Non-additive interval-valued F-transform

O. Strauss *

LIRMM Université Montpellier II, 61 rue Ada, 34392 Montpellier cedex 5, France

Abstract

This article proposes a new interval-valued fuzzy transform. Its construction is based on a possibilistic interpretation of the partition on which the fuzzy transform is built. The main advantage of this approach is that it provides specific interval valued functions whose interpretation is straightforward. This interpretation relates to a traditional sampling/reconstruction framework where little is known about the sampling and/or reconstructing kernels. Numerous properties of the proposed approach are proved that could be useful for function analysis and comparison. In the experimental section, we illustrate some properties of the proposed transform while highlighting interesting features of the obtained framework.

Key words: Choquet integral, fuzzy transform, sampling, summative and maxitive kernels, possibility theory, non-additive confidence measure.

1 Introduction

Nowadays, signal processing is mainly achieved through algorithms on discrete and quantified representations of real signals. In this domain, transformations are essential to transpose a signal to another space in order to obtain a more compact and meaningful representation of the signal. Transformations are extensively used for analysis, compression, encryption, filtering, inversion, information retrieval, etc. Many transformations have been proposed in the relevant literature. Fourier and Laplace transforms are the most widely used, which associate a complex decomposition with any real signal in the frequency domain. The advent of the fast Fourier transform (FFT) enabled real-time computation of the Fourier transform and thus its wide use in many applications. The discrete cosine transform can be seen as a simplification of FFT that only keeps

* Corresponding author.
Email address: strauss@lirmm.fr (O. Strauss).

the real part of the Fourier transform, thus associating a real decomposition with a real signal. It became popular through its use in the jpeg compression method. The wavelet transform was more recently proposed as a better solution for analyzing signals having compact support. In image processing, more dedicated transformations have also been proposed. For example, the Hough transform places the image in a parametric space, thus facilitating the retrieval of specific parametrized features (e.g. lines, curves, etc.). Its close relative, the Radon transform, has been proposed to solve the problem of reconstructing an image from its projections. Sampling and subsampling can also be seen as transformations. Sampling consists of associating a bounded set of real valued samples with a continuous signal, while subsampling associates a reduced number of representative samples with a real high resolution discrete signal. The idea underlying working with a sampled (or subsampled) signal is to use an inherent redundancy of the signal, so as to make it possible to work with the original signal by only manipulating a reduced set of samples. It is used to solve continuous problems by finite computation or to reduce the computation complexity of a signal processing algorithm. Other transforms exist like those of Hilbert, Gabor, Zack, etc. Most of those transformations consist of convolving the real signal with an appropriate set of kernels, which is the basis of the transform.

The fuzzy transform (or F-transform) recently proposed by Irina Perfilieva [23] belongs to this transform family. It consists of associating, with an original continuous or discrete real signal, a reduced set of real samples by projecting this signal on a fuzzy partition *à la* Ruspini [29]. This has drawn a great deal of attention from the scientific community since it is one of the rare uses of the fuzzy framework to directly handle real functions without any linguistic interpretation. It has been used for data analysis [26], compression (see references in [23]), segmentation [17], coding [16], solving differential equations, forecasting [18], scheduling [12], trading [34], etc. A special issue of this journal was recently dedicated to advances in fuzzy transform theory and applications [27].

The fuzzy transform involves two operations: a direct fuzzy transform (F-transform) which is the decomposition itself and an inverse fuzzy transform (IF-transform) that goes from the sampled space to the original space. The word "inverse" may seem somewhat inappropriate since applying the IF-transform to the F-transform of a signal leads to an estimate that is not equal to the original signal. However, as shown in [23], an appropriate choice of fuzzy partition can make the reconstructed signal an approximation of the original signal with any arbitrary precision.

With most transforms, the question arises as to the existence of an inverse transform, i.e. *is it possible to reconstruct the original signal from its transformation*? Except for the wavelet transform, the answers to this question

given by the authors of most of the aforementioned transforms are highly debatable when applied to numerical signal processing. More precisely, if a signal is represented by a reduced number of values, the reconstruction of the original signal is risky. For example, the discrete Fourier transform has an inverse, i.e. it is possible to exactly reconstruct the original sampled signal from its transformed values. However, manipulations in the transformed space are often meaningless in the original space since the Fourier transform of a discrete signal is continuous while the FFT associates a discrete representation with a discrete signal. The Radon transform also has an inverse form in the continuous domain. However, this inverse form does not exist in the discrete space and a certain number of dedicated tools, including regularizations and optimizations, are required to invert a discrete Randon transform.

The inverse transform issue also naturally exists within the F-transform framework. The first proposition of Perfileva was to use the same shape function as that used to generate the partition to achieve both F- and IF-transforms. She proved that the obtained reconstruction locally minimizes a L_2 distance between the original and the reconstructed signal. However, as shown first by Crouzet [2] and then by Patané [21], this local criterion is not very relevant from a signal processing standpoint. Thus, in this setting, another basis should be preferred that leads to minimizing a global L_2 distance. As shown in [3], this leads to the least square interpolation procedure conventionally used in signal processing. However, this approach only applies to discrete functions.

Other techniques have been proposed to enhance the ability of the fuzzy transform framework to work with a simple representation of a signal. For example, Bede and Rudas [1] question the shape function of the weighting fuzzy numbers used to form fuzzy partitions. It appears, from a qualitative comparison, that the optimality of a particular shape function highly depends on the function to be represented. The partition can thus be adapted to the signal, as shown by Sefanini in [32]. The position of the partition nodes can also be adapted to have a higher concentration of atoms where the signal has more variations. However, adapting the partition to signals can require the use of two different partitions for two different signals, thus limiting the ability of using fuzzy transforms for combining or comparing two signals. For this kind of application, it is more interesting to use a fixed regular partition with a known approximation ability.

From a signal processing standpoint, the fuzzy transform framework looks like a sampling/interpolation process, classically used to solve continuous problems by discrete computations or to perform computations that are equivalent to continuous processes [15] [36]. Most results reported in the fuzzy transform literature are classical signal processing results. Thus, how does this fuzzy framework apply in the signal processing context? As base functions? As a

complementary tool?

A very interesting answer to this question was proposed by Perfileva in [23] whereby new fuzzy transforms based on residuated lattice operations were constructed. These new transforms lead to interval-valued direct and inverse fuzzy transforms. In this paper, we try to go further in this interval-valued direction by reinterpreting the fuzzy transform framework in the light of non-additive measures theory [6].

The paper is organized as follows. Section 2 presents useful definitions and notations. Section 3 introduces the ordinary F-transform as proposed in [23] as well as some new properties. Section 4 is dedicated to our proposition of an interval-valued non-additive F-transform. We provide some illustative experimentations in Section 5 and we conclude in Section 6 with remarks on the limitations of the current approach and propose some avenues for future study.

2 Preliminary considerations, definitions and notations

2.1 Notations

This study is restricted to a one-dimensional case of continuous and discrete real functions. Extending this study to more than one dimension is straightforward (see Section 5.1).

Let $\mathbb{R}$ be the real line and $\mathbb{IR}$ the set of all intervals of $\mathbb{R}$. The continuous functions we consider are defined on an interval $\Omega = [a, b] \subset \mathbb{R}$. We define $\mathcal{P}(\Omega)$ as the set of all Lebesgue measurable subsets of Ω. Let p be a positive integer, then $\Theta_p = \{0, \ldots, p\} \subset \mathbb{N}$ denotes the set of $(p + 1)$ positive integers and $\Delta_p = \frac{b-a}{p}$ is the sampling step. The k^{th} sampling location $(\omega_k)_{k\in\Theta_p}$ is defined as: $\forall k \in \Theta_p$, $\omega_k = a + k\Delta_p$. We also define $\mathcal{P}(\Theta_p)$ as the set of all subsets of Θ_p.

2.2 Confidence measures and integrals

A capacity, also called confidence measure, can be defined on both continuous and discrete reference sets.

Definition 1 *A continuous capacity ν is a set function $\nu : \mathcal{P}(\Omega) \to [0, 1]$ such that $\nu(\varnothing) = 0$, $\nu(\Omega) = 1$ and $\forall A, B \in \mathcal{P}(\Omega)$, $A \subseteq B \Rightarrow \nu(A) \leq \nu(B)$.*

Given a capacity ν, its conjugate ν^c, is defined as: $\nu^c(A) = 1 - \nu(A^c)$, for any subset $A \in \mathcal{P}(\Omega)$, with A^c being the complementary set of A in Ω. Note that, in this paper, to avoid confusion with other notations the classical $\bar{\nu}$ notation will not be used. A capacity ν such that for all A, B in $\mathcal{P}(\Omega)$, $\nu(A \cup B) + \nu(A \cap B) \leq \nu(A) + \nu(B)$ is said to be concave (or submodular or 2-alternating). This paper only considers such capacities. The core of a concave capacity ν, denoted $\mathcal{M}(\nu)$, is the set of probability measures P on $\mathcal{P}(\Omega)$ such that $\nu(A) \geq P(A)$ for all subsets $A \in \mathcal{P}(\Omega)$.

Remark 2 *If ν is a concave capacity, its conjugate capacity ν^c is convex, i.e. $\forall A, B \in \mathcal{P}(\Omega)$, $\nu^c(A \cup B) + \nu^c(A \cap B) \geq \nu^c(A) + \nu^c(B)$. Due to the conjugation relationship between ν and ν^c, the core of ν can be rewritten:*
$\mathcal{M}(\nu) = \{P \text{ probability measure on } \mathcal{P}(\Omega), \forall A \in \mathcal{P}(\Omega), \nu^c(A) \leq P(A) \leq \nu(A)\}$.

Remark 3 *A concave capacity that equals its conjugate is a probability measure, i.e. an additive capacity.*

Definition 4 *Let $f : \Omega \to \mathbb{R}^+$ be a L_1 bounded positive function and let ν be a capacity on $\mathcal{P}(\Omega)$, the Choquet integral of f with respect to ν is the real value $\mathbb{C}_\nu(f)$ defined by:*

$$\mathbb{C}_\nu(f) = \int_0^\infty \nu\{x \in \Omega / f(x) \geq \alpha\} d\alpha. \tag{1}$$

This definition can be easily extended to non-positive functions when considering the so-called asymmetric Choquet integral [6].

Definition 5 *Let $f : \Omega \to \mathbb{R}$ be a L_1 bounded function and let ν be a capacity on $\mathcal{P}(\Omega)$, let f^+ (resp. f^-) be the function defined by $\forall x \in \Omega$, $f^+(x) = max(f(x), 0)$ (resp. $f^-(x) = max(-f(x), 0)$), the asymmetric Choquet integral of f with respect to ν is the real value $\check{\mathbb{C}}_\nu(f)$ defined by:*

$$\check{\mathbb{C}}_\nu(f) = \mathbb{C}_\nu(f^+) - \mathbb{C}_\nu(f^-). \tag{2}$$

All of those definitions can be easily particularized to the discrete case.

Definition 6 *A discrete capacity ν is a set function ν: $\mathcal{P}(\Theta_p) \to [0, 1]$ such that $\nu(\varnothing) = 0$, $\nu(\Theta_p) = 1$, and $\nu(A) \leq \nu(B)$ for all $A \subseteq B \subseteq \Theta_p$.*

The definitions of conjugate capacity, concavity and core of a concave capacity exactly match the continuous case.

Definition 7 *Let $F : \Theta_p \to \mathbb{R}^+$ be a L_1 bounded positive function and let ν be a capacity on $\mathcal{P}(\Theta_p)$, the Choquet integral of F with respect to ν is the real*

value $\mathbb{C}_\nu(F)$ *defined by:*

$$\mathbb{C}_\nu(F) = \sum_{n \in \Theta_p} F_{\sigma(n)}(\nu(A_{\sigma(n)}) - \nu(A_{\sigma(n+1)})), \tag{3}$$

where σ *is a permutation of* Θ_p *such that* $F_{\sigma(0)} \leq \ldots \leq F_{\sigma(p)}$*, and* $A_{\sigma(n)}$ *is the subset of* Θ_p *such that* $\forall i \in A_{\sigma(n)}$*,* $F_i \geq F_{\sigma(n)}$*:* $A_{\sigma(n)} = \{\sigma(n), \ldots, \sigma(p)\}$*. By convention* $A_{\sigma(p+1)} = \varnothing$*.*

If F is not positive, then the asymmetric Choquet integral can also be defined in the discrete case.

Definition 8 *Let* $F : \Theta_p \to \mathbb{R}$ *be a* L_1 *bounded function and let* ν *be a capacity on* $\mathcal{P}(\Theta_p)$*, let* F^+ *(resp.* F^-*) be the function defined by* $\forall k \in \Theta_p$*,* $F_k^+ = max(F_k, 0)$ *(resp.* $F_k^- = max(-F_k, 0)$*), the asymmetric Choquet integral of* F *with respect to* ν *is the real value* $\check{\mathbb{C}}_\nu(F)$ *defined by:*

$$\check{\mathbb{C}}_\nu(F) = \mathbb{C}_\nu(F^+) - \mathbb{C}_\nu(F^-). \tag{4}$$

2.3 Summative and maxitive kernels

Kernels are used in signal processing to define weighted neighborhoods of real locations.

A **continuous summative kernel** [13] is a continuous function $\kappa : \Omega \longrightarrow \mathbb{R}^+$ such that $\int_\Omega \kappa(x)dx = 1$. This function is formally equivalent to the density of a Lebesgue-measurable probability distribution P_κ defined by: $\forall A \in \mathcal{P}(\Omega)$, $P_\kappa(A) = \int_A \kappa(x)dx$. $\mathcal{K}(\Omega)$ is the set of all summative kernels defined on Ω.

A **continuous maxitive kernel** [13] is a continuous function $\pi : \Omega \longrightarrow [0, 1]$ such that $\sup_{x \in \Omega} \pi(x) = 1$. This function is equivalent to the density of a possibility distribution, thus defining a possibility measure (Π_π) and a necessity measure (N_π) on Ω: $\Pi_\pi(A) = \sup_{x \in A} \pi(x)$ and $N_\pi(A) = 1 - \sup_{x \notin A} \pi(x)$. Π_π is a concave capacity and N_π its conjugate (convex) capacity.

Defining a maxitive kernel is equivalent to defining a subset of $\mathcal{K}(\Omega)$. Let π be a maxitive kernel, then $\mathcal{M}(\pi) = \{\kappa \in \mathcal{K}(\Omega) / \forall A \in \mathcal{P}(\Omega), N_\pi(A) \leq P_\kappa(A) \leq \Pi_\pi(A)\}$ is called the *core* of π. This definition coincides with that given in Section 2.2.

Those concepts can easily be extended to a discrete space [13].

A **discrete summative kernel** [13] is a discrete function $\eta : \Theta_p \longrightarrow [0, 1]$ such that $\sum_{k \in \Theta_p} \eta_k = 1$. This function defines a probability measure P_η by:

$\forall A \in \mathcal{P}(\Theta_p)$, $P_\eta(A) = \sum_{k \in A} \eta_k$. The set of all discrete summative kernels defined on Θ_p is denoted $\mathcal{K}(\Theta_p)$.

A **discrete maxitive kernel** [13] is a discrete function $\pi : \Theta_p \longrightarrow [0, 1]$ such that $\sup_{k \in \Theta_p} \pi_k = 1$. This function defines two dual confidence measures on Θ_p called a possibility measure (Π_π) and a necessity measure (N_π): $\Pi_\pi(A) = \sup_{k \in A} \pi_k$ and $N_\pi(A) = 1 - \sup_{k \notin A} \pi_k$.

Similar to the continuous case, a discrete maxitive kernel defines a convex subset of discrete summative kernels, denoted $\mathcal{M}(\pi)$ [28].

2.4 Precise and imprecise expectations

Let $f : \Omega \to \mathbb{R}$ be a L_1 bounded function and let P be an additive confidence measure (i.e. a probability). The precise expectation of f with respect to P is the real value $\mathbb{E}_P(f)$ defined by:

$$\mathbb{E}_P(f) = \int_0^\infty f dP, \tag{5}$$

with P being additive. As P is fully characterized by its density function κ through the equality: $\forall A \in \mathcal{P}(A)$, $P(A) = P_\kappa(A) = \int_A \kappa(x)dx$, Expression (5) can be rewritten as:

$$\mathbb{E}_P(f) = \mathbb{E}_{P_\kappa}(f) = \int_\Omega f(x)\kappa(x)dx. \tag{6}$$

The expectation concept can easily be extended to concave capacities (see [28]). Let ν be a concave capacity and let $f : \Omega \to \mathbb{R}$ be a L_1 bounded function. The imprecise expectation of f with respect to ν is the real interval $\overline{\underline{\mathbb{E}}}_\nu(f)$ defined by:

$$\overline{\underline{\mathbb{E}}}_\nu(f) = \left[\underline{\mathbb{E}}_\nu(f), \overline{\mathbb{E}}_\nu(f)\right] = \left[\check{\mathbb{C}}_{\nu^c}(f), \check{\mathbb{C}}_\nu(f)\right].$$

Two fundamental properties come from the work of Denneberg [6].

Proposition 9 *Let $f : \Omega \to \mathbb{R}$ be a L_1 bounded function and let ν be a capacity defined on Ω, $\forall P \in \mathcal{M}(\nu)$, $\mathbb{E}_P(f) \in \overline{\underline{\mathbb{E}}}_\nu(f)$ and $\forall y \in \overline{\underline{\mathbb{E}}}_\nu(f)$, $\exists P \in \mathcal{M}(\nu)$ such that $y = \mathbb{E}_P(f)$.*

Proposition 10 *Let $f, g : \Omega \to \mathbb{R}$ be two L_1 bounded functions and let ν be a concave capacity defined on Ω, $\overline{\mathbb{E}}_\nu(f + g) \leq \overline{\mathbb{E}}_\nu(f) + \overline{\mathbb{E}}_\nu(g)$ and $\underline{\mathbb{E}}_\nu(f + g) \geq \underline{\mathbb{E}}_\nu(f) + \underline{\mathbb{E}}_\nu(g)$.*

Precise and imprecise expectations coincide when considering a probability measure, i.e. if P is a probability measure on $\mathcal{P}(\Omega)$, then $\underline{\overline{\mathbb{E}}}_P(f) = \mathbb{E}_P(f)$.

In the same way, precise and imprecise expectations can be defined in the discrete space. Let $F : \Theta_p \to \mathbb{R}$ be a bounded function. Let P_η be a discrete probability measure on $\mathcal{P}(\Theta_p)$ generated by the discrete summative kernel η. The precise expectation of F with respect to P_η is the precise value $\mathbb{E}_{P_\eta}(F)$ defined by:

$$\mathbb{E}_{P_\eta}(F) = \sum_{k \in \Theta_p} F_k \eta_k. \tag{7}$$

Let ν be a concave capacity on $\mathcal{P}(\Theta_p)$. The imprecise expectation of F with respect to ν is the real interval $\underline{\overline{\mathbb{E}}}_\nu(F)$ defined by:

$$\underline{\overline{\mathbb{E}}}_\nu(F) = \left[\underline{\mathbb{E}}_\nu(F), \overline{\mathbb{E}}_\nu(F)\right] = \left[\check{\mathbb{C}}_{\nu^c}(F), \check{\mathbb{C}}_\nu(F)\right].$$

Properties 9 and 10 are also true in the discrete space [30].

Finally, expectation operators can be easily extended to interval valued functions. Let $\overline{\underline{f}} : \Omega \to \mathbb{IR}$ be an interval valued function, i.e. $\forall x \in \Omega$, $\overline{\underline{f}}(x) = [\underline{f}(x), \overline{f}(x)]$, let P be a probability measure,
$\mathbb{E}_P(\overline{\underline{f}}) = [\mathbb{E}_P(\underline{f}), \mathbb{E}_P(\overline{f})] = \{\mathbb{E}_P(g)/g \in \overline{\underline{f}}\}$ [7]. Let ν be a concave capacity, this extension also applies with the imprecise expectation operator $\underline{\overline{\mathbb{E}}}_\nu$:
$\underline{\overline{\mathbb{E}}}_\nu(\overline{\underline{f}}) = [\underline{\mathbb{E}}_\nu(\underline{f}), \overline{\mathbb{E}}(\overline{f})] = \{\mathbb{E}_P(g)/g \in \overline{\underline{f}} \text{ and } P \in \mathcal{M}(\nu)\}$. These extensions also apply with discrete functions (proofs see [33]).

2.5 Kernel: an instrumental tool in discrete signal processing

In signal processing, kernels are instrumental to obtain discrete operations on discrete signals (or functions) which are equivalent to continuous operations on continuous signals (or functions) [36]. Computing the derivative of a digital image is a good example of such a process that requires estimating the gradient value (which is a continuous concept) at each pixel location of the input image. The sampling process is modeled by associating a summative kernel $\kappa^k \in \mathcal{K}(\Omega)$ with each sampling location $\omega_k \in \Omega$ ($k \in \Theta_p$). The F_k value associated with the k^{th} sampling location is computed by $F_k = \int_\Omega f(x)\kappa^k(x)dx$. In most applications, each summative kernel κ^k can be deduced from a generic kernel $\kappa \in \mathcal{K}(\Omega)$ by $\forall x \in \Omega$, $\kappa^k(x) = \kappa(\omega_k - x)$ (see Figure (1)). Each sampled value F_k can thus be seen as an average value of the original signal in a weighted neighborhood of each sampling location ω_k obtained by:

$$\forall k \in \Theta_p, F_k = \int_\Omega f(x)\kappa(\omega_k - x)dx. \tag{8}$$

Hereafter, the procedure whereby the value F_k is associated with each sampling location ω_k via Expression (8) will be referred to as *sampling f via κ at the sampling location ω_k*.

From a signal processing standpoint, the sampling is said to be perfect if $F_k = f(\omega_k)$. This situation is true for any signal f if and only if the sampling kernel is δ, i.e. the fictive function associated with the Dirac distribution. Generally, F_k can be seen as the perfect sampling of g, a function obtained by smoothing f with the kernel κ: $F_k = g(\omega_k)$, with $g = f \otimes \kappa$, $\otimes$ being the convolution operation defined by: $\forall x \in \Omega$, $(f \otimes \kappa)(x) = \int_{\mathbb{R}} f(u)\kappa(x-u)du$.

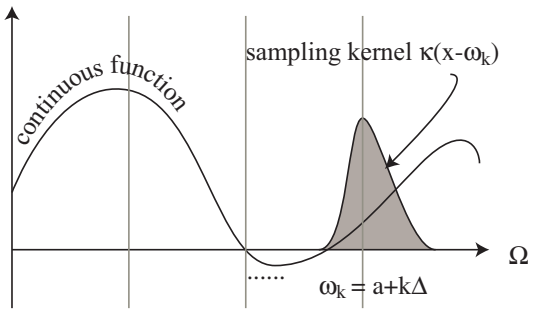

Fig. 1. Sampling.

Sampling is done to preserve the main information contained in the continuous signal in the samples, i.e. the loss of information due to sampling is negligible. It is thus possible to obtain a reliable reconstruction of f based on these samples. This reconstruction is conditioned by the Nyquist–Shannon sampling theorem: no signal whose frequency is higher than $\frac{1}{2\Delta_p}$ can be reconstructed from a sampled signal whose sampling period is Δ_p. In that case, theoretically, the value of signal f, at each $x \in \Omega$, can be reconstructed by convoluting F_k with a kernel whose impulse response is a sine cardinal function centered on x (see Figure (2)). However, this theoretical reconstruction is usually not possible since the signal has a bounded support, while the sine cardinal kernel is adapted to reconstruction of a signal whose support is not bounded. As suggested by Unser [36], it is often better and more significant to reconstruct function f by using a band limited reconstruction kernel η^x that, contrary to the sampling kernel, is not translation invariant, i.e. the shape of the reconstruction kernel depends on the position of location x where the signal has to be reconstructed. A fitting condition is used to define the reconstruction kernel for the reconstructed signal $\hat{f}$ to be as close as possible to the original signal f. Some computational complexity arguments for or against a particular kernel function can also be put forward. This reconstruction step can be written:

$$\forall x \in \Omega, \hat{f}(x) = \sum_{k \in \Theta_p} F_k \eta_k^x, \tag{9}$$

where η^x is the reconstruction kernel defined at position x.

This reconstruction step is usually called interpolation. Interpolation is a spe-

cial reconstruction case where the reconstructed continuous function has to coincide with the sampled function at the sampling locations, i.e. $\forall k \in \Theta_p$, $\hat{f}(\omega_k) = F_k$. Therefore, an interpolation kernel has to comply with $\forall k \in \Theta_p$, $\eta_k^{\omega_k} = 1$. This entails $\forall l \neq k \in \Theta_p$, $\eta_l^{\omega_k} = 0$. However, interpolating the F_k leads to reconstructing $g = f \otimes \kappa$ and not f since $F_k \neq f(\omega_k)$. Thus, reconstructing f usually requires a high-pass filtering step, that can be combined with the interpolation to generate the reconstruction kernel. In that case, the obtained reconstruction kernel is not summative since it is non-positive: a positive kernel can only induce low-pass filtering. It however sums to one.

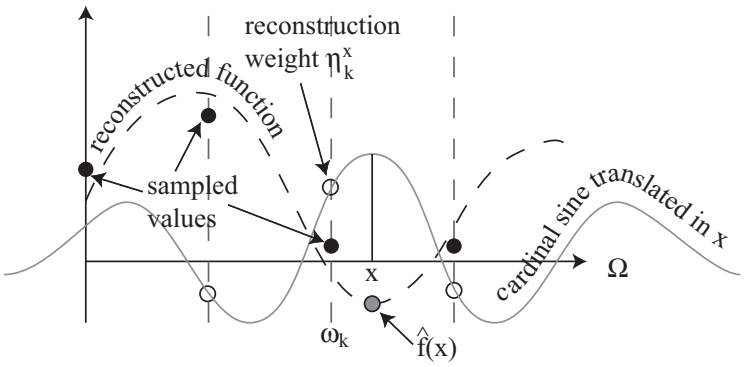

Fig. 2. Reconstruction

In a pioneer work, Unser [36] proposed B-spline kernels as a tool to represent both sampling and reconstructing kernels, thus making a convenient bridge between continuous and discrete domains. The main assets are the computation simplicity, the reconstruction optimality and the possibility of coupling the sampling and reconstructing kernels in a *perfect fit* scheme. The perfect fit scheme consists of defining a couple of summative kernels (κ, η) which enables us to define discrete operations based on continuous operations (e.g. derivation). This couple is such that, when reconstructing a continuous signal $\hat{f}$ from a sampled signal F by using η and then sampling $\hat{f}$ via κ, the obtained sampled signal $\hat{F}$ is identical to the original sampled signal F. Defining a perfect fit couple is easy within the B-spline framework. Within this approach, the real sampling kernel, i.e. the kernel associated with the imager, is unknown. Note that F-transform and perfect fit approaches have different purposes. Where the F-transform tries to simplify the processing of a signal by an appropriate sampling-reconstruction scheme, the perfect fit approach aims at processing digital signals with operations that have been defined in the continuous domain.

Regardless of the purpose, a sampling/reconstruction couple is a tool that should be able to approximate any continuous function f with arbitrary precision. For any arbitrary value $\epsilon \in \mathbb{R}$, there is always a triplet (Δ_p, κ, η) that makes the distance between f and $\hat{f}$ lower than ϵ. Within this framework, summative kernels are mostly used since they have easy interpretations in terms of weighted neighborhoods.

However, summative kernels are not the only kernels that can be used to achieve this continuous to discrete interplay. In fact, the reconstruction kernel is not positive in the perfect fit approach. It however sums to one. The fact that the kernels sum to one is usually justified by some energy conservation arguments.

When the kernels are summative, both sampling and reconstruction can be rewritten by using precise expectation operators. Indeed, since κ^k is the kernel κ translated in ω_k, the sample $F_k = \mathbb{E}_{P_{\kappa^k}}(f) = \int_\Omega f(x)\kappa(\omega_k - x)dx$. Similarly, when considering the summative kernel η^x defined in Section 2.5, $\hat{f}$, the estimation of f based on the samples can be written: $\forall x \in \Omega$, $\hat{f}(x) = \mathbb{E}_{P_{\eta^x}}(F)$.

3 Fuzzy transform

The fuzzy transform (or F-transform) was proposed and extensively studied by Irina Perfilieva [23]. This transform can be viewed as the decomposition of a continuous (or discrete) function on a fuzzy partition of its domain. We present some nice properties of this approach.

3.1 Direct F-transform

Let $f : \Omega \to \mathbb{R}$ be a function. Let $\{C_k\}_{k\in\Theta_p}$ be the $(p+1)$ atoms of a fuzzy partition *à la* Ruspini of Ω [23], [29], i.e. a set of unimodal symmetrical fuzzy intervals complying with $\forall x \in \Omega$:

- $\sum_{k\in\Theta_p} C_k(x) = 1$,
- $\exists! k \in \Theta_p, C_k(x) > 0$, $C_{k+1}(x) \geq 0$,
- $\forall k \in \Theta_p$, C_k is continuous.

The (integral) fuzzy transform of f with respect to the fuzzy partition $\{C_k\}_{k\in\Theta_p}$ is the (p+1)-tuple $\{F_k\}_{k\in\Theta_p}$ obtained by decomposing (averaging) f on each atom of the partition by:

$$\forall k \in \Theta_p, F_k = \frac{\int_\Omega f(x)C_k(x)dx}{\int_\Omega C_k(x)dx}. \tag{10}$$

In many applications, the partition upon which the F-transform is built is uniform [23]. A straightforward way to build such a uniform partition consists of defining a generic unimodal symmetrical fuzzy subset E of Ω having a bounded support $[-\Delta_p, \Delta_p]$, with $\Delta_p = \frac{b-a}{p}$. Each atom C_k of the partition is then deduced from the membership function of E by $\forall x \in \Omega$, $C_k(x) =$

$E(x-\omega_k)$, with $\omega_k = a+k\Delta_p$. For the $\{C_k\}_{k\in\Theta_p}$ to be a regular fuzzy partition, the fuzzy subset E should comply with $\forall x \in [0, \frac{\Delta_p}{2}]$, $E(x) = 1 - E(\Delta_p - x)$.

The granularity of the partition [13] is a way to characterize how the transformation approximates the original function. It is a measure of the non-specificity of the fuzzy subset E and therefore the roughness of the partition. The granularity $\gamma(E)$ of a fuzzy subset E is defined by: $\gamma(E) = \int_{\mathbb{R}} E(x)dx$.

Proposition 11 *If E is the generic unimodal symmetrical fuzzy subset defining a partition à la Ruspini with a sampling step Δ_p then $\gamma(E) = \Delta_p$.*

Proof: Due to the symmetry of the membership function of E, $\int_{\mathbb{R}} E(x)dx = \int_{-\Delta_p}^{\Delta_p} E(x)dx = 2\int_0^{\Delta_p} E(x)dx$. Though $\int_0^{\Delta_p} E(x)dx = \int_0^{\frac{\Delta_p}{2}} E(x)dx + \int_{\frac{\Delta_p}{2}}^{\Delta_p} E(x)dx = \int_0^{\frac{\Delta_p}{2}} 1 - E(\Delta_p - x)dx + \int_{\frac{\Delta_p}{2}}^{\Delta_p} E(x)dx = \int_{\frac{\Delta_p}{2}}^{\Delta_p} 1 - E(x)dx + \int_{\frac{\Delta_p}{2}}^{\Delta_p} E(x)dx = \frac{\Delta_p}{2}$ □

Thus, when considering a uniform partition [23], Expression (10) is simplified in:

$$\forall k \in \Theta_p, F_k = \frac{1}{\Delta_p}\int_\Omega f(x)C_k(x)dx. \tag{11}$$

Note that when considering the support of function f as being unbounded on Ω, F_0 and F_p are rather obtained by: $F_0 = \frac{2}{\Delta_p}\int_\Omega f(x)C_0(x)dx$ and $F_p = \frac{2}{\Delta_p}\int_\Omega f(x)C_p(x)dx$.

With the partition being fixed, the F-transform of a function can be considered as a rough representation of this function. As noted in [25], the components of an F-transform are average values of the original function. Moreover, any sequence $\{F_k\}_{k\in\Theta_p}$ could be considered as being the F-transform of an unknown function f [23]. In fact, there are very close similarities between Equation (11) and Equation (8). From a signal processing standpoint, an F-transform is simply p samples of a signal (function) based on the sampling kernel κ_E defined by $\forall x \in [-\Delta_p, \Delta_p]$, $\kappa_E(x) = E(-x)/\int_{-\Delta_p}^{\Delta_p} E(x)dx = \frac{E(x)}{\Delta_p}$, since $\gamma(E) = \int_{-\Delta_p}^{\Delta_p} E(x)dx = \Delta_p$. By construction κ_E is a summative continuous kernel since $\forall x \in \mathbb{R}$, $\kappa_E(x) \geq 0$ and $\int_{\mathbb{R}} \kappa_E(x)dx = 1$ [24]. Thus Expression (11) can be rewritten via an expectation operation:

$$\forall k \in \Theta_p, F_k = \mathbb{E}_{P_E^k}(f), \tag{12}$$

where P_E^k is the probability measure induced by the summative kernel $\kappa_E^{\omega_k}$ defined by: $\forall x \in \Omega$, $\kappa_E^{\omega_k}(x) = \kappa_E(x - \omega_k)$.

The inverse F-transform (IF-transform) aims to reconstruct the original function f from $(p+1)$ components of its F-transform. The exact reconstruction is possible with few very smooth functions since the F-transform, like any subsampling method, loses some information. Thus, the IF-transform should be perceived as a method that provides an estimate $\hat{f}$ of the original function f by using the $(p+1)$ components of F by:

$$\forall x \in \Omega, \hat{f}(x) = \sum_{k \in \Theta_p} F_k C_k(x). \tag{13}$$

This reconstruction can be viewed as an interpolation by simply replacing $C_k(x)$ in Expression (13) by η_k^x in Expression (9). By construction, η^x is a discrete summative kernel. Moreover, when considering the fuzzy partition defined in Section 3.1, Expression (13) can be easily rewritten as a linear interpolation, for each $x \in \Omega$: let $k(x) \in \mathbb{N}$ be the integer value such that $x - \Delta_p < k(x)\Delta_p + a \leq x$, thus:

$$\forall x \in \Omega, \hat{f}(x) = \alpha_x F_{k(x)} + (1 - \alpha_x) F_{k(x)+1}, \tag{14}$$

with $\alpha_x = C_{k(x)}(x)$. This interpolation involves the summative kernel η^x completely defined by the value α_x, by $\eta^x_{k(x)} = \alpha_x$, $\eta^x_{k(x)+1} = 1 - \alpha_x$, and $\eta^x_k = 0$ for any $k \in \mathbb{N}/\{k(x), k(x)+1\}$. It is linear in the discrete space Θ_p, but its linearity in the continuous space Ω depends on the linearity of the base membership function E.

Expression (14) can be rewritten as an expectation operation involving the above defined summative kernel η^x:

$$\forall x \in \Omega, \hat{f}(x) = \mathbb{E}_{P_{\eta^x}}(F). \tag{15}$$

As proved in [23] (Theorem 2), the F-transform and IF-transform combination is able to approximate the original continuous function f with arbitrary precision: $\forall \epsilon > 0$, then we have $\Theta_p \subset \mathbb{N}$ and a fuzzy partition $\{C_k\}_{k \in \Theta_p}$ of Ω that leads to a decomposition $\{F_k\}_{k \in \Theta_p}$ such that $||f - \hat{f}|| \leq \epsilon$, with $\hat{f}$ being the IF-transform of $\{F_k\}_{k \in \Theta_p}$ given by Expression (13).

Proposition 12 *Let E be a fuzzy subset generating a fuzzy partition $\{C_k\}_{k \in \Theta_p}$ (see Section 3.1). Let F be the decomposition of $f : \Omega \to \mathbb{R}$ obtained through Expression (10), then reconstructing $\hat{f}$ through Expression (13) can be expressed as convoluting f with the continuous summative kernel φ^x defined on each $x \in \Omega$ by: $\forall u \in \mathbb{R}$, $\varphi^x(u) = \frac{1}{\Delta_p} \sum_{k \in \Theta_p} C_k(x) C_k(u)$.*

Proof: The φ^x formula comes directly from the composition of Expressions (10) and (13): $\forall x \in \Omega$, $\hat{f}(x) = \sum_{k\in\Theta_p} F_k C_k(x)$, yet $F_k = \frac{1}{\Delta_p} \int_\Omega f(u)C_k(u)du$, thus $\hat{f}(x) = \int_\Omega f(u) \left(\frac{1}{\Delta_p} \sum_{k\in\Theta_p} C_k(x)C_k(u)\right) du = \int_\Omega f(u)\varphi^x(u)du$. The kernel φ^x is positive by construction. Now $\int_\Omega \varphi^x(u)du = \int_\Omega \left(\frac{1}{\Delta_p} \sum_{k\in\Theta_p} C_k(x)C_k(u)\right) du$ $= \frac{1}{\Delta_p} \sum_{k\in\Theta_p} C_k(x) \left(\int_\Omega C_k(u)du\right)$. Though, $\int_\Omega C_k(u)du = \int_\Omega E(u)du = \Delta_p$. Thus $\int_\Omega \varphi^x(u)du = \sum_{k\in\Theta_p} C_k(x)\frac{1}{\Delta_p}\Delta_p = 1$. φ^x is a summative kernel. □

The approximation of f by $\hat{f}$ completely depends on the properties of the summative kernel φ^x whose specificity depends directly on Δ_p, the granularity of the partition. Note that φ^x is not translation invariant. There is, however, a kind of stepwise translation invariance in the sense that $\forall u \in \Omega$, $\varphi^x(u) = \varphi^{x+\Delta_p}(u + \Delta_p)$.

As shown by Perfilieva, the IF-transform provided by Expression (13) is the best local least square approximation (see [23] pages 1002-1004). Other reconstructions have been proposed by Crouzet [2] and then Patané [21], leading to a global least square approximation. Although this reconstruction also leads to an interpolation-like scheme, the interpolation kernel defined in each $x \in \Omega$ is not summative because it is not positive.

3.3 Discrete F-transform

Since this article is mainly focused on discrete signals, we should mention the discrete F-transform. This discrete transform applies when the function to be transformed is only known at some discrete locations $x_1 \dots x_n$. Most digital signals (including images) are regularly sampled. However, particularly in agricultural applications [20], the case when the sampling is not uniform should be considered. The proposition of Perfilieva in [23] is to simply replace the integral in Expression (10) by a discrete sum. This definition is a rough approximation of a classical discrete-to-continuous approach.

Let $f : \Omega \to \mathbb{R}$ be a function. Let $\{C_k\}_{k\in\Theta_p}$ be the considered fuzzy partition of Ω. Let $x_1 \dots x_n$ be n locations where function f is known. By using the discrete-to-continuous approach presented in Section 2.5, it is possible to reconstruct a continuous function $\tilde{f}$ with an appropriate interpolation kernel ρ^x defined at each location x where the function has to be reconstructed:

$$\forall x \in \Omega, \tilde{f}(x) = \sum_{i=1}^{n} \rho_i^x f(x_i). \tag{16}$$

Moreover, each kernel ρ^x can be defined by a generic kernel ρ by $\forall u \in \Omega$, $\rho^x(u) = \rho(u - x)$. Thus, Expression (10) can be rewritten as:

$$\forall k \in \Theta_p, F_k = \beta_k . \int_\Omega \tilde{f}(x) C_k(x) dx = \beta_k . \sum_{i=1}^{n} \eta_i^k f(x_i), \tag{17}$$

with $\eta_i^k = \int_\Omega \rho(x_i - x) C_k(x) dx$, and β_k being a normalization factor ensuring the energy conservation, i.e. if $\forall x \in \Omega$, $f(x) = 1$ then $\forall k \in \Theta_p$, $F_k = 1$.

In the discrete F-transform proposed by [23], the Dirac impulse is used as an interpolation kernel. But this solution is not unique, and any discrete kernel of the form $\forall k, i \in \Theta_p, \eta_i^k = \int_\Omega \rho(x_i - x) C_k(x)$ can be used. This approach can be convenient when the translation invariant kernel ρ aims at representing the uncertainty (or imprecision) due to the sensor.

A dual approach can be used that consists of defining a weighted neighborhood around each sample x_i. It aims to account for the sampling. According to the F-transform framework, these weighted neighborhoods can be defined by using a non-regular fuzzy partition, i.e. a set of n continuous fuzzy subsets X_i such that $\forall x \in \Omega$, $\sum_{i=1}^{n} X_i(x) = 1$. In that case, Expression (17) becomes:

$$\forall k \in \Theta_p, F_k = \sum_{i=1}^{n} \eta_i^k f(x_i), \tag{18}$$

with $\eta_i^k = \frac{1}{\Delta_p} \int_\Omega X_i(x) C_k(x) dx$.

η^k is a summative kernel since it is positive by construction and $\sum_{i=1}^{n} \eta_i^k = \frac{1}{\Delta_p} \sum_{i=1}^{n} \int_\Omega X_i(x) C_k(x) dx = \frac{1}{\Delta_p} \int_\Omega \Big(\sum_{i=1}^{n} X_i(x) \Big) C_k(x) dx = \frac{1}{\Delta_p} \int_\Omega C_k(x) dx = 1$.

3.4 Discussion

As noted by numerous authors (see e.g. [1,18,25,26,32]), the F-transform can be considered as a transform since it allows us to work with a simplified representation of a function through a reduced number of samples. Moreover, many operations achieved in the transformed space have a relevant meaning in the original space due to the linearity of the F-transform. Therefore, the F-transform framework enables some complex combinations or analysis of functions with low computational cost due to its linearity: let F and G be the F-transforms of two functions f and g, then $\forall (\alpha, \beta) \in \mathbb{R}^2$, $\alpha F + \beta G$ is the F-transform of $(\alpha f + \beta g)$. In addition, a distance between F and G can be considered as a distance between smoothed versions of f and g. However, with this property, the functions have to be decomposed on the same partition, which prohibits the use of any method for adapting the partitioning to

only one function. Moreover, if the partition is adapted to f and g, nothing can be guaranteed about its adaption to $(\alpha f + \beta g)$.

However, the word *transform* can be somewhat misleading, and the comparison with other transforms like Laplace, Fourier, Radon or wavelet may seem irrelevant. In fact, within the Laplace, Fourier, Radon or wavelet transforms, the inverse transform is really the inverse form of the direct transform in the sense that, if F is the direct transform of f, then f is the inverse transform of F. Moreover, the aforementioned transforms are not downsampled, i.e. the number of elements of the transformed signal is usually equal (or close) to the number of elements of the original signal. Thus the F-transform is not a transform like the Laplace, Fourier, Radon or wavelet transforms: the direct transform is rather a sampling-like process, and therefore some information about the original function is lost. Working with the F-transform is equivalent to working with a smooth downsampled version of the original function. At best, it is possible to ensure that the distance between the original function f and its reconstruction $\hat{f}$ is bounded by controlling the number, shape and position of the atoms of the partition. But a partition ensuring a certain bounded error for a function cannot ensure the same bound for another function.

From a signal processing standpoint, the F-transform/IF-transform couple is equivalent to defining a couple of sampling and interpolating kernels based on the same fuzzy subset E. This seems to be the only obvious difference between the F-transform framework and classical kernel-based continuous-to-discrete interplay framework. The F-transform approach does not fulfill the *best fit* conditions of Unser (see Section 2.5), except for the linear (and crisp) membership function, in the sense that there is no orthogonality between the obtained sampling and the interpolating kernels at the sampling locations. All the nice properties of the Unser approach are thus generally lost. Meanwhile, because of the definition of the fuzzy partition, by construction there is complete identity between F_k, the k^{th} component of the F-transform, and $\hat{f}(\omega_k)$, the value of the reconstructed function at the k^{th} sampling location. Moreover, defining the reconstruction step as an interpolation process leads to a systematic bias at the sampling locations since, by construction, $\hat{f}(\omega_k) = F_k \neq f(\omega_k)$. In the sampling-interpolation framework, the sampling kernel has to be as specific as possible, while within the F-transform framework the specificity of the kernel is fixed by the granularity of the generic subset E. In fact, the ideal Dirac-sampling does not belong to the F-transform framework. Moreover, Property (12) shows that the reconstructed function is simply a smoothed version of the original function. The distance between f and $\hat{f}$ is thus completely defined by the shape of the membership function of subset E which ensures the interpolation between two sampling locations. Adapting the partition to the function to be projected reduces the possibility of combining or comparing two functions in the transformed space.

Thus a question arises: *What is the advantage of defining both sampling and interpolation kernels with a single fuzzy subset?* More precisely, what does fuzzy subset theory offer in this framework other than constraining the specificity of the kernel involved in both sampling and interpolation processes?

In this article, we propose another view of the F-transform framework that justifies the reference to the fuzzy subset theory. We construct a new interval valued F-transform based on a possibilistic interpretation of the involved fuzzy subsets.

4 Non-additive F-transform

The non-additive F-transform (NF-transform) framework we propose is based on the extension of the convolution operation proposed in [11]. Within this framework, the fuzzy subset E that generates the fuzzy partition $\{C_k\}_{k\in\Theta_p}$ is considered as a maxitive kernel [13]. Constructing the NF-transform requires some preliminary properties.

4.1 Preliminary properties

Let us consider some very constructive properties of a partition *à la* Ruspini.

Proposition 13 *Let $\{C_k\}_{k\in\Theta_p}$ be a regular fuzzy partition of Ω, as defined in Section 3.1. Let $\kappa \in \mathcal{K}(\Omega)$ and κ^x be the kernel κ translated on $x \in \Omega$: $\forall u \in \Omega$, $\kappa^x(u) = \kappa(x-u)$. The function η^x defined on each $x \in \Omega$ by $\forall k \in \Theta_p$, $\eta^x_k = \int_\Omega \kappa^x(u)C_k(u)du$ is a discrete summative kernel.*

Proof: Since $\forall u \in \Omega$, $C_k(u) \geq 0$ and $\kappa(u) \geq 0$ is positive, then η^x is positive. Thus it is enough to prove that $\sum_{k\in\Theta_p} \eta^x_k = 1$. $\sum_{k\in\Theta_p} \eta^x_k = \sum_{k\in\Theta_p} \int_\Omega \kappa^x(u)C_k(u)du = \int_\Omega \kappa^x(u)(\sum_{k\in\Theta_p} C_k(u))du = \int_\Omega \kappa^x(u)du = 1$. $\square$

Proposition 14 *Let $\{C_k\}_{k\in\Theta_p}$ be a regular fuzzy partition of Ω, as defined in Section 3.1. Let κ and κ' be two summative kernels of Ω defining two reconstruction kernels η^x and η'^x as defined in Proposition 13. $\forall x \in \Omega$, $\eta^x = \eta'^x$ is equivalent to $\kappa = \kappa'$.*

Proof: It is sufficient to note that $\forall x \in \Omega$, $\eta^x = \eta'^x$ is equivalent to writing $\forall k \in \Theta_p$, $\int_\Omega(\kappa(u-x)-\kappa'(u-x))C_k(u)du = 0$, i.e. $\forall x \in \Omega$, $\int_\Omega(\kappa(u)-\kappa'(u))E(u+x)du = (h \otimes g)(x) = 0$, with $\forall u \in \Omega$, $h(u) = \kappa(u) - \kappa'(u)$ and $g(u) = E(u)$. By construction, functions h and g have a bounded support and g is positive. Let $\mathcal{F}$ be the Fourier transform. The preceding condition can be rewritten $\forall s$, $\mathcal{F}\{h\}(s)\mathcal{F}\{g\}(s) = 0$ (with s being the frequency). Since g is bounded and

positive, $\mathcal{F}\{g\}$ is null for only sparse values of s. Thus, $\mathcal{F}\{h\}$ can be non-null only at these sparse frequency values. Therefore, h cannot have a bounded support, which is a contradiction. Then h is null and thus $\kappa = \kappa'$.
The reverse implication is straightforward. $\square$

The summative discrete interpolation kernel used in the ordinary IF-transform is a particular case of the proposed construction. In that case, the involved summative kernel is the Dirac impulse δ. Thus $\forall x \in \Omega$, $\forall k \in \Theta_p$, $\eta_k^x = \int_\Omega \delta^x(u) C_k(u) du = C_k(x)$.

Proposition 15 *A discrete summative reconstruction kernel obtained by convoluting a partition à la Ruspini with a continuous summative kernel is Δ_p translation invariant, i.e. $\forall x \in \Omega$, $\forall k \in \Theta_p$ $\eta_{k+1}^{x+\Delta_p} = \eta_k^x$.*

Proof: This property is straightforward since $\eta_{k+1}^{x+\Delta_p} = \int_\Omega \kappa^{x+\Delta_p}(u) C_{k+1}(u) du = \int_\Omega \kappa(u - x - \Delta_p) E(u - \omega_{k+1}) du = \int_\Omega \kappa(u - x - \Delta_p) E(u - \omega_k - \Delta_p) du = \int_\Omega \kappa(u - x) E(u - \omega_k) du = \eta_k^x$. $\square$

Such kernels can be expressed with the expectation operator defined in Section 2.4. Indeed, $\eta_k^x = \int_\Omega \kappa^x(u) C_k(u) du = \mathbb{E}_{P_{\kappa^x}}(C_k)$, which coincides with the extension proposed in [37] for defining a probability measure of a fuzzy subset: $\eta_k^x = P_{\kappa^x}(C_k)$. This summative kernel induces a probability measure P_{η^x} defined by $\forall A \subseteq \Theta_p$, $P_{\eta^x}(A) = \sum_{k \in A} \eta_k^x = \sum_{k \in A} \int_\Omega \kappa^x(u) C_k(u) du = \int_\Omega \kappa^x(u) \sum_{k \in A} C_k(u) du = P_{\kappa^x}(\bigcup_{k \in A} C_k)$, with the union being defined using the Łukasievicz T-conorm. Therefore, the reconstruction of the continuous function within the precise reconstruction approach can be seen as an additive aggregation of values associated with each atom of the partition, with the weights being defined by the probability of each atom to belong to the neighborhood defined by the summative kernel κ^x around each location $x \in \Omega$.

4.2 Direct non-additive F-transform

Let E be a symmetrical fuzzy subset used to generate a fuzzy partition as defined in Section 3.1: $C_k(x) = E(x - \omega_k)$. The direct non-additive F-transform (NF-transform) is defined by:

$$\overline{\underline{F}}_k = \left[\underline{F}_k, \overline{F}_k\right] = \overline{\underline{\mathbb{E}}}_{\Pi_{\pi^k}}(f) = \left[\underline{\mathbb{E}}_{\Pi_{\pi^k}}(f), \overline{\mathbb{E}}_{\Pi_{\pi^k}}(f)\right] = \left[\check{\mathbb{C}}_{N_{\pi^k}}(f), \check{\mathbb{C}}_{\Pi_{\pi^k}}(f)\right], \quad (19)$$

with $\forall u \in \Omega$, $\pi(u) = E(u)$, $\pi^k(u) = C_k(u)$ and Π_{π^k} (rsp. N_{π^k}) is the possibility (rsp. necessity) measure based on the possibility distribution π^k .

Proposition 16 *Let $f : \Omega \to \mathbb{R}$ be a continuous function, and $\overline{\underline{F}}$ its NF-transform (Equation (19)), then $\forall \kappa \in \mathcal{M}(\pi)$, the discrete function F obtained*

by sampling f via κ (Equation (8)) is included in $\underline{\overline{F}}$, and $\forall k \in \Theta_p$, $\forall y \in \underline{\overline{F}}_k$, $\exists \kappa \in \mathcal{M}(\pi)$ such that $y = F_k$, with F being obtained by sampling f via κ.

Proof: It is sufficient to note that, since κ is a summative kernel, then sampling f via κ at the sampling location ω_k can be written $F_k = \mathbb{E}_{P_{\kappa^k}}(f)$, with κ^k being the summative kernel κ translated in ω_k and P_{κ^k} being the probability measure induced by the summative kernel κ^k. Due to Property 9, then $\forall \kappa \in \mathcal{M}(\pi)$, $\mathbb{E}_{P_{\kappa^k}}(f) \in \overline{\mathbb{E}}_{\Pi_{\pi^k}}(f)$ and $\forall y \in \overline{\mathbb{E}}_{\Pi_{\pi^k}}(f)$, $\exists \kappa \in \mathcal{M}(\pi)$ such that $y = \mathbb{E}_{P_{\kappa^k}}(f)$. $\square$

Proposition 17 *Let π be a maxitive kernel, let $f, g : \Omega \to \mathbb{R}$ be two continuous functions, and $\underline{\overline{F}}$ and $\underline{\overline{G}}$ be the two interval-valued functions obtained by the NF-transform of f and g based on the maxitive kernel π (Equation (19)), then the Minkowski addition of $\underline{\overline{F}}$ and $\underline{\overline{G}}$ contains the NF-transform of $f + g$ based on the maxitive kernel π.*

This property can be seen as a kind of extension, to the NF-transform, of the linearity of the F-transform.

Proof: Proving Proposition 17 is straightforward. In fact, $\forall k \in \Theta_p$, the Minkowski addition $\oplus$ is defined by: $\underline{\overline{F}}_k \oplus \underline{\overline{G}}_k = [\underline{F}_k + \underline{G}_k, \overline{F}_k + \overline{G}_k]$. Due to Proposition 10, $\underline{F}_k + \underline{G}_k = \underline{\mathbb{E}}_{\Pi_{\pi^k}}(f) + \underline{\mathbb{E}}_{\Pi_{\pi^k}}(g) \leq \underline{\mathbb{E}}_{\Pi_{\pi^k}}(f + g)$ and $\overline{F}_k + \overline{G}_k = \overline{\mathbb{E}}_{\Pi_{\pi^k}}(f) + \overline{\mathbb{E}}_{\Pi_{\pi^k}}(g) \geq \overline{\mathbb{E}}_{\Pi_{\pi^k}}(f + g)$. $\square$

Remark 18 *It has been shown in [13,8] that the triangular maxitive kernel defined on $[-\Delta_p, \Delta_p]$ is the most specific maxitive kernel whose core contains every symmetric unimodal summative kernel whose support is $[-d, d]$ with $d \leq \Delta_p$. This property combined with Proposition 16 implies that a direct NF-transform of a function $f : \Omega \to \mathbb{R}$ on the fuzzy partition of Ω generated by a triangular fuzzy subset whose support is $[-\Delta_p, \Delta_p]$ at each sampling location ω_k ($k \in \Theta_p$) contains all discrete functions obtained by sampling f on each sampling position ω_k ($k \in \Theta_p$) with a symmetric sampling kernel whose support is $[-d, d]$ with $d \leq \Delta_p$. In addition, note that the triangular maxitive kernel is also the only non-crisp kernel for which the Unser approach coincides with the FT approach.*

Remark 19 *A direct consequence of Remark 18 is that any direct F-transform obtained by using Equation (12) and a basic kernel E having $[-\Delta_p, \Delta_p]$ for support is included in the direct NF-transform obtained by using Equation (19) and a triangular maxitive kernel.*

4.3 Inverse non-additive F-transform

Ideally, this inverse non-additive F-transform (INF-transform) should comply with one of these three possible definitions:

(1) an interval-valued inverse transform that provides a convex set of all continuous functions whose precise F-transform belongs to the interval valued NF-transform, $\underline{\overline{F}}$.
(2) an interval-valued inverse transform that provides a convex set of all continuous functions whose precise F-transform is F (knowing the fuzzy partition),
(3) an interval-valued inverse transform that provides a convex set of all continuous functions whose sampling is F (without knowing the sampling kernel).

Unfortunately, finding an inverse form with a non-additive based approach is not easier than finding an inverse form with an additive based approach. Moreover, finding an INF-transform complying with definition (2) is quite impossible since high frequency information is lost in the precise sampling process that cannot be retrieved from the sampled signal. Our current proposition is to build an INF-transform that reconstructs an interval valued continuous function that contains a set of precise valued continuous functions obtained by reconstructing the discrete function with an appropriate set of kernels.

Let us first introduce, for any subset $A \subseteq \Theta_p$, the shortcut notation $\Upsilon_A = \bigcup_{k\in A} C_k$, with the union being defined using the Łukasievicz T-conorm: $\forall u \in \Omega$, $\Upsilon_A(u) = \min(1, \sum_{k\in A} C_k(u)) = \sum_{k\in A} C_k(u)$ (since $\{C_k\}_{k\in\Theta_p}$ makes a partition *à la* Ruspini). The following proposition is required to build this INF-transform.

Proposition 20 *Let $\{C_k\}_{k\in\Theta_p}$ be a regular fuzzy partition of Ω and ν be a concave continuous capacity dominating a probability measure P, then the discrete capacity υ defined by: $\forall A \subseteq \Theta_p$, $\upsilon(A) = \overline{\mathbb{E}}_\nu(\Upsilon_A)$ is concave and dominates the discrete probability Q measure defined by: $\forall A \subseteq \Theta_p$, $Q(A) = \mathbb{E}_P(\Upsilon_A)$.*

Proof: The dominance part of the proposition is easy to prove using the domination Property 9. In fact, since P is dominated by ν, regardless of the real function f, then $\mathbb{E}_P(f) \leq \overline{\mathbb{E}}_\nu(f)$. Now, let A, B be two subsets of Θ_p. With $\{C_k\}_{k\in\Theta_p}$ being a fuzzy partition *à la* Ruspini, then $\forall u \in \Omega$, $\Upsilon_{A\cup B}(u) = \Upsilon_A(u) + \Upsilon_B(u) - \Upsilon_{A\cap B}(u)$. Due to Property 10, $\overline{\mathbb{E}}_\nu(\Upsilon_{A\cup B}) \leq \overline{\mathbb{E}}_\nu(\Upsilon_A) + \overline{\mathbb{E}}_\nu(\Upsilon_B) - \overline{\mathbb{E}}_\nu(\Upsilon_{A\cap B})$, with $\cap$ being the Łukasievicz T-norm. Thus, $\upsilon(A\cup B) + \upsilon(A\cap B) \leq \upsilon(A) + \upsilon(B)$. □

A straightforward consequence of Proposition 20 is that any discrete summative reconstruction kernel constructed by convoluting atoms of the partition

$\{C_k\}_{k\in\Theta_p}$ with a continuous summative kernel (see Proposition 13) belonging to $\mathcal{M}(\nu)$, belongs to $\mathcal{M}(\upsilon)$. Moreover, due to Proposition 9, any discrete summative reconstruction kernel belonging to $\mathcal{M}(\upsilon)$ can be obtained by convoluting a continuous summative kernel belonging to $\mathcal{M}(\nu)$ with the atoms of the partition $\{C_k\}_{k\in\Theta_p}$.

The INF-transform we propose is based on Proposition 20.

Let $F:\Theta_p\to\mathbb{R}$ be a discrete bounded function. Let $\{C_k\}_{k\in\Theta_p}$ be a regular fuzzy partition of Ω, as defined in Section 3.1. Let π^x be a maxitive continuous kernel defined in each $x\in\Omega$. Let υ^x be the discrete capacity defined by:

$$\forall A\subseteq\Theta_p, \upsilon^x(A)=\overline{\mathbb{E}}_{\Pi_{\pi^x}}(\Upsilon_A). \tag{20}$$

The INF-transform is defined by:

$$\forall x\in\Omega, \underline{\overline{f}}(x)=\underline{\overline{\mathbb{E}}}_{\upsilon^x}(F)=\left[\check{\mathbb{C}}_{(\upsilon^x)^c}(F),\check{\mathbb{C}}_{\upsilon^x}(F)\right]. \tag{21}$$

Proposition 21 *Let $F:\Theta_p\to\mathbb{R}$ be a discrete bounded function. Let $\{C_k\}_{k\in\Theta_p}$ be a regular fuzzy partition of Ω, as defined in Section 3.1. Let π^x be a maxitive continuous kernel defined in each $x\in\Omega$ and υ^x be the discrete capacity defined by Expression (20). Let $\underline{\overline{f}}$ be the interval-valued function reconstructed according to Equation (21). Reconstructing F with a summative discrete kernel $\eta^x\in\mathcal{M}(\upsilon^x)$ by Equation (13) leads to a reconstructed value $\hat{f}(x)\in\underline{\overline{f}}(x)$. In the same way, $\forall y\in\underline{\overline{f}}(x)$, $\exists\eta^x\in\mathcal{M}(\upsilon^x)$ such that $y=\hat{f}(x)$, with $\hat{f}(x)$ being obtained by reconstructing F with η^x according to Expression (13).*

Proof: Proving Proposition 21 is straightforward. It is sufficient to note that $\hat{f}(x)=\mathbb{E}_{\eta^x}(F)$ and $\underline{\overline{f}}(x)=\underline{\overline{\mathbb{E}}}_{\upsilon^x}(F)$ and to use Property 9. □

The INF-transform can be easily extended to an interval-valued function $\underline{\overline{F}}$. The INF-transform of $\underline{\overline{F}}$ is defined by:

$$\forall x\in\Omega, \underline{\overline{f}}(x)=\underline{\overline{\mathbb{E}}}_{\upsilon^x}(\underline{\overline{F}})=\left[\check{\mathbb{C}}_{(\upsilon^x)^c}(\underline{F}),\check{\mathbb{C}}_{\upsilon^x}(\overline{F})\right]. \tag{22}$$

Proposition 21 also holds with this last extension for the same reasons. In fact, $\underline{\overline{\mathbb{E}}}_{\upsilon^x}(\underline{\overline{F}})$ is the set of all values that can be reconstructed at location $x\in\Omega$ from a discrete function $G\in\underline{\overline{F}}$ with a discrete kernel $\eta^x\in\mathcal{M}(\upsilon^x)$: $\underline{\overline{\mathbb{E}}}_{\upsilon^x}(\underline{\overline{F}})=\left\{y=\mathbb{E}_{\eta^x}(G)/G\in\underline{\overline{F}},\eta^x\in\mathcal{M}(\upsilon^x)\right\}$.

The non-additive version of the inverse F-transform is thus instrumental in applications involving guaranteed calculus [10] since it allows reconstruction of a continuous signal with its samples when the appropriate reconstruction kernel is imprecisely known. To ensure this, it is sufficient to use a maxitive

kernel that defines an appropriate family of reconstruction kernels. The domination guarantees the inclusion of the desired function into the interval-valued reconstructed function. Note that Proposition 17 can be easily extended to to the INF-transform, which can be useful when linear operations are involved. Once again, the triangular maxitive kernel should play a particular role in this reconstruction scheme.

The usual IF-transform is a special case of the proposed INF-transform when π^x is a Krœnecker impulse translated in x, i.e. $\pi^x(x) = 1$ and $\forall u \in \Omega$, $u \neq x$, $\pi^x(u) = 0$.

4.4 Discrete non-additive F-transform

As in Section 3.3, it is possible to account for the fact that the function to be transformed is only known at some discrete locations $x_1 \dots x_n$. Instead of constructing a function $\tilde{f}$ by interpolating the discrete valued function, this framework allows us to account for the fact that the appropriate interpolation function is unknown, and thus leads to an interval-valued F-transform that also accounts for this ill-knowledge.

Let $f : \Omega \to \mathbb{R}$ be a function. Let $\{C_k\}_{k\in\Theta_p}$ be the considered fuzzy partition of Ω. Let $\Phi = \{1, \dots, n\}$. Let $X_1 \dots X_n$ be n be the ill known (fuzzy) locations where function f is known such that $\sum_{i\in\Phi} X_i(u) \geq 1$, then the interval-valued discrete F-transform is defined by:

$$\forall k \in \Theta_p, \overline{F}_k = \overline{\mathbb{E}}_{\upsilon^k}(f) = \left[\check{\mathbb{C}}_{(\upsilon^k)^c}(f), \check{\mathbb{C}}_{\upsilon^k}(f)\right], \tag{23}$$

with $\forall A \subseteq \Phi$, $\upsilon^k(A) = \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\Upsilon_A)$, $\Upsilon_A = \bigcup_{i\in A} X_i$ ($\cup$ being the Łukasievicz union) and $\forall u \in \Omega$, $\pi^k(u) = C_k(u)$.

Proposition 22 *$\forall k \in \Theta_p$, the function υ^k defined by $\forall A \in \Phi$, $\upsilon^k(A) = \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\Upsilon_A)$, $\Upsilon_A = \bigcup_{i\in A} X_i$ and $\forall u \in \Omega$, $\pi^k(u) = C_k(u)$, is a concave capacity.*

Proof: The proof is very close to the proof of Property 20.
$\upsilon^k(\Phi) = \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\Upsilon_\Phi) = \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\mathbb{1}_\Omega) = 1$.
$\upsilon^k(\varnothing) = \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\varnothing) = 0$. $\forall u \in \Omega$, $\forall A, B \subseteq \Phi$,
$\min(\sum_{i\in A\cup B} X_i(u), 1) \leq \min(\sum_{i\in A} X_i(u), 1) + \min(\sum_{i\in B} X_i(u), 1) - \min(\sum_{i\in A\cap B} X_i(u), 1)$.
Thus, $\Upsilon_{A\cup B}(u) \leq \Upsilon_A(u) + \Upsilon_B(u) - \Upsilon_{A\cap B}(u)$. Due to Property 10,
$\overline{\mathbb{E}}_{\Pi_{\pi^k}}(\Upsilon_{A\cup B}) \leq \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\Upsilon_A) + \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\Upsilon_B) - \overline{\mathbb{E}}_{\Pi_{\pi^k}}(\Upsilon_{A\cap B})$ and thus $\upsilon^k(A \cup B) + \upsilon^k(A \cap B) \leq \upsilon^k(A) + \upsilon^k(B)$. □

Proposition 23 *Let $\rho \in \mathcal{K}(\Omega)$ be a continuous summative kernel used for interpolating a continuous function $\tilde{f}$, as presented in Expression (16). Let*

$\{C_k\}_{k\in\Theta_p}$ be a fuzzy partition of Ω. Let F be the discrete F-transform computed by using Expression (17). Let π be a continuous maxitive kernel on Ω defining, at each discrete location x_i, a fuzzy subset X_i by $\forall u \in \Omega$, $X_i(u) = \pi(x_i - u)$. Let $\underline{\overline{F}}$ be the discrete NF-transform computed by using Expression (23). If $\rho \in \mathcal{M}(\pi)$, then $F \in \underline{\overline{F}}$.

This property is easy to prove by using, once again, the domination Property 9. There are two interesting particular cases. First, if ρ is the Dirac impulse, then Expression (17) coincides with Expression (21) of [23]. In this first case, any possibility function can be used to ensure Property 14. Second, if $\rho = \frac{\pi}{\Gamma(\pi)}$ ($\Gamma(\pi) = \int_\Omega \pi(u)du$) then, by construction, $\rho \in \mathcal{M}(\pi)$ (see [4]).

Finally, it is also easy to include the fact that the discrete values of f are quantified in this representation. The simplest strategy is to account for this quantification by replacing the precise values $f(x_i)$ in Equation (23) by imprecise values.

4.5 Convergence of the non-additive F-transform

As pointed out by Perfilieva and most of the authors [23,21,3,1,32], a function cannot be reconstructed by using its F-transform, since transforming consists of sampling, and sampling leads to information loss. However, the reconstructed function can approximate its original continuous function with an arbitrary precision. The proof of this convergence is based on the ordinary proof of convergence of a band limited sampling-reconstruction scheme.

Proposition 24 *Let $f : \Omega \to \mathbb{R}$ be a uniformly continuous bounded function. Let E be a symmetric normalized fuzzy subset on Ω such that $\forall x \notin [-1, 1]$, $E(x) = 0$ and $\forall x \in [0, \frac{1}{2}]$, $E(x) = 1 - E(\frac{1}{2} - x)$. For any $p \in \mathbb{N}$, let $\underline{\overline{F}}^p$ be the NF-transform of f (Expression (19)), with $\forall x \in \Omega$, $\pi^k(x) = C_k(x) = E(\frac{x-\omega_k}{\Delta_p})$, $\Delta_p = \frac{b-a}{p}$ and $\forall k \in \Theta_p$, $\omega_k = a + k\Delta_p$.*
Let μ be a symmetric maxitive kernel on Ω such that $\forall u \notin [-1, 1]$, $\mu(u) = 0$ and μ^x be the maxitive kernel defined for any $x \in \Omega$ by : $\forall u \in \Omega$, $\mu^x(u) = \mu(\frac{u-x}{\Delta_p})$. Let υ^x be the discrete concave capacity defined by : $\forall A \subseteq \Theta_p$, $\upsilon^x(A) = \overline{\mathbb{E}}_{\Pi_{\mu^x}}(\Upsilon_A)$. Let $\overline{\underline{f}}^p$ be the INF-transform of $\overline{\underline{F}}_p$ (Expression (22)). Then for any $\epsilon > 0$, there is p_ϵ such that: $\max\left(|\overline{f}^{p_\epsilon}(x) - f(x)|, |\underline{f}^{p_\epsilon}(x) - f(x)|\right) < \epsilon$.

Proof: First, let us define the operator τ by: $\forall p \in \mathbb{N}, \forall i \in \mathbb{Z}$, $\tau(i, p) = \max(0, \min(i, p))$. Now, let us note that, due to its construction, the discrete capacity υ^x dominates any discrete summative kernel obtained by convoluting the partition $\{C_k\}_{k\in\Theta_p}$ with any continuous kernel κ belonging to $\mathcal{M}(\pi^x)$. Since the support of π^x is bounded, any summative kernel dominated by υ^x is bounded: we have n such that $\forall \eta \in \mathcal{M}(\upsilon^x)$, $\forall x \in \Omega, \forall i \notin$

$[\tau(k(x)-n,p),\tau(k(x)+n,p)]$, $\eta_i = 0$.

Note also that, since E is a normalized fuzzy subset, its membership function can be interpreted as a possibility distribution π_E inducing a possibility measure on Ω.

Since f is uniformly continuous on Ω, then for any $\epsilon > 0$, there is $\delta > 0$ such that $\forall u, v \in \Omega$, $|u - v| < \delta$ implies $|f(u) - f(v)| < \epsilon$. Let us choose a value of p inducing regularly spaced values $(\omega_k)_{k\in\Theta_p}$ such that $\forall k \in \Theta_p$, $\forall u, v \in [\omega_{\tau(k-n,p)}, \omega_{\tau(k+n,p)}]$, $|f(u) - f(v)| < \epsilon$.

Thus, $\forall x \in \Omega$, $\forall i \in \{\tau(k(x)-n,p),\ldots,\tau(k(x)+n,p)\}$, $\forall \kappa \in \mathcal{M}(\pi_E)$
$|f(x)-F_i^p| = |f(x)-\int_\Omega \frac{1}{\Delta_p} f(\omega)\kappa(\frac{\omega-\omega_i}{\Delta_p})d\omega| \leq \int_\Omega \frac{1}{\Delta_p}|f(x)-f(\omega)|\kappa(\frac{\omega-\omega_i}{\Delta_p})d\omega \leq \epsilon$.
Due to Propositions 9 and 10, $\exists \kappa_1, \kappa_2 \in \mathcal{M}(\pi_E)$ such that $\underline{F}_i^p = \int_\Omega \frac{1}{\Delta_p} f(\omega)\kappa_1(\frac{\omega-\omega_i}{\Delta_p})d\omega$ and $\overline{F}_i^p = \int_\Omega \frac{1}{\Delta_p} f(\omega)\kappa_2(\frac{\omega-\omega_i}{\Delta_p})d\omega$. Therefore $|f(x)-\underline{F}_i^p| < \epsilon$ and $|f(x)-\overline{F}_i^p| < \epsilon$

Now, $\forall x \in \Omega$, due to Propositions 9 and 10, $\exists \eta \in \mathcal{M}(\upsilon^x)$ such that $\overline{f}^p(x) = \sum_{i=0}^{i=p} \overline{F}_i^p \eta_i$,
thus $|\overline{f}^p(x) - f(x)| = |\sum_{i=0}^p \overline{F}_k^p \eta_i - \sum_{i=0}^p f(x)\eta_i| \leq \sum_{i=0}^p |\overline{F}_i^p - f(x)|\eta_i$.
Since $\eta_i = 0$ if $i \notin [\tau(k(x)-n,p),\tau(k(x)+n,p)]$,
$|\overline{f}^p(x) - f(x)| \leq \sum_{i=\tau(k(x)-n,p)}^{\tau(k(x)+n,p)} |\overline{F}_i^p - f(x)|\eta_i \leq \epsilon \sum_{i=\tau(k(x)-n,p)}^{\tau(k(x)+n,p)} \eta_i = \epsilon$. The same scheme can be used to prove that $|\underline{f}(x) - f(x)| \leq \epsilon$. □

4.6 A simple example with crisp partitioning

A very easy way to understand the proposed interval-valued transform is to consider a crisp partition, i.e. a partition generated by the crisp subset $E = [-\frac{\Delta_p}{2}, \frac{\Delta_p}{2}]$. In that case, Equation (19) simply leads to computing the upper and lower values of f within the crisp subset C_k: $\overline{F}_k = \left[\underline{F}_k, \overline{F}_k\right] = \left[\inf_{u\in C_k} f(u), \sup_{u\in C_k} f(u)\right]$. Note that the F-transform based on the residuated lattice proposed by Perfilieva gives exactly the same interval in that binary case. When using a smoother maxitive kernel, the interval-valued F-transform is more specific in both cases.

Such a decomposition on a crisp partition can also be seen as a granulation process, related to the Pawlak rough set framework [22] where the sets of the partition create an equivalence relation on Ω. Thus, replacing the crisp partition by a fuzzy partition leads to considering a weighted granulation as proposed in [5]. Each atom of the partition can be thought of as a granule on which the value of function f is imprecisely known.

In the same way, Equation (21) leads to computing, at each $x \in \Omega$, the upper

and lower values of F in a discrete neighborhood equal to $\{k(x), k(x)+1\}$, where $k(x) \in \mathbb{N}$ is the integer value such that $x - \Delta_p < k(x)\Delta_p + a \leq x$. Therefore, in that case, since $\forall x \in \Omega$, $f(x) \in \overline{\underline{F}}_{n(x)}$ (with $n(x)$ being the integer value such that $x \in C_{n(x)}$), $f(x)$ also belongs to $[\inf_{k\in\{k(x),k(x)+1\}}, \sup_{k\in\{k(x),k(x)+1\}}]$ since either $n(x) = k(x)$ or $n(x) = k(x)+1$. Naturally, the inverse transform would have been more precise if the interval $\{k(x), k(x)+1\}$ had been replaced by the singleton $n(x)$. This situation corresponds to using a possibility density π^x as the Krœnecker impulse translated in x, i.e. using, as an inverse transform, the usual inverse F-transform. When using a smoother maxitive kernel, the interval-valued F-transform is smoother within the NF-transform framework.

5 Experiments

This experimental section aims to illustrate some properties of the NF-transform and show what this approach brings to the classical approach. The experiments are carried out on digital images, i.e. bidimensional discrete signals. We thus need to define how to extend both the F-transform and NF-transform in two dimensions.

5.1 Bidimensional extension

Extending the F-transform in two dimensions is rather straightforward. In fact, since F-transforming a unidimensional signal consists of convoluting this signal with a unidimensional summative kernel, F-transforming a bidimensional signal consists of convoluting this signal with a bidimensional summative kernel.

Let $\Omega = [a, b] \times [c, d]$ be a box of $\mathbb{R}^2$. Let $I : (x, y) \in \Omega \to I(x, y) \in \mathbb{R}$ be an image on Ω. Let $p_x, p_y \in \mathbb{N}$ be two positive intergers. Let $\Delta_x = \frac{b-a}{p_x}$ and $\Delta_y = \frac{d-c}{p_y}$. Let E be a fuzzy subset of Ω complying with $\forall x, y \in [0, \frac{\Delta_x}{2}] \times [0, \frac{\Delta_y}{2}]$, $E(x, y) = 1 - E(\Delta_x - x, \Delta_y - y)$, and 0 elsewhere. Let $\Theta = [0, p_x] \times [0, p_y]$ be be a box of $\mathbb{N}^2$. Let $\{C_{k_x,k_y}\}_{(k_x,k_y)\in\Theta}$ be $(p_x + 1).(p_y + 1)$ fuzzy subsets of Ω defined by $\forall (x, y) \in \Omega$, $C_{k_x,k_y}(x, y) = E(x - a - k_x.\Delta_x, y - c - k_y.\Delta_y)$. By construction, the $\{C_{k_x,k_y}\}_{(k_x,k_y)\in\Theta}$ form a partition *à la Ruspini* of Ω.

Then, Expression (10) becomes:

$$\forall (k_x, k_y) \in \Theta, F_{k_x,k_y} = \frac{\int_\Omega I(x, y)C_{k_x,k_y}(x, y)dxdy}{\int_\Omega C_{k_x,k_y}(x, y)dxdy}. \tag{24}$$

In the same way, Expression (13) becomes:

$$\forall (x,y) \in \Omega, \hat{I}(x,y) = \sum_{(k_x,k_y)\in\Theta} F_{k_x,k_y} C_{k_x,k_y}(x,y). \tag{25}$$

Extending the NF-transform in two dimensions is as straightforward as extending the F-transform in two dimensions. It follows the same process and involves bidimensional maxitive kernels.

Expression (19) becomes:

$$\forall (k_x,k_y) \in \Theta, \overline{F}_{k_x,k_y} = \overline{\mathbb{E}}_{\Pi_{\pi^{k_x,k_y}}}(f) = \left[\underline{\mathbb{E}}_{\Pi_{\pi^{k_x,k_y}}}(f), \overline{\mathbb{E}}_{\Pi_{\pi^{k_x,k_y}}}(f)\right] \tag{26}$$

with $\forall (x,y) \in \Omega$, $\pi(x,y) = E(x,y)$, $\pi^{k_x,k_y}(x,y) = C_{k_x,k_y}(x,y)$ and $\Pi_{\pi^{k_x,k_y}}$ (rsp. $N_{\pi^{k_x,k_y}}$) is the possibility (rsp. necessity) measure based on the possibility distribution π^{k_x,k_y} .

In the same way, Expression (21) becomes:

$$\forall (x,y) \in \Omega, \overline{I}(x,y) = \overline{\mathbb{E}}_{\upsilon^{x,y}}(\overline{F}) = \left[\check{\mathbb{C}}_{(\upsilon^{x,y})^c}(\underline{F}), \check{\mathbb{C}}_{\upsilon^{x,y}}(\overline{F})\right], \tag{27}$$

$\upsilon^{x,y}$ being the discrete capacity defined by: $\forall A \subseteq \Theta, \upsilon^{x,y}(A) = \overline{\mathbb{E}}_{\Pi_{\pi^{x,y}}}(\Upsilon_A)$, with Υ_A being the continuous fuzzy subset defined by $\forall (x,y) \in \Omega$, $\Upsilon_A(x,y) = \sum_{(k_x,k_y)\in A} C_{k_x,k_y}(x,y)$.

Extending the discrete transforms proceeds in the same way. The granularity of the partition can be seen either as the couple (Δ_x, Δ_y) or as the value $\Delta_x * \Delta_y = \int_\Omega E(x,y)dxdy$.

Image processing generally considers separable kernels. Expressions (24), (25), (26) and (27) can be highly simplified by considering E as separable. E is said to be separable if $\forall (x,y) \in \Omega$, $E(x,y) = E_x(x).E_y(y)$. In that case, computing a bidimensional transform turns into computing two consecutive unidimensional transforms. The same applies for inverse transforms.

In the rest of the experimental section, we consider centered linear symmetric kernels (i.e. triangular kernels).

5.2 Illustrating the dominations

In this section, we propose to illustrate different behaviors and properties of the NF-transform, e.g. Properties 16 and 21. This illustration is based on the high resolution discrete image depicted in Figure (3). The experiment consists of downsampling then upsampling this image with a linear (pyramidal) fuzzy

partition (see Remark 18) whose granularity equals $\Delta = 16$ (4×4) and with different sampling and reconstruction kernels complying with the domination property (i.e. sampling (reconstruction) kernels belonging to the core of the sampling (reconstruction) capacity generated by the partition).

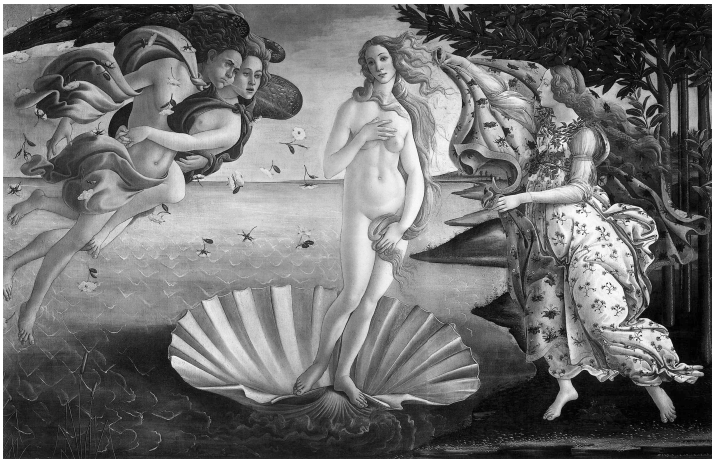

Fig. 3. 1560×2435 image of the painting *Nascita di Venere* by Sandro Botticelli.

Figure (4) shows the lower (a) and upper (b) downsampled image obtained using the NF-transform. Figure (5-a) shows the downsampled image obtained using the F-transform while Figure (5-b) shows the median of the interval-valued image obtained using the NF-transform. Figure (6-a) shows the image obtained by reconstructing the downsampled image (5-a) using the IF-transform while Figure (6-b) shows the median of the interval-valued downsampled image whose median is drawn in Figure (5-b) reconstructed using the INF-transform.


a)

b)

Fig. 4. Lower (a) and upper (a) images obtained by downsampling the original image using the NF-transform.

As a first remark, the image sampled using the F-transform (Figure (5-a)) and the median of the image sampled using the NF-transform (Figure (5-b)) seem to be very close. The same applies for the reconstructed images (Figures (6-a) and (6-b)). In fact, the L_2 distance between the image reconstructed using the F-transform approach (Figure (6-a)) and the original image equals 0.15 while the L_2 distance between the median image reconstructed using the NF-transform approach (Figure (6-b)) and the original image equals 0.17.

 a)  b)

Fig. 5. Downsampled image obtained using the F-transform (a) and median of the interval-valued downsampled image obtained using the NF-transform (b).

 a)  b)

Fig. 6. Reconstruction of image (5-a) using the IF-transform (a) and median of the interval-valued reconstruction of the interval-valued image whose median is plotted in Figure (5-b) using the INF-transform (b).

Naturally, as proved by Perfilieva [23], if a signal has been downsampled using the F-transform on a partition whose granularity is Δ, the best Δ bounded reconstructing kernel (in L_2 distance) is the sampling kernel itself. Thus the median reconstructed image has no reason to be L_2-closer to the original image. This result was confirmed by Unser concerning the crisp partition (spline 0) and the triangular partition (spline 1). The question of how close the reconstructed images are to the original image will be discussed in the next section.

From an image processing standpoint, the lower and upper images presented in Figure (4) look like eroded and dilated images. In fact, maxitive-based image processing is a kind of intermediate approach between conventional linear image processing and mathematical morphology [11].

In this experiment, 95% of the original image is included in the imprecise valued reconstructed image. This percentage is quite representative of what happens generally (see next section). Using a crisp partition would have ensured complete inclusion (see Section 4.6). However, the quality of the image obtained with a crisp partition is very poor compared to using a fuzzy partition, as illustrated by the detailed images in Figure (11). The L_2 distance between the original image and the median of the obtained reconstructed image is 0.25. This highlights the advantage of using a fuzzy partition instead of

a crisp partition in this framework.

The last part of this experiment aims at illustrating Properties 16 and 21. We considered 1000 different randomly selected sampling kernels and 1000 different randomly selected reconstruction kernels. Every sampled image is included in the NF-transformed image. Every reconstructed image is included in the INF-transformed image. Figure (7) plots the 32^{nd} line of the sampled images while Figure (9) plots the 128^{th} line of the reconstructed images, superimposed with the 128^{th} line of the original image.

The kernels were chosen in the class of symmetric centered kernels that comply with the domination property, i.e. each sampling (reconstruction) kernel is included in the core of the sampling (reconstruction) capacity induced by the fuzzy partition. As can be seen in Figures (8) and (10), the precise sampled and reconstructed images are not uniformly spread in the interval-valued image. This is the main reason why the median cannot be considered as the best representative image. Figure (10) illustrates the fact that the original image is not completely included in the interval-valued reconstructed image. Note, however, that in some regions where the original signal goes out of the bounds of the imprecise valued reconstructed image, images reconstructed using the traditional approach are very far from the original signal.

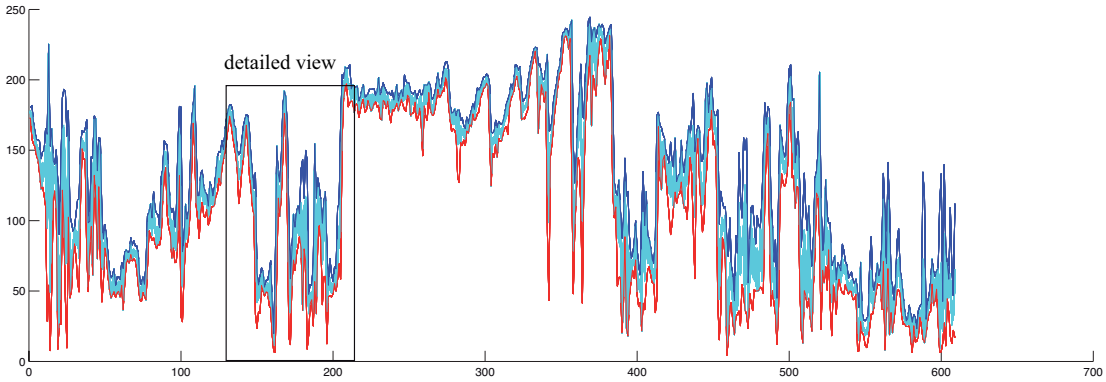

Fig. 7. Lower (in red) and upper (in blue) 32^{nd} line of the NF-transformed image superimposed with the 32^{nd} line of 1000 sampled images (in cyan).

5.3 Statistical properties

As shown in [14], one interesting property of a maxitive kernel-based approach is its ability to quantify the variability in the processed signal. In this experiment, we will show that the imprecision of an image reconstructed by an INF-transform is a marker of the roughness of the obtained reconstruction. To carry out this experiment, we consider the set of $10,000$ images derived from the *BOWS2-Original* image processing database [9].

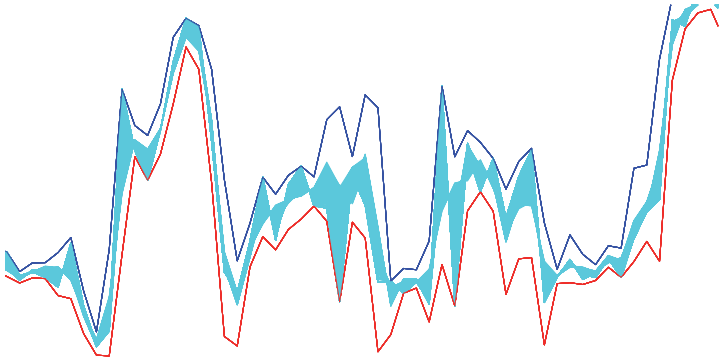

Fig. 8. Detailed part of Figure (7).

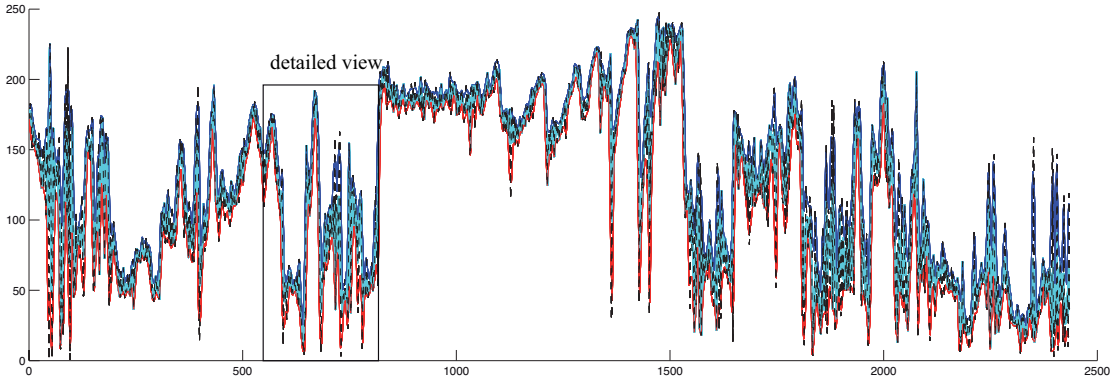

Fig. 9. Lower (in red) and upper (in blue) 128^{th} line of the interval valued reconstructed image superimposed with the 128^{th} line of 1000 reconstructed images (in cyan) and the 128^{th} line of the original image (black dotted line).

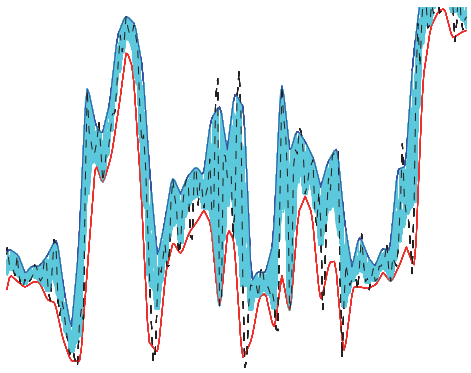

Fig. 10. Detailed part of Figure (9).

For each image I, we compute its F-transform F and NF-transform $\underline{\overline{F}}$. We then reconstruct $\hat{I}$, a precise estimate of the original image, by IF-transforming F and $\underline{\overline{I}} = [\underline{I}, \overline{I}]$, an imprecise estimate of the original image, by INF-transforming $\underline{\overline{F}}$. We compute $\triangle I = \frac{1}{2}(\overline{I} - \underline{I})$, the imprecision of $\underline{\overline{I}}$ and $\tilde{I} = \frac{1}{2}(\overline{I} + \underline{I})$ its median. We finally compute $\square I = |I - \tilde{I}|$. The median is chosen as an objective representative of the interval-valued image since it is the image that is the closest, in L_1 distance, to the set of images included in the interval-valued

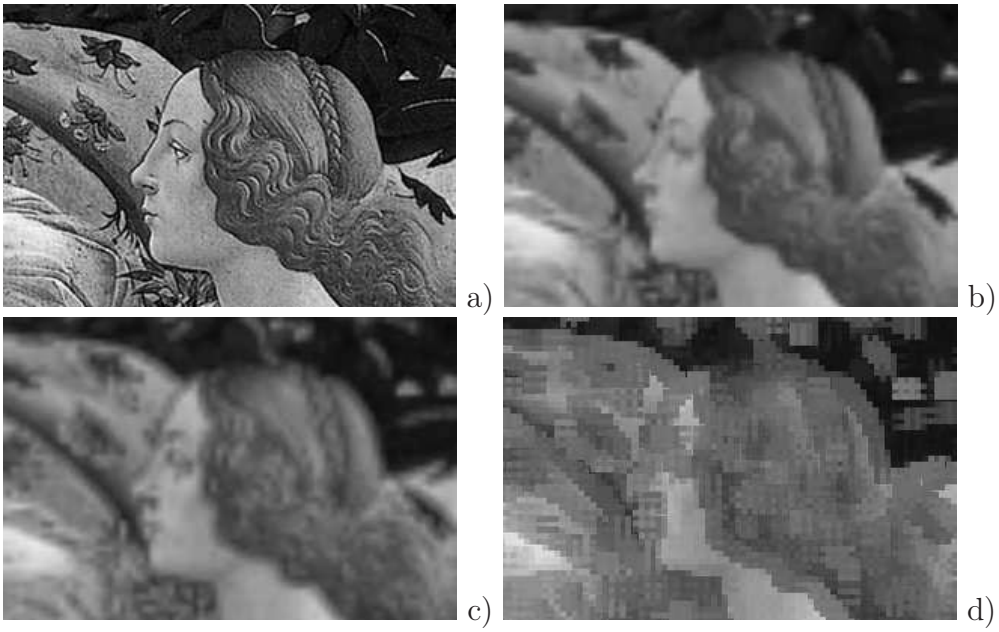

Fig. 11. Details of the original image (a), the image reconstructed with the F-transform approach (with a triangular kernel) (b), median of the image reconstructed with the NF-transform approach with a triangular kernel (c) and with a rectangular (crisp) kernel (d).

image.

Figure (12) plots the signal-to-noise ratio (SNR) of each reconstructed image. This SNR is computed as $10.log_{10}(\frac{\Sigma^2(I)}{\Sigma^2(\Box I)})$ for the interval-valued INF-transformed image and as $10.log_{10}(\frac{\Sigma^2(I)}{\Sigma^2(I-\hat{I})})$ for the F-transformed image, where $\Sigma^2(I)$ is the sum of the squared values of the image I. The higher the SNR, the closer the reconstructed image is to the original image. As expected, the SNR is always better for the precise image than for the median of the imprecise image. However, as illustrated in Figure (12), the SNR are very close. Moreover, Properties 16 and 21 ensure that the "best" image obtained using the F-transform and the IF-transform is included in the interval-valued reconstructed image obtained using the NF-transform and the INF-transform (see also Remark 19).

Generating an image that is closer to the original image is not the main purpose of the NF-transform approach. In fact, as shown by Unser [35], the appropriate kernel to reconstruct an image that has been sampled by a bounded positive kernel is not positive (and often not bounded). This idea corresponds to the least square reconstruction approach [2]. What makes the NF-transform approach new and interesting is that the imprecision of the obtained interval-valued reconstructed image encodes the loss of information induced by the sampling/reconstruction process. To illustrate this property, in Figure (13), we have plotted the cloud of the mean of $\Box I$ versus the mean of $\triangle I$. The

correlation between those two values is obvious. The Pearson correlation coefficient is 0.983.

As we mentioned previously, even if the inclusion of the original image in the interval-valued reconstructed image cannot be guaranteed, the inclusion rate is generally higher than 90%. This property is illustrated in Figure (14). To further illustrate this, we artificially constructed, for each precise reconstructed image (using the F-transform), an interval valued image having a constant imprecision equal to the mean imprecision of the corresponding interval valued reconstructed image (NF-transform). Figure (14) plots the inclusion rate of the original image within each interval-valued image. It can be seen that, even when using the best reconstructed image as the median image, the inclusion rate is much higher in the non-additive than in the additive approach. This confirms that the imprecision really quantifies the roughness of the approximation induced by both direct and inverse transformations. This quantification property can also be seen in Figure (10).

The distance between those interval-valued reconstructions and the original image can also be questioned. We thus compute the Hausdorff generalization of the L_1 distance ([19]) defined by:

$$L_1([a,b],c) = \begin{cases} (a-c) & \text{, if } c < a \\ (c-b) & \text{, if } c > b \\ 0 & \text{, else} \end{cases}$$

Figure(15) plots this distance for the two interval-valued images. The distance is much lower for the non-additive than for the additive approach.

6 Concluding remarks

In this article, we have proposed a new F-transform based on a possibilistic interpretation of the fuzzy subsets involved in the partition. One of the main original features of this framework is that it provides interval-valued transformed signals instead of precise-valued signals. We have defined direct and inverse transforms that are easy interpreted within the signal processing framework: NF-transforming a signal provides the interval of all sampled signals that would have been obtained by using a conventional approach with a kernel belonging to a convenient convex set of kernels. Transforming a discrete signal with an INF-transform provides the interval of all reconstructed signals that would have been obtained by using a conventional approach with a kernel belonging to a relevant convex set of kernels. Numerous properties of this new transform have been mathematically proved that can be useful for

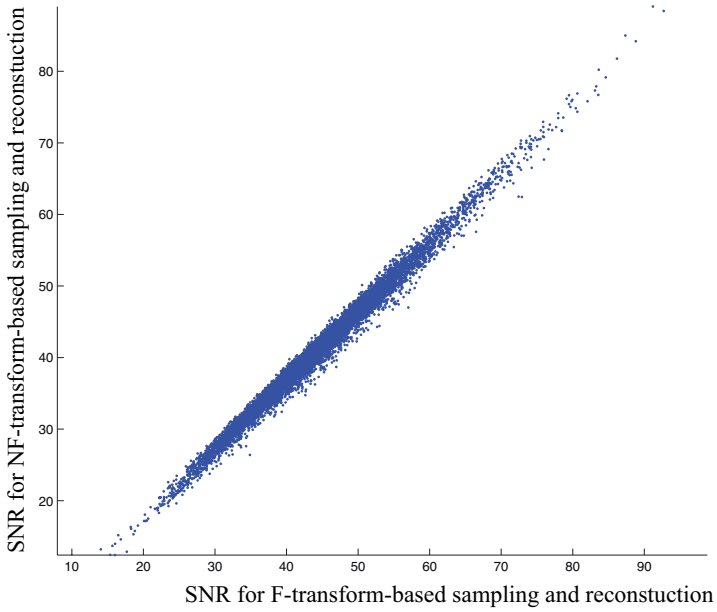

Fig. 12. The SNR of the F-transform-based method versus the SNR of the median of the NF-transform-based method

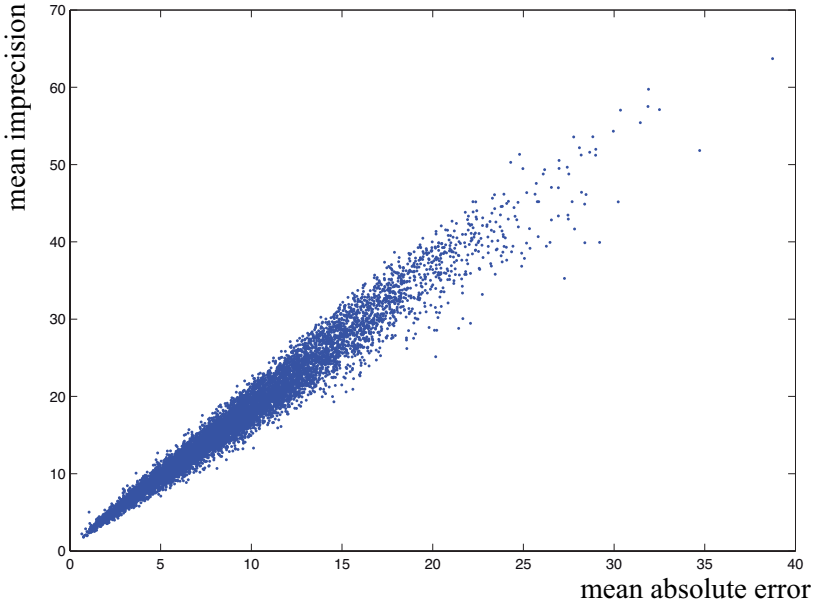

Fig. 13. Correlation of the mean error and the imprecision of the NF-transform-based reconstruction.

guaranteed analysis or comparison of functions. For example, this framework can be used to decide whether or not a digital signal can be considered as a sampled version of a continuous (or a high resolution) original signal. It can also be used to compare two digital signals that are not sampled at the same sampling locations, with the same sampling step, or with the same sensor – with the point spread function being ill-known. It can also be helpful to obtain a guaranteed calculus on sub-sampled signals when the complete signal

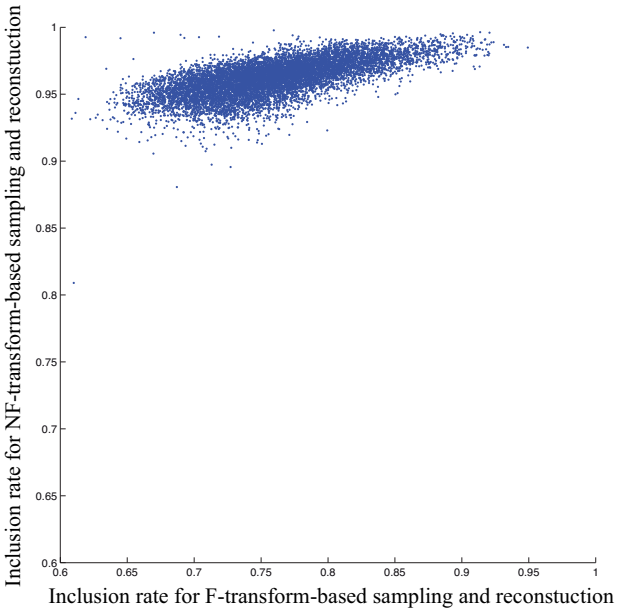

Fig. 14. Inclusion of the original image in the interval-valued images based on the F-transform and on the NF-transform.

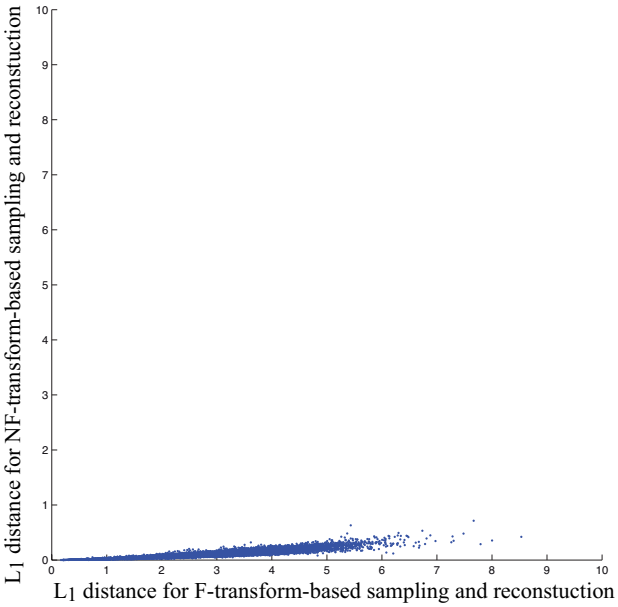

Fig. 15. L_1 distance between the original image and the F-transform and median of the NF-transform-based reconstructions.

cannot be processed for computational complexity or limited memory reasons. Within this kind of application, interval-valued downsampling keeps track of the roughness of the approximation throughout the processing.

The ideal NF-transform/INF-transform couple should provide a convex envelope that contains the original signal. Within the actual framework, this guarantee cannot be ensured. However, the inclusion of an original signal in

the reconstructed interval-valued signal generally exceeds 90%. Moreover, as shown in the experimental part of this article, the imprecision of the reconstructed signal can be seen as a local marker of the roughness of the approximation. The convergence of sampling and reconstructing with the current NF-transform/INF-transform couple has been proved.

A convenient framework should comply with at least one of the properties listed in Section 4.3. One possibility is to define a capacity-based set of non-positive kernels. Another possibility is to enlarge the proposed approach to mimic the least square reconstruction proposed by Crouzet and Patané.

As a future work, we also plan to more thoroughly analyze analogies between the F-transform framework and the *best fit* approach to see if we could define a coupling between the kernel used in the partition and the reconstruction kernel (since the triangular kernel is a B-spline). Perhaps the F-transform or NF-transform framework could be instrumental for defining continuous-based discrete operators.

Some other partitions should also be envisaged, e.g. partitions such that $\forall u \in \Omega$, $\sup_k C_k(u) = 1$. It would also be relevant to try and bridge the gap between the non-additive F-transform approach and the F-transform based on residuated lattices. This could lead to a better reconstruction of the transformed function.

Acknowledgment

The author would like to thank Kevin Loquin and Sebastien Destercke for their helpful comments on the paper.

References

[1] B. Bede and I.J. Rudas. Approximation properties of fuzzy transforms. *Fuzzy Sets and Systems*, 180(1):20–40, 2011.

[2] J-F. Crouzet. The fuzzy projection: a simple adaptive representation of signals. In *IEEE International Conference on Signal Processong*, pages 139 – 142, Beijing, China, October 2010.

[3] J-F. Crouzet. Fuzzy projection versus inverse fuzzy transform as sampling/interpolation schemes. *Fuzzy Sets and Systems*, 193(2):108–121, 2012.

[4] J-F. Crouzet and O. Strauss. Interval-valued probability density estimation based on quasi-continuous histograms: Proof of the conjecture. *Fuzzy Sets and Systems*, 183(1):92–100, 2011.

[5] D.Dubois and H. Prade. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17:191–200, 1990.

[6] D. Denneberg. *Non-Additive Measure and Integral.* Kluwer Academic Publishers, 1994.

[7] D. Dubois. Possibility theory and statistical reasoning. *Computational Statistics and Data Analysis*, 51(1):47–69, 2006.

[8] D. Dubois, H. Prade, L. Foulloy, and G. Mauris. Probability-possibility transformations, triangular fuzzy sets and probabilistic inequalities. *Reliable Computing*, 10:273–297, 2004.

[9] Teddy Furon and Patrick Bas. Broken arrows. *EURASIP Journal on Information Security*, pages 1–13, 2008.

[10] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter. *Applied Interval Analysis with Examples in Parameter and State Estimation, Robust Control and Robotics.* Springer Verlag, 2001.

[11] O. Strauss K. Loquin. Linear filtering and mathematical morphology on an image: a bridge. In *Proc. IEEE International Conference on Image Processing*, page 3965–3968, Le Caire, Egypt, 2009.

[12] M. Liu, D. Chen, C. Wu, and H. Li. Approximation theorem of the fuzzy transform in fuzzy reasoning and its application to the scheduling problem. *Computers and Mathematics with Applications*, 51(3-4):515–526, February 2006.

[13] K. Loquin and O. Strauss. On the granularity of summative kernels. *Fuzzy Sets and Systems*, 159(15):1952–1972, August 2008.

[14] K. Loquin, O. Strauss, and J.F. Crouzet. Possibilistic signal processing: How to handle noise? *International journal of approximate reasoning*, 51(9):1129–1144, 2010.

[15] A. Aldroubi M. Unser and M. Eden. Fast b-spline algorithms for continuous image representation and interpolation. *IEEE Trans. Pattern Anal. Machine Intell.*, 13:277–285, 1991.

[16] F. Di Martino, V. Loia, I. Perfilieva, and S. Sessa. An image coding/decoding method based on direct and inverse fuzzy transforms. *International Journal of Approximate Reasoning*, 48(1):110–131, 2008.

[17] F. Di Martino, V. Loia, and S. Sessa. A segmentation method for images compressed by fuzzy transforms. *Fuzzy Sets and Systems*, 161(1):56–74, 2010.

[18] F. Di Martino, V. Loia, and S. Sessa. Fuzzy transforms method in prediction data analysis. *Fuzzy Sets and Systems*, 180(1):146–163, 2011.

[19] Sam B. Nadler. *Hyperspaces of Sets.* Marcel Dekker, New York, 1978.

[20] J.-N. Paoli, O. Strauss, B. Tisseyre, J.-M. Roger, and S. Guillaume. Spatial data fusion for qualitative estimation of fuzzy request zones: Application on precision viticulture. *Fuzzy Sets and Systems*, 158:535–554, 2007.

[21] G. Patané. Fuzzy transform and least-squares approximation: Analogies, differences, and generalizations. *Fuzzy Sets and Systems*, page Available online, 2011.

[22] Z. Pawlak. *Rough sets: Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publisher, 1991.

[23] I. Perfilieva. Fuzzy transforms: Theory and applications. *Fuzzy Sets and Systems*, 157(8):993–1023, April 2006.

[24] I. Perfilieva and P. Hodáková. Fuzzy and fourier transforms. In *EUSFLAT-LFA 2011*, 2011.

[25] I. Perfilieva and V. Kreinovich. Fuzzy transforms of higher order approximate derivatives: A theorem. *Fuzzy Sets and Systems*, 180(1):55–68, 2011.

[26] I. Perfilieva, V. Novák, and A. Dvořák. Fuzzy transform in the analysis of data. *International Journal of Approximate Reasoning*, 48(18):36–46, 2008.

[27] Irina Perfilieva and Vladik Kreinovich. Fuzzy transform as a new paradigm in fuzzy modeling. *Fuzzy Sets and Systems*, 180:1–2, 2011.

[28] A. Rico and O. Strauss. Imprecise expectations for imprecise linear filtering. *International Journal of Approximate Reasoning*, 51(8):933–947, October 2010.

[29] E. Ruspini. New experimental results in fuzzy. *Information Sciences*, 6:273–284, 1973.

[30] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[31] Laurent Schwartz. *Théorie des distributions*. Hermann & Cie, Paris, 1966.

[32] L. Stefanini. F-transform with parametric generalized fuzzy partitions. *Fuzzy Sets and Systems*, 180(1):98–120, 2011.

[33] O. Strauss and A. Rico. Towards interval-based non-additive deconvolution in signal processing. *Soft computing*, 16(5):809–820, 2012.

[34] L. Troiano and P. Kriplani. Supporting trading strategies by inverse fuzzy transform. *Fuzzy Sets and Systems*, 180(1):121–145, 2011.

[35] M. Unser, A. Aldroubi, and M Eden. Polynomial spline signal approximations: filter design and asymptotic equivalence with shannon's sampling theorem. *IEEE Transactions on Information Theory*, 38(1):95–103, 1992.

[36] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing: Part i-theory. *IEEE Transactions on Signal Processing*, 41:821–833, 1993.

[37] L. Zadeh. Fuzzy sets. *Journal of Mathematical Analysis and Applications*, 23:421–427, 1968.