation data are vectors, *i.e.*, $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N \in \mathbf{R}^K$ and the distance function is $\ell_p$ distance $\rho_{\mathbf{Agg}}(\mathbf{x}, \mathbf{y}) = \rho_p(\mathbf{x}, \mathbf{y}) = \left\|\mathbf{x} - \mathbf{y}\right\|_p$ then we have

$$\mathbf{c}_{opt} = \mathbf{R}^K\text{-}\mathbf{AgFrechPt}\left(\mathbf{Agg}_{\mathbf{CF}}, \rho_p; w^1, ..., w^N \mid \mathbf{x}^1, ..., \mathbf{x}^N\right) = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\mathbf{Agg}_{\mathbf{CF}}\left\{w_1\left\|\mathbf{c} - \mathbf{x}^1\right\|_p, ..., w_N\left\|\mathbf{c} - \mathbf{x}^N\right\|_p\right\}\right],$$

In particular, if $\mathbf{Agg}_{\mathbf{CF}} = \mathbf{Hol}_q$ (and $w^1 = w^2 = ... = w^N$) then

$$\mathbf{c}_{opt} = \mathbf{R}^K\text{-}\mathbf{AgFrechPt}\left(\mathbf{Hol}_q, \rho_p \mid \mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^N\right) = \mathbf{arg}\min_{\mathbf{c}\in\mathbf{R}^K}\left[\sqrt[q]{\sum_{i=1}^{N}\left(\sum_{k=1}^{K}\left|c_k - x_k^i\right|^p\right)^{q/p}}\right].$$

If $p = 1, 2, \infty$ and $q = 1, 2, \infty$ then we obtain nine Fréchet points and nine Fréchet MIMO-filters

$$\hat{s}_{opt}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{R}\text{-}\mathbf{AgFrechPt}}\left[\mathbf{Hol}_q, \rho_p \mid \mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left[\sqrt[q]{\sum_{(m,n)\in M_{(i,j)}}\left(\sum_{k=1}^{K}\left|s_k - f_k(m,n)\right|^p\right)^{q/p}}\right], \quad (20)$$

For each pair of aggregation operators, we get the unique class of new nonlinear filters. If one can accurately model the noise distribution, then the filtering results can be significantly improved by using a suitable metric $\rho_{\mathbf{Agg}}(\mathbf{c}, \mathbf{x})$ or aggregation cost function $\mathbf{Agg}_{\mathbf{CF}}$. The link between the noise distribution and the metric is given by the maximum likelihood theory.

## 4. Results and Discussion

We performed a number of experiments with the proposed MIMO-filters using several images. The results of some of them are presented here. We developed five the following filtering algorithms:

1) Classical arithmetic mean MIMO-filter (**Mean**)

$$\hat{\mathbf{s}}(i,j) = \underset{(m,n)\in M_{(i,j)}}{\mathbf{FrechMean}}\left[\mathbf{Agg_{CF}} = \mathrm{Hol}_1, \rho_1^2 \mid \mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\mathbf{s}\in\mathbf{R}^K}\left(\sum_{(m,n)\in M_{(i,j)}} \left\|\mathbf{s}-\mathbf{f}(m,n)\right\|_1^2\right) = \frac{1}{N}\sum_{(m,n)\in M_{(i,j)}}\mathbf{f}(m,n),$$

a) Original image. b) Noise images, PSNR=25.08.

$c)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{ob} \in \mathbf{R}_r \times \mathbf{R}_g \times \mathbf{R}_b,$ $PSNR = 27.69.$

$d)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{dig} \in \mathbf{R}_r \times \mathbf{R}_g \times \mathbf{R}_b,$ $PSNR = 29.26.$

$e)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{ob} \in \mathbf{R}^3_{rgb},$ $PSNR = 29.92.$

$f)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{dig} \in \mathbf{R}^3_{rgb},$ $PSNR = 30.48.$

**Fig. 3.** Original (a) and noise (b) images.
Impulse noise: "Uniform PD". Denoised images (c)-(f).

a) Original image. b) Noise images, PSNR=15.27.

$c)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{ob} \in \mathbf{R}_r \times \mathbf{R}_g \times \mathbf{R}_b,$ $PSNR = 20.60.$

$d)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{dig} \in \mathbf{R}_r \times \mathbf{R}_g \times \mathbf{R}_b,$ $PSNR = 29.70.$

$e)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{ob} \in \mathbf{R}^3_{rgb},$ $PSNR = 29.61.$

$f)$ $\mathbf{Agg_{CF}} = \mathbf{Hol}_1,\ \rho_{\mathbf{Agg}} = \rho_2,$ $\mathbf{D} = \mathbf{D}_{dig} \in \mathbf{R}^3_{rgb},$ $PSNR = 29.81.$

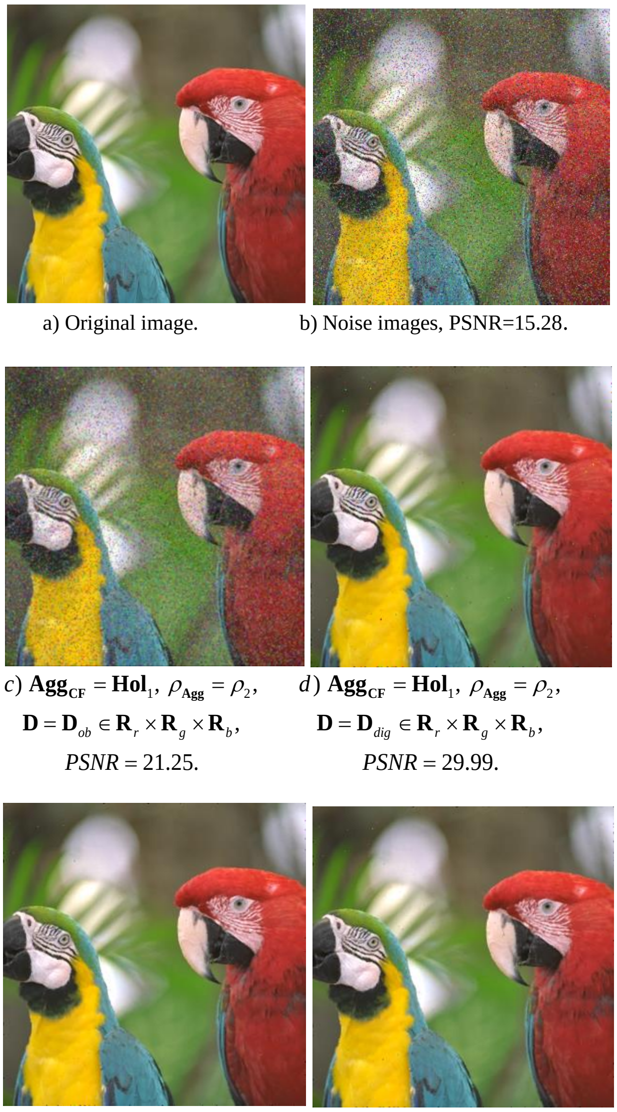
**Fig. 4.** Original (a) and noise (b) images.
Impulse noise: "Black-white salt-pepper". Denoised images (c)-(f).

2) Classical vector-valued median filter - independent median filtering along every channel $\mathbf{R}_r, \mathbf{R}_g$ and $\mathbf{R}_b$ with the research domain in the form of observed data $s_r \in D_{r,\mathrm{ob}} \subset [0,255]_r$, $s_g \in D_{g,\mathrm{ob}} \subset [0,255]_g$, $s_b \in D_{b,\mathrm{ob}} \subset [0,255]_b$ (three median SIS0-filters acting in each channel. Our designate $D_{ob}$-$SISO^3$ -filter or **Med**)

$$\hat{\mathbf{s}}^{ob}_{subopt}(i,j) = \mathbf{D}_{ob}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg_{CF}} = \mathrm{Hol}_1, \rho_1, w(m,n) \mid \mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\substack{s_r\in D_{r,\mathrm{ob}}\subset[0,255]_r\\ s_g\in D_{g,\mathrm{ob}}\subset[0,255]_g\\ s_b\in D_{b,\mathrm{ob}}\subset[0,255]_b}}\left[\sum_{(m,n)\in M_{(i,j)}} \left|\mathbf{s}-\mathbf{f}(m,n)\right|\right]. \quad (21)$$

3) Elaborated vector-valued median filter - independent median filtering along every channel $\mathbf{R}_r, \mathbf{R}_g$ and $\mathbf{R}_b$ with the research domain in the form of digital domains $s_r \in [0,255]_r$, $s_g \in [0,255]_g$, $s_b \in [0,255]_b$ (three elaborated median SIS0-filters acting in each channel - $D_{dig}$-$SISO^3$-filter or **ElabMed**)

$$\hat{\mathbf{s}}^{dig}_{subopt}(i,j) = \mathbf{D}_{ob}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg_{CF}} = \mathrm{Hol}_1, \rho_1, w(m,n)\,|\,\mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\substack{s_r\in[0,255]_r\\ s_g\in[0,255]_g\\ s_b\in[0,255]_b}}\left[\sum_{(m,n)\in M_{(i,j)}} \left|\mathbf{s}-\mathbf{f}(m,n)\right|\right]. \quad (22)$$



a) Original image. b) Noise images, PSNR=15.28.

c) $\mathbf{Agg_{CF}} = \mathbf{Hol}_1$, $\rho_{\mathbf{Agg}} = \rho_2$, $\mathbf{D} = \mathbf{D}_{ob} \in \mathbf{R}_r \times \mathbf{R}_g \times \mathbf{R}_b$, $PSNR = 21.25$.

d) $\mathbf{Agg_{CF}} = \mathbf{Hol}_1$, $\rho_{\mathbf{Agg}} = \rho_2$, $\mathbf{D} = \mathbf{D}_{dig} \in \mathbf{R}_r \times \mathbf{R}_g \times \mathbf{R}_b$, $PSNR = 29.99$.

e) $\mathbf{Agg_{CF}} = \mathbf{Hol}_1$, $\rho_{\mathbf{Agg}} = \rho_2$, $\mathbf{D} = \mathbf{D}_{ob} \in \mathbf{R}^3_{rgb}$, $PSNR = 31.86$.

f) $\mathbf{Agg_{CF}} = \mathbf{Hol}_1$, $\rho_{\mathbf{Agg}} = \rho_2$, $\mathbf{D} = \mathbf{D}_{dig} \in \mathbf{R}^3_{rgb}$, $PSNR = 32.07$.

**Fig. 5.** Original (a) and noise (b) images.
Noise: "Color Salt-Pepper". Denoised images (c)-(f).

**Table 1.** Noise wit uniform pdf

| % | PSNR | **Mean** | **Med** | **ElabMed** | **VecMed** | **ElabVecMed** |
|---|---|---|---|---|---|---|
| 05 | 19.05 | 26.68 | 24.16 | 24.16 | 22.36 | 25.83 |
| 10 | 15.71 | 23.99 | 20.88 | 20.88 | 19.09 | 22.71 |
| 20 | 13.48 | 21.89 | 18.57 | 18.57 | 16.64 | 20.52 |
| 40 | 11.79 | 20.11 | 16.72 | 18.57 | 15.11 | 18.77 |
| 50 | 10.57 | 18.69 | 15.37 | 15.37 | 13.84 | 17.46 |
| 70 | 08.83 | 16.41 | 13.26 | 13.26 | 11.91 | 15.44 |

**Table 2.** Black-white salt-pepper noise

| % | PSNR | **Mean** | **Med** | **ElabMed** | **VecMed** | **ElabVecMed** |
|---|---|---|---|---|---|---|
| 01 | 15.27 | 23.47 | 32.97 | 32.92 | 32.92 | 33.02 |
| 05 | 12.49 | 20.60 | 29.80 | 29.80 | 29.72 | 29.80 |
| 10 | 10.94 | 18.87 | 25.54 | 25.46 | 25.46 | 25.50 |
| 20 | 09.90 | 17.47 | 21.81 | 21.67 | 21.67 | 21.74 |
| 50 | 09.15 | 16.51 | 18.92 | 18.92 | 18.75 | 18.82 |
| 70 | 08.54 | 15.67 | 16.61 | 16.61 | 16.40 | 16.51 |

**Table 3.** Color salt-pepper noise

| % | PSNR | **Mean** | **Med** | **ElabMed** | **VecMed** | **ElabVecMed** |
|---|---|---|---|---|---|---|
| 01 | 15.20 | 23.49 | 32.77 | 32.77 | 33.26 | 33.54 |
| 05 | 12.40 | 20.56 | 29.58 | 29.58 | 31.84 | 32.09 |
| 10 | 10.93 | 18.78 | 25.57 | 25.57 | 30.00 | 30.10 |
| 20 | 09.81 | 17.43 | 21.78 | 21.78 | 27.32 | 27.32 |
| 50 | 09.13 | 16.48 | 19.00 | 19.00 | 24.77 | 24.82 |
| 70 | 08.82 | 15.63 | 16.60 | 16.60 | 22.06 | 22.26 |

4) Classical vector-valued median MIMO-filter [13-14] with the research domain in the form of observed data $\mathbf{s}\in \mathbf{D}_{ob} = \{\mathbf{f}(m,n)\}_{(m,n)\in M(i,j)} \subset [0,255]^3_{rgb}$ ($\mathbf{D}_{ob}$-$MIMO$ or **VecMed**)

$$\hat{\mathbf{s}}^{ob}_{subopt}(i,j) = \mathbf{D}_{ob}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg_{CF}} = \mathrm{Hol}_1, \rho_1\,|\,\mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\mathbf{s}\in\mathbf{D}_{ob}\subset[0,255]^3_{rgb}}\left[\sum_{(m,n)\in M_{(i,j)}} \left|\mathbf{s}-\mathbf{f}(m,n)\right|\right], \quad (23)$$

5) Elaborated vector-valued median MIMO-filter with the research domain in the form of RGB-cube $\mathbf{s}\in[0,255]^3_{rgb}$ ( $\mathbf{D}_{dig}$-$MIMO$ or **ElabVecMed**)

$$\hat{\mathbf{s}}^{dig}_{subopt}(i,j) = \mathbf{D}_{dig}\text{-}\underset{(m,n)\in M_{(i,j)}}{\mathbf{AgFrechPt}}\left[\mathbf{Agg_{CF}} = \mathrm{Hol}_1, \rho_1\,|\,\mathbf{f}(m,n)\right] = \mathbf{arg}\min_{\mathbf{s}\in[0,255]^3_{rgb}}\left[\sum_{(m,n)\in M_{(i,j)}} \left|\mathbf{s}-\mathbf{f}(m,n)\right|\right], \quad (24)$$

For the experiments presented here, the "Macaw" images (Figures 2a,3a,4a, respectively) are used. Salt-Pepper and Unichannel Uniform-PDF noises are added to the images to obtain noised images with different peak signal-to-noise (PSNR). The performance evaluation of the filtering operation is quantified by the PSNR (Peak Signal to Noise Ratio). The proposed suboptimal Fréchet MIMO-filters (21)-(24) has been applied to noised $3\times3$ image "Macaw". We use $3\times3$ -window. The denoised images are shown in Fig. 3-5. Tables 1-3 are the filtering results at different intensities and types of noise. All Fréchet MIMO-filters ( $\mathbf{D}_{ob}$-*MIMO* and $\mathbf{D}_{dig}$-*MIMO* ) have very good denoised properties. It is easy to see that results for Fréchet filters $\mathbf{D}_{ob}$-*MIMO* and $\mathbf{D}_{dig}$-*MIMO* are better, compared to the classical **Mean- and-** $D_{ob}$-$SISO^3$ filters. Filter $D_{dig}$-$SISO^3$ gives elaborated results with respect to their classical counterpart $D_{ob}$-$SISO^3$. These facts confirm that further investigation of these new filters is perspective.

## 5. Conclusion

A new class of nonlinear generalized MIMO-filters for multichannel image processing are introduced in this paper. These filters are based on an arbitrary pair of aggregation operators, which could be changed independently of one another. For each pair of parameters, we get the unique class of new nonlinear MIMO-filters. The main goal of the work is to show that generalized Fréchet aggregation means can be used to solve problems of image filtering in a natural and effective manner.


## Acknowledgements

This work was supported by grants the RFBR No. 17-07-00886 and by Ural State Forest Engineering’s Center of Excellence in ”Quantum and Classical Information Technologies for Remote Sensing Systems”.


## References


[1] Astola, J. Fundamentals of Nonlinear Digital Filtering / J. Astola, P. Kuosmanen. – New York: CRC Press LLC. – 1997. – 288 p.
[2] Dougherty, E.R. An Introduction to Nonlinear Image Processing / E.R. Dougherty, J. Astola. – Washington: SPIE Press. – 1994. – 182 p.
[3] Mitra, S.K. Nonlinear Image Processing // Academic Press Series in Communications, Networking, and Multimedia. – San Diego, New York.– 2001.–248 p.
[4] Heath, T., A History of Greek Mathematics: From Thales to Euclid / Dover, New York. – 1981. – Vol. 1. – 172 p.
[5] Bullen, P. S. Means and their Inequalities / Mitrinovic, D. S., Vasic, P. M. // Mathematics and its Applications (East European Series), Reidel Publishing Co., Dordrecht, The Netherlands. – 1988. – P. 225-245.
[6] Labunets, V. G. Systematic approach to nonlinear filtering associated with aggregation operators. Part 1. SISO-filters // In this Proceedings.
[7] Labunets, V. G. Filters based on aggregation operators. Part 1. Aggregation Operators // Proc. of the 24th Int. Crimean Conference “Microwave & Telecommunication Technology”. – Sevastopol, Crimea, Russia. – 2014. – Vol. 24. – P. 1239 – 1240
[8] Labunets, V.G. Filters based on aggregation operators. Part 2. The Kolmogorov filters / Gainanov, D.N., Ostheimer, E. // Proc. of the 24th Int. Crimean Conference “Microwave & Telecommunication Technology” . – Sevastopol, Crimea, Russia. – Vol. 24. – P. 1241 – 1242.
[9] Labunets, V.G. Filters based on aggregation operators. Part 3. The Heron filters / Gainanov, D.N., Tarasov A.D., Ostheimer E. // Proc. of the 24th Int. Crimean Conference Microwave & Telecommunication Technology” – Sevastopol, Crimea, Russia. – Vol. 24. – P. 1243 – 1244.
[10] Labunets, V.G. Filters based on aggregation operators. Part 4. Generalized vector median flters/ Gainanov, D.N., Arslanova, R.A., Ostheimer, E. // Proc. of the 24th Int. Crimean Conference “Microwave& Telecommunication Technology” . – Sevastopol, Crimea, Russia. – Vol. 24. – P. 1245—1246.
[11] Fréchet, M. Les elements aleatoires de nature quelconque dans un espace distancie // Ann. Inst. H. Poincare. – 1948. – Vol.10. – No. 3. – P. 215–310.
[12] Chandrasekaran, R. Algebraic optimization: The Fermat-Weber problem/ Tamir, F. // Mathematical Programming. –1990. –Vol. 46. – P. 219-224.
[13] Astola, J. Vector median filters/ Haavisto, P., Neuvo, Y. // Proc. IEEE Trans. Image Processing. – 1990. – Vol. 78. – P. 678–689.
[14] Tang, K. Nonlinear multivariate image filtering techniques/ Astola, J., Neuvo, Y. // IEEE Trans. Image Processing. –1996. – Vol. 4. – P. 788–798.
[15] Mayor, G. On the representation of some Aggregation functions/ Trillas, E. // Proceeding of ISMVL. – 1986. – P. 111-114.