6-2009

Probabilistic Interpretation of Fuzzy Transforms and Fuzzy Control

Irina Perfilieva

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-09-17

Probabilistic Interpretation of Fuzzy Transforms and Fuzzy Control

Irina Perfilieva[1] and Vladik Kreinovich[2]
[1]University of Ostrava
Inst. for Research and Applications of Fuzzy Modeling
Ostrava, Czech Republic, Irina.Perfilieva@osu.cz
[2]Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA, vladik@utep.edu

Abstract

In many practical applications, it turns out to be useful to use the notion of *fuzzy transform*: once we have functions $A_1(x) \geq 0, \ldots, A_n \geq 0$, with $\sum\limits_{i=1}^{n} A_i(x) = 1$, we can then represent each function $f(x)$ by the coefficients $F_i = \dfrac{\int f(x) \cdot A_i(x)\, dx}{\int A_i(x)\, dx}$. Once we know the coefficients F_i, we can (approximately) reconstruct the original function $f(x)$ as $\sum\limits_{i=1}^{n} F_i \cdot A_i(x)$. The original motivation for this transformation came from fuzzy modeling, but the transformation itself is a purely mathematical transformation. Thus, the empirical successes of this transformation suggest that this transformation can be also interpreted in more traditional (non-fuzzy) mathematics as well.

Such an interpretation is presented in this paper. Specifically, we show that fuzzy transform has a natural probabilistic interpretation – related to the known interpretation of fuzzy sets as equivalence classes of random sets. We also show that a similar interpretation is possible for fuzzy control techniques.

1 Introduction: Fuzzy Transform and the Need for Its Probabilistic Interpretation

Fuzzy transform: a definition. The notion of a fuzzy transform (*F-transform*, for short) turned out to be very useful in many application areas such as image compression, solving differential equations under initial uncertainty, etc.; see, e.g., [6, 7] and references therein.

Generally speaking, the F-transform of function f is a vector with weighted local mean values of f as components. The first step in the definition of the F-transform of $f : X \longrightarrow \mathbb{R}$ is a selection of a *fuzzy partition* of universal set X (e.g, a bounded interval $[a, b]$ on $\mathbb{R}$) by a finite set of *basic functions*

$$A_1(x) \geq 0, \ldots, A_n(x) \geq 0,$$

which are continuous and satisfy the condition: $\sum\limits_{i=1}^{n} A_i(x) = 1$. Basic functions are called *membership functions* of respective fuzzy sets, or, alternatively, *granules*, information pieces, etc. Their choice reflects the type of uncertainty which is related to the knowledge of x.

Once the basic functions are selected, we define the F-transform of a continuous function $f : X \longrightarrow \mathbb{R}$ as a vector $(F_1, \ldots, F_n)$ where

$$F_i \stackrel{\text{def}}{=} \frac{\int f(x) \cdot A_i(x)\, dx}{\int A_i(x)\, dx}. \tag{1}$$

F-transform satisfies the following properties [6, 7]:

- $y = F_i$ minimizes $\int_a^b (f(x) - y)^2 A_i(x) dx$,
- for a twice continuously differentiable function f, $F_i = f(x_i) + O(h_i^2)$, where h_i is the length of the support of A_i.

F-transform is used in applications as a "skeleton model" of f. This model provides a compressed image if f is an image [1], values of a trend if f is a time series [8], a numeric model if f is used in numeric computations (integration, differentiation) [9], etc.

Once we know the F-transform components F_i, we can (approximately) reconstruct the original function f as

$$\overline{f}(x) = \sum_{i=1}^{n} F_i \cdot A_i(x). \tag{2}$$

In [6], the formula (2) is called the *F-transform inversion formula*. (2) represents a continuous function that approximates f. Under certain reasonable conditions, a sequence of functions represented by (2) uniformly converges to f (see [6] for more details).

F-transform: original motivation. The original motivation for F-transform came from fuzzy modeling [6, 7]. The purpose was to show that this type of modeling can be as useful in applications as more traditional techniques such as Fourier transform and wavelet transform. Moreover, F-transform has a potential advantage over Fourier and wavelet transforms: while the Fourier transform uses a single type of basic functions $(\exp(\mathrm{i} \cdot \omega \cdot x))$ and the wavelet transform uses a single "mother wavelet" that determines all the basic functions, F-transform can use several different basic functions A_i. An additional

advantage is that, in contrast to the purely mathematical basic functions used in Fourier and wavelet transforms, the basic functions A_i in a fuzzy partition usually come from natural language terms like "low" or "high". (For a detailed description of fuzzy modeling e.g., [2, 4].)

Just like any other tool of applied mathematics, F-transform is not a panacea. It is more successful in some problems, and in other problems, it is less successful. It is therefore desirable to combine F-transform with other mathematical tools, so as to combine relative advantages of different techniques. For combining F-transform with other mathematical tools, it is desirable to come up with a purely mathematical (non-fuzzy) interpretation for this transform.

In particular, since most mathematical data processing tools are based on probability and statistics, it is desirable to come up with a probabilistic interpretation for F-transform.

What we do in this paper. In this paper, we provide a non-fuzzy interpretation of F-transform. Specifically, we show that F-transform can be naturally interpreted in probabilistic terms.

In view of the above pragmatic argument, this interpretation will (hopefully) make it easier to combine F-transform with traditional (non-fuzzy) statistical data processing tools – the main tools of modern applied mathematics.

We also show that the resulting probabilistic interpretation is related to the known relation between fuzzy sets and equivalence classes of random sets; see, e.g., [3].

We also extend our comparison of probabilistic and fuzzy approaches to more general modeling situations, and provide examples where one of these approaches has an advantage over the other.

2 A Natural Practical Problem that Leads to F-transform

Physical setting: general discussion. Let us assume that we have a physical process that is characterized by two quantities x and z, and we know that these quantities are related by a functional dependence $z = f(x)$.

In the ideal situation of complete knowledge,

- we know the exact value of x, and
- we have the exact description of the function f.

In this case, we can get the corresponding *exact* value $z = f(x)$. of the second quantity.

In practice, we know the value x with uncertainty, i.e., several different values of x are consistent with our knowledge. We must therefore provide a reasonable *estimate* for z. Finding such an estimate will be the *first problem* with which we will be dealing. In this first problem, we assume that the function f is known *exactly.*

If this function has to be determined *empirically*, then we shall transform the empirical (often, partial) knowledge about f into a reasonable estimate for this function. This will be the *second problem* with which we will be dealing in this section.

First problem: estimating the value $f(x)$ for an imprecisely known x. If we only know one piece of information X_i about x, what is the reasonable estimate for $z = f(x)$?

Second problem: estimating the function $z = f(x)$ based on partial information about the dependence between x and z. Assume that for every information piece X_i, $1 \le i \le n$, we have the corresponding measured value F_i of z. Since we know only n numerical characteristics F_i of the unknown function f, we cannot exactly reconstruct this function. Instead, we need to provide a good estimate for each value $f(x)$ of this function.

3 A Natural Probabilistic Problem that Leads to the Probabilistic Interpretation of F-transform

Uncertainty in x: a general probabilistic description. Assume that we have a *model* of the estimation procedure, that enables us, given the actual value x, to compute the probability $P(X_i \mid x) \ge 0$ of this procedure resulting in X_i – under the condition that the actual (unknown) value of the estimated quantity is x.

To simplify formulas, we denote

$$A_i(x) \stackrel{\text{def}}{=} P(X_i \mid x). \tag{3}$$

Since for every x, we must have exactly one of the n possible outcomes, we thus conclude that the probabilities $P(X_1 \mid x)$, $\ldots$, $P(X_n \mid x)$ of different estimation results must add up to one, i.e., we must have

$$P(X_1 \mid x) + \ldots + P(X_n \mid x) = 1. \tag{4}$$

In the above simplified notation, this formula takes the form

$$A_1(x) + \ldots + A_n(x) = 1. \tag{5}$$

First problem: estimating the value $f(x)$ for an imprecisely known x. Let us consider the first problem. In practice, we do not know the exact value of the quantity x. Instead, we only have one of the information pieces X_i, $1 \le i \le n$. Under the assumption that we know X_i, what is the reasonable estimate for $z = f(x)$?

In terms of probability theory, we would like to find the conditional expected value $F_i \stackrel{\text{def}}{=} E[z \,|\, X_i] = E[f(x) \,|\, X_i]$ of $z = f(x)$ under the condition X_i.

By definition, this expected value is equal to

$$F_i = E[f(x) \,|\, X_i] = \int f(x) \cdot P(x \,|\, X_i)\, dx. \tag{6}$$

Thus, to compute this expected value, we must know the probabilities $P(x \,|\, X_i)$. Instead, we know the probabilities $P(X_i \,|\, x)$.

In general, the problem of reconstructing

- probabilities $P(H_x \,|\, X_i)$ of different hypotheses H_x based on the observation X_i from
- conditional probabilities $P(X_i \,|\, H_x)$ of this observation under different hypotheses H_x

is well known in probability theory; it is solved by applying the Bayes theorem. The continuous version of this theorem is

$$P(H_x \,|\, X_i) = \frac{P(X_i \,|\, H_x) \cdot P(H_x)}{\int P(X_i \,|\, H_y) \cdot P(y)\, dy}, \tag{7}$$

in which $P(H_x)$ is a prior probability of the hypothesis H_x (strictly speaking, $P(H_x \,|\, X_i)$ and $P(H_x)$ are probability densities).

In our case, different hypotheses H_x correspond to different possible values x of the quantity of interest. Thus, (7) takes the form

$$P(x \,|\, X_i) = \frac{P(X_i \,|\, x) \cdot P(x)}{\int P(X_i \,|\, y) \cdot P(y)\, dy}. \tag{8}$$

Since there is no a priori reason to prefer one value of x to the other, it is reasonable to assume that all the values x are equally probable, i.e., that all prior values $P(x)$ are equal to each other: $P(x) = P_0$.

Substituting $P(x) = P_0$ into the formula (8) and dividing both the numerator and the denominator by the common factor P_0, we get the expression

$$P(x \,|\, X_i) = \frac{P(X_i \,|\, x)}{\int P(X_i \,|\, y)\, dy}.$$

Substituting this expression into formula (6) (and renaming the variable in the denominator), we get

$$F_i = E[f(x) \,|\, X_i] = \frac{\int f(x) \cdot P(X_i \,|\, x)\, dx}{\int P(X_i \,|\, x)\, dx}.$$

In terms of the simplified notation (3), we thus get

$$F_i = E[f(x) \,|\, X_i] = \frac{\int f(x) \cdot A_i(x)\, dx}{\int A_i(x)\, dx}, \tag{9}$$

i.e., exactly the formula (1) corresponding to F-transform.

Second problem: estimating the function $z = f(x)$ based on partial information about the dependence between x and z. In some practical situations, we do not know the exact expression for the function $f(x)$. Instead, we must estimate $f(x)$ from the empirical data, i.e., from the previous results of simultaneous measuring x and z.

In each such measurement, the only information that we get about x is one of the values $X_1, \ldots, X_n$. For each case when the information about x is X_i, we have one or several values z.

Ideally, we should have a large number of values z corresponding to each x-measurement result X_i. Based on these values z, we should then be able to reconstruct the conditional distribution of z under the condition of X_i. Based on these conditional distributions, we should be able to reconstruct the values $f(x)$ for all x.

In practice, however, we have only a few values z corresponding to each x-measurement result X_i. In this case, at best, instead of the entire conditional probability distribution, we can only reconstruct a single parameter – the conditional mean $F_i = E[z \mid X_i]$. Since we only know n characteristics F_i of the unknown function $f(x)$, we cannot exactly reconstruct this function. Instead, we need to describe a good estimates for each value $f(x)$ of this function.

Similarly to the first problem, we take the mean as a reasonable estimate. Thus, in the above practical setting, the problem of estimating the function $f(x)$ takes the following form:

- for every i, we know the conditional mean $F_i = E[f(x) \mid X_i]$;
- based on these conditional means, for every x, we want to estimate the mean value $\overline{f}(x) \stackrel{\text{def}}{=} E[z \mid x]$.

For this problem, the formula of full probability leads to the following result:

$$E[z \mid x] = \sum_{i=1}^{n} E[z \mid X_i] \cdot P(X_i \mid x). \tag{10}$$

By using the notations $\overline{f}(x)$ for $E[z \mid x]$, F_i for $E[z \mid X_i]$, and $A_i(x)$ for $P(X_i \mid x)$, we can transform the formula (10) into the form

$$\overline{f}(x) = \sum_{i=1}^{n} F_i \cdot A_i(x), \tag{11}$$

i.e., exactly the F-transform inversion formula (2).

Conclusion. Thus, both basic formulas (1) and (2) related to F-transform have been interpreted in probabilistic terms.

Relation with the random set interpretation of fuzzy sets. It is worth mentioning that the above probabilistic interpretation is related to the random set interpretation of fuzzy sets; see, e.g., [3].

In this interpretation, the meaning of an imprecise (fuzzy) term like "small" is based on the following idea. The fact that the term is imprecise means that for the same value x, some people will say that this value is small, while other people will say that this value is not small. To take this imprecision into account, we can store, for each person, a set of all the values that this person considers small.

Since there is no prior reason to prefer the opinion of one of these folks, we consider their opinions equally reasonable. We can then take the ratio $\mu_{\text{small}}(x)$ of people who consider x to be small as a reasonable measure of smallness. (This is actually one of the standard ways to construct a membership function corresponding to a certain term.)

We can describe this ratio in probabilistic terms if we assume that all the persons are equally probable. In these terms, the value $\mu_{\text{small}}(x)$ can be interpreted as the probability $P(\text{small}\,|\,x)$ that a randomly selected person would consider x to be small.

This interpretation of the membership function $A_i(x)$ as the conditional probability $P(X_i\,|\,x)$ is exactly what we came up with in our probabilistic interpretation of F-transform.

Terminological comment. For completeness, let us explain why the above interpretation is called the random sets interpretation.

For crisp (well-defined) properties, each property can be described by the set of all the values that satisfy this property.

For each imprecise property like "small", instead of a *single* set describing all the values that satisfy this property, we have *several* sets describing the opinions of several persons. We consider the opinions of all these persons to be equally valid, so each of N persons has the exact same probability $1/N$ of being correct. In this case, we have different sets, each occurring with probability $1/N$.

In mathematical terms, we can describe this situation by saying that we have a probability distribution on the class of all possible sets. In probability theory, such a distribution is called a *random set* – similarly to the fact that a probability distribution on the class of all possible numbers is called a *random number.*

4 A Similar Probabilistic Interpretation Is Possible For Fuzzy Control

Fuzzy modeling: one of the origins of F-transform. One of the original motivation for F-transform came from *fuzzy modeling*, an area closely related to the most successfully application of fuzzy techniques – to intelligent control. In the previous section, we have shown that F-transform can be interpreted in probabilistic terms. It is therefore reasonable to search for a similar probabilistic interpretation for intelligent control techniques.

In this section, we show that such an interpretation is indeed possible. Moreover, we show that this probabilistic interpretation enables us to improve the traditional intelligent control techniques.

Need for fuzzy control. Before we start this explanation, let us briefly explain what is fuzzy control and how it is different from the traditional control.

In the traditional engineering control,

- we know the equations that describe the dynamics of the system,
- we know the objective – e.g., we want the system to be stable and to return to the desired trajectory even when its state has been perturbed at some moment.

Based on this information, the traditional control theory generates a *control strategy*, i.e., a function $u(x)$ that describes what control value u we should apply at any given state x so as to guarantee the given objectives.

In practice, we often do not know the exact equations describing the system. Instead, we have an experience of skilled human controllers who know how to control the system so as to guarantee the given objective. For example, we may have an experience of a driver how to drive a car, an experience of a professional pilot on how to pilot a plane, etc. It is desirable to apply the experience of the best human controllers to all similar control situations.

In the ideal case, we simply write down what control u the expert controller applies for different possible input values x. In practice, however, an expert controller cannot formulate his or her expertise in these precise mathematical terms. For example, an experienced fighter pilot cannot exactly say for how long and at what angle he turns the corresponding controller. Instead, the expert can describe his or her control strategy by using natural-language rules like "if the deviation x from the desired trajectory is small, apply the small value of the control u for a short period of time".

To use this control in an automatic controller, we must therefore transform such natural-language ("fuzzy") rules into a precise control strategy $u(x)$. Methodologies for such a transformation are called *fuzzy control methodologies.*

Mamdani's approach to fuzzy control: a brief reminder. As we have mentioned, in fuzzy control methodology, we start with rules like

"if x is small, then u should be medium",

and then use membership functions for "small" and "medium" to transform these rules into an exact control strategy.

In general, we have rules

"if x has a property A_i then u has the property B_i" $(1 \le i \le n)$,

with known membership functions $A_i(x)$ and $B_i(u)$ for the corresponding properties. Mamdani's methodology is based on saying that for each input x, the

value u is a reasonable value of control if and only if one of the above n rules is applicable, i.e.,

- either the first rule is applicable, i.e., x satisfies the property A_1 and u satisfies the property B_1,
- or the second rule is applicable, i.e., x satisfies the property A_2 and u satisfies the property B_2,
- ...
- or the n-th rule is applicable, i.e., x satisfies the property A_n and u satisfies the property B_n.

Once we select functions $f_\&(a,b)$ and $f_\vee(a,b)$ to represent "and" and "or" (these functions are called *t-norm* and *t-conorm*), we can thus describe the degree of our belief $\mu_x(u)$ that u is reasonable (for a given input x) as

$$\mu_x(u) = f_\vee(f_\&(A_1(x), B_1(u)), \ldots, f_\&(A_n(x), B_n(u))). \tag{12}$$

In particular, if we select $f_\&(a,b) = a \cdot b$ and $f_\vee(a,b) = \min(a+b,1)$ (and if the added values do not go beyond 1), we get

$$\mu_x(u) = \sum_{i=1}^{n} A_i(x) \cdot B_i(u). \tag{13}$$

Once we know this membership function, we can find the appropriate value of u by using the so-called *centroid defuzzification*:

$$\overline{u}(x) = \frac{\int u \cdot \mu_x(u)\, du}{\int \mu_x(u)\, du}. \tag{14}$$

Towards a probabilistic interpretation of fuzzy control. Similarly to the above probabilistic interpretation of F-transform, let us assume that we have possible pieces of information $X_1, \ldots, X_n$ about the quantity x, and that for each piece of information, we also know the corresponding probability $P(X_i \mid x)$ which we will be denoted by $A_i(x)$.

Similarly, let us assume that we have possible pieces of information $U_1, \ldots, U_m$ about u, and we know the corresponding probabilities $P(U_i \mid u)$ which we will denote by $B_i(u)$.

We know that u depends on x, but we do not know the exact dependence. Instead, for each information X_i about x, we know the corresponding information U_j about the corresponding u.

Since we did not select any specific order for the informations U_i, we can select the value corresponding to X_1 as U_1, the value corresponding to X_2 by U_2, etc. Under this selection, the available information simply means that if x is described by the piece of information X_i, then the corresponding u is described by the piece of information U_i.

Our objective is, given these rules and given a new value x, to find a good estimate for the appropriate u.

Due to the formula of full probability, the conditional probability density $P(u\,|\,x)$ of u under the condition x has the form

$$P(u\,|\,x) = \sum_{i=1}^{n} P(u\,|\,U_i) \cdot P(X_i\,|\,x). \tag{15}$$

We know the probabilities $P(X_i\,|\,x) = A_i(x)$. The probability densities $P(u\,|\,U_i)$ can be determined by using the Bayes theorem – similarly to the F-transform case – as

$$P(u\,|\,U_i) = \frac{P(U_i\,|\,u)}{\int P(U_i\,|\,y)\,dy}, \tag{16}$$

i.e., in terms of the values $B_i(u)$, as

$$P(u\,|\,U_i) = \frac{B_i(u)}{\int B_i(y)\,dy}. \tag{17}$$

Substituting the formula (17) and the expression (3) into the formula (15) (and changing the multiplication order), we get the formula

$$P(u\,|\,x) = \sum_{i=1}^{n} A_i(x) \cdot \frac{B_i(u)}{\int B_i(y)\,dy}. \tag{18}$$

Once we know these probabilities, we can produce the mean $\overline{u}$ as a reasonable estimate for u:

$$\overline{u}(x) = \frac{\int u \cdot P(u\,|\,x)\,du}{\int P(u\,|\,x)\,du}. \tag{19}$$

Probabilistic and fuzzy approaches to control: comparing the resulting formulas. If we compare the formulas (18) and (19) for the probabilistic control with the formulas (13) and (14) for Mamdani's approach to fuzzy control, we conclude that:

- the formula (19) is exactly the same as (14), with $P(u\,|\,x)$ instead of $\mu_x(u)$;
- however, the formula (18) is slightly different from Mamdani's formula (13) – by the integral in the denominator.

In some cases, the two controls lead to the same result. For F-transform (and, more generally, in all the cases when the value $\int B_i(y)\,dy$ is the same for all i), this additional denominator simply divides all the values $P(u\,|\,x)$ by the constant. This constant appears both in the numerator and in the denominator of the formula (18) and thus, it does not affect the resulting value $\overline{u}(x)$.

In some cases, probabilistic control is better. When the values $\int B_i(y)\,dy$ are different, probabilistic control and fuzzy control lead, in general, to a different value $\overline{u}$. We will show, on an example originally proposed by R. Yager, that in this case, the result of the probabilistic control is closer to common sense that the result of Mamdani's control.

Indeed, let us consider the situation in which we have two rules:

- the first rule is a more general rule saying that if x is small, then u should be small;
- the second rule is a very specific rule, saying that if x is very close to 0.11, then u should be very close to 0.15.

Intuitively, if we have a value x for which a very specific rule is applicable, e.g., the value $x = 0.11$, then this specific rule should have a priority over the general rule. However, since the width of the membership function $B_2(u)$ is small, the corresponding term in (13) will practically not affect the resulting estimate (14).

In contrast, in the probabilistic control, the effect of $B_2(u)$ is normalized by, crudely speaking, the total width of the corresponding function $B_2(u)$. Thus, even the most specific rules will have – as desired – the significant influence on the result (19).

Comment. It should be mentioned that the problem with specific rules occurs only in Mamdani's approach to fuzzy control. In the alternative *logical* approach, this problem does not appear; see, e.g., [5].

Another case when probabilistic control is better. The probabilistic interpretation enables us to naturally consider more general situations in which the rules are themselves probabilistic, i.e., when, for each i and j, we know the conditional *probability* $P(U_i \,|\, X_j)$ that if x has the property X_j, then u has the property U_i.

In other words, instead of the original rules

"if x has the property X_i, then u has the property U_i",

we now have rules

"if x has the property X_j, then u has the property U_i
with probability $P(U_i \,|\, X_j)$".

Indeed, in this case, due to the formula of full probability, the conditional probability density $P(u \,|\, x)$ of z under the condition x has the form

$$P(u \,|\, x) = \sum_{i=1}^{n} \sum_{j=1}^{n} P(u \,|\, U_i) \cdot P(U_i \,|\, X_j) \cdot P(X_j \,|\, x). \tag{20}$$

Here, we know the original probabilities $P(U_i \,|\, X_j)$ and the probabilities $P(X_i \,|\, x) = A_i(x)$. The probability densities $P(u \,|\, U_i)$ can be determined by

using the Bayes theorem as an expression (17). Substituting the formula (17) and the expression $P(X_i \,|\, x) = A_i(x)$ into the formula (20) (and changing the multiplication order), we get the formula

$$P(u \,|\, x) = \sum_{i=1}^{n} \sum_{j=1}^{n} P(U_i \,|\, X_j) \cdot A_j(x) \cdot \frac{B_i(u)}{\int B_i(y)\, dy}. \tag{21}$$

Once we know these probabilities, we can produce the mean $\overline{u}$ by using the formula (19).

In some cases, fuzzy control is better. We have shown that in some situations, probabilistic control is better than the original Mamdani's fuzzy control. However, in other situations, the fuzzy control is better. Let us give two examples.

In some cases, fuzzy control is easily applicable but the probabilistic control is difficult to apply. The above probabilistic formulas only work for the case when $\sum\limits_{i=1}^{n} A_i(x) = 1$ – i.e., in the probabilistic terms, when the properties A_i are mutually exclusive. In practice, we may have non-exclusive properties, in which case we may have $\sum\limits_{i=1}^{n} A_i(x) > 1$.

It is not clear how to handle this situation within the probabilistic approach. However, such situations are not a problem if we apply fuzzy control: its formulas are applicable no matter whether we satisfy the requirement $\sum\limits_{i=1}^{n} A_i(x) = 1$ or not.

In some cases, fuzzy control leads to a better quality control. The probabilistic interpretation is only possible when we use multiplication and addition as "and" and "or" operations $f_\&$ and $f_\vee$.

Fuzzy control does not necessarily have to use these operations, it can use different t-norms and t-conorms. It is an empirical fact that in many control situations, the use of t-norm different from the product and of the t-conorm different from the sum leads to a much better quality control – e.g., a more stable or a smoother one.

In [12], we have formulated the problem of selecting the t-norm and the t-conorm as a precise optimization problem, and for several objective functions like smoothness or stability, we gave an explicit analytical solutions to these optimization problem – specifically, we described the selection that leads to the optimal values of smoothness or stability. In many of these case, the optimal selection is indeed different from the probabilistic case of product and sum. Thus, fuzzy control methodology indeed leads to a better quality control.

5 Conclusion

The fuzzy transform (F-transform) techniques have been lately shown to be very successful in various applications, including applications where until recently, only more traditional tools like Fourier transform or wavelet transform have been applied. In many other applications, however, the traditional tools have a clear advantage. It is therefore desirable to combine F-transform with the more traditional tools, so as to combine the relative advantages of both techniques. To make this combination easier, it is desirable to interpret F-transform in traditional mathematical terms.

A probabilistic interpretation is also desirable for fuzzy control, so that we will be able to combine the relative advantages of fuzzy and probabilistic approaches.

Our analysis have shown that

- F-transform and fuzzy control can be adequately interpreted in probabilistic terms – i.e., in terms of the most widely used approach to uncertainty;
- the need to combine probabilistic and fuzzy approaches to control comes from the fact that
  - in certain situations, fuzzy techniques lead to better results than the probabilistic approach: e.g., fuzzy techniques result, in general, in a more stable and/or more smooth control strategy;
  - in other situations, probabilistic approach leads to a more reasonable control: e.g., when some control rules are more specific than the others.

Acknowledgments.

This work was supported in part by the National Science Foundation grant HRD-0734825, by Grant 1 T36 GM078000-01 from the National Institutes of Health, and by Grant MSM 6198898701 from MŠMT of Czech Republic.

The authors are thankful to Josef Štěpán and Ron Yager for motivations and valuable discussions.

References

[1] F. Di Martino, S. Sessa, V. Loia and I. Perfilieva, "An image coding/decoding method based on direct and inverse fuzzy transforms", *International Journal of Approximate Reasoning*, 2008, Vol. 48, pp. 110131.

[2] G. Klir amd B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applicatins*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[3] H. T. Nguyen, *An Introduction to Random Sets*, Chapman & Hall/CRC, Boca Raton, Florida, 2006.

[4] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman & Hall/CRC, Boca Raton, Florida, 2006.

[5] V. Novak, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Dordrecht, 1999.

[6] I. Perfilieva, "Fuzzy transforms: theory and applications", *Fuzzy Sets and Systems*, 2006, Vol. 157, pp. 993–1023.

[7] I. Perfilieva, "Fuzzy transforms: a challenge to conventional transforms", *Advances in Imaging and Electron Physics*, 2007, Vol. 47, pp. 137–196.

[8] I. Perfilieva, V. Novák, V. Pavliska, A. Dvořák, and M. Štěpnička, "Analysis and Prediction of Time Series Using Fuzzy Transform". In: *Proceedings of WCCI 2008, IEEE International Conference on Neural Networks IJCNN'2008*, Hong Kong, June 1–6, 2008, pp. 3875–3879.

[9] I. Perfilieva, H. De Meyer, B. De Baets, and D. Plšková, "Cauchy Problem with Fuzzy Initial Condition and Its Approximate Solution with the Help of Fuzzy Transform". In: *Proceedings of WCCI 2008, IEEE International Conference on Fuzzy Systems FUZZ-IEEE'2008*, Hong Kong, June 1–6, 2008, pp. 2285–2290.

[10] S. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, Springer Verlag, New York, 2005.

[11] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC, Boca Raton, Florida, 2004.

[12] M. H. Smith and V. Kreinovich, "Optimal strategy of switching reasoning methods in fuzzy control", In: H. T. Nguyen, M. Sugeno, R. Tong, and R. Yager (eds.), *Theoretical aspects of fuzzy control*, J. Wiley, N.Y., 1995, pp. 117–146. 2005.