Probabilistic representation of spatial fuzzy sets

Lubomír Soukup[1]

[1]Institute of Information Theory and Automation of the Czech Academy of Sciences
soukup@utia.cas.cz

Abstract

Membership function of a given fuzzy set is expressed by probability that a point belongs in the fuzzy set. Such a membership function is derived from probability distribution of points on the boundary of the fuzzy set. Polygonal boundary is considered. Spatial operations (conjunction, disjunction, complement) are defined accordingly. Several application areas are mentioned, namely classification of land cover, cadastral mapping, material quality analysis, interferometric monitoring of bridges.

1 Introduction

Several attempts had been made to represent uncertain real objects by means of precise mathematical tools. Two most successful approaches – probability theory and fuzzy sets theory – yet seem to be non-compatible. These two approaches have been heavily applied to plenty of real-world problems, but still there are little understanding of their mutual relationship. The early attempt to bring together probability and fuzziness was made by the founder of the fuzzy sets theory (Zadeh, 1968). His achievements were further worked out from more general point of view in (Singpurwalla and Jane M. Booker, 2004). Researchers who have tried making fuzzy set theory and probability theory work in concert usually agree with (Zadeh, 1995) that the two approaches are complementary rather than compatible or competitive. These authors customarily conclude that vaugness and randomness demonstrate two different aspects of uncertainty of the real world. Therefore fuzzy sets differ from imprecise regions, membership function differs from probability measure. Overview of such results is presented in (Schmitz and Morris, 2006).

Nevertheless, there are some cases that allow membership function of a fuzzy set to gain probabilistic meaning. Passing reference of this possibility can be found in book (Viertl, 1996). This case typically occures when fuzzy sets are spatially defined, namely as geographic regions. Some interesting attempts to represent imprecise regions stem from geographical information sciences, e. g. (Cunha and Martins, 2014), (Bruin, 2000). Therefore, geographical motivation stands beyond the approach addressed in this contribution.

2 Formulation of the problem

We are searching for a spatial fuzzy region whose boundary is uncertain due to imprecise position of points on its boundary. Provided that the boundary of the region has a polygonal shape the problem can be concisely formulated as follows:

2.1 Given:

1. two–dimensional closed polygonal region with imprecise vertices,
2. probability distribution of each vertex of the polygonal boundary.

2.2 Required:

1. probability that a point in 2D plane belongs to the given region,
2. fuzzy set whose membership function is defined by the probabilities evaluated by means of 1.,
3. fuzzy set operations (conjunction, disjunction, complement) of probabilistic fuzzy sets created by means of 2.

The given polygonal region can look as Figure 1 shows.

3 Solution — 1D case

3.1 Membership function

Principle of creating the probabilistic membership function can be easily explained in 1D case. Let us suppose that the spatial fuzzy set is formed by a line segment $\mathcal{F}_{A,B}$. Boundary of this set consists of two imprecise points, position of which is given by random variables X_A, X_B. Probability that some fixed point x_U belongs in the fuzzy set $\mathcal{F}_{A,B}$ can be evaluated by

$$\begin{aligned} P(x_U \in \mathcal{F}_{A,B}) &= P((X_A < x_U) \wedge (X_B > x_U) \,|\, X_A < X_B) = \\ &= \frac{\displaystyle\int_{-\infty}^{x_U} f_A(x)\,dx \int_{x_U}^{\infty} f_B(y)\,dy}{\displaystyle\int_{-\infty}^{\infty} f_B(y) \int_{-\infty}^{y} f_A(x)\,dx\,dy} \end{aligned} \tag{1}$$

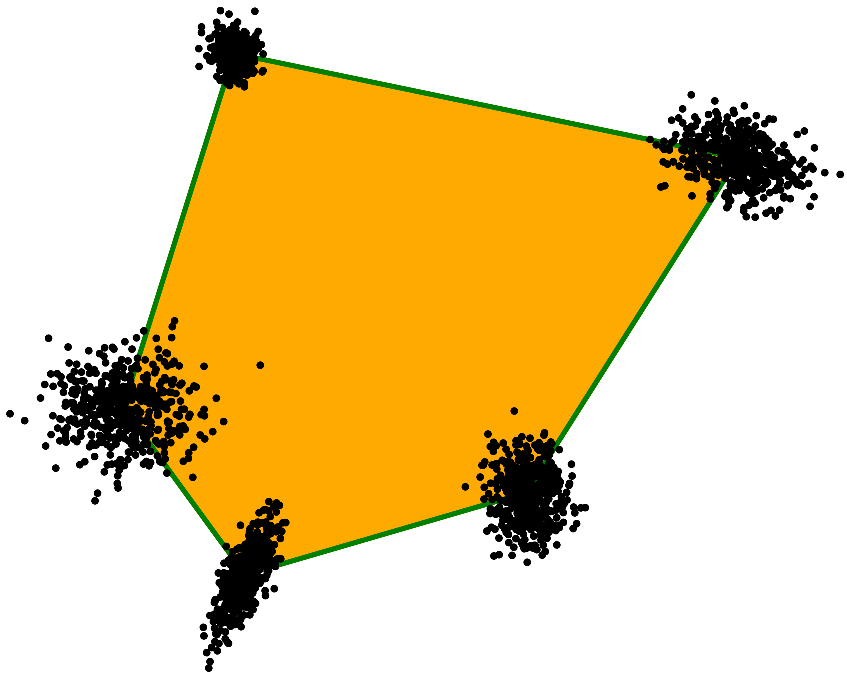

Figure 1: Polygonal region with imprecise vertices. Precision of each vertex is modeled by a 2D probability distribution (black clusters).

Function f_A, resp. f_B stands for probability density function of random variable X_A, resp. X_B. If the probability density functions are Gaussian $\mathcal{N}(\hat{x}_A, \sigma_A)$, resp. $\mathcal{N}(\hat{x}_B, \sigma_B)$, the resulting probability is quite simple:

$$P(x_U \in \mathcal{F}_{A,B}) = \frac{\left(1 + \operatorname{erf}\left(\frac{x_U - \hat{x}_A}{\sqrt{2}\,\sigma_A}\right)\right)\left(1 - \operatorname{erf}\left(\frac{x_U - \hat{x}_B}{\sqrt{2}\,\sigma_B}\right)\right)}{2\left(1 + \operatorname{erf}\left(\frac{\hat{x}_B - \hat{x}_A}{\sqrt{2\,(\sigma_A^2 + \sigma_B^2)}}\right)\right)} . \quad (2)$$

Function erf stands for error function, i. e.

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt .$$

Membership function of fuzzy set $\mathcal{F}_{A,B}$ can therefore be defined by probability (1).

$$\mu_{\mathcal{F}_{A,B}} : \mathbb{R} \to \langle 0, 1 \rangle : \xi \mapsto \mu_{\mathcal{F}_{A,B}}(\xi) := P(\xi \in \mathcal{F}_{A,B}) . \quad (3)$$

Illustration of the probabilistic membership function gives Figure 2.

3.2 Fuzzy set operations

3.2.1 Probabilistic fuzzy conjunction

Probability that some fixed point x_U belongs in conjunction of two fuzzy sets $\mathcal{F}_{A,B}$, $\mathcal{F}_{C,D}$ can be expressed similarly as in (1).

$$\begin{aligned} & P(x_U \in \mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D}) = \\ = \; & P((X_A < x_U < X_B) \wedge (X_C < x_U < X_D) \,|\, (X_A < X_B) \wedge (X_C < X_D)) = \\ = \; & \frac{\int_{-\infty}^{x_U} f_A(w)\,dw \int_{x_U}^{\infty} f_B(x)\,dx \int_{-\infty}^{x_U} f_C(y)\,dy \int_{x_U}^{\infty} f_D(z)\,dz}{\int_{-\infty}^{\infty} f_B(x) \int_{-\infty}^{x} f_A(w)\,dw\,dx \int_{-\infty}^{\infty} f_D(z) \int_{-\infty}^{z} f_C(y)\,dy\,dz} \,. \end{aligned} \tag{4}$$

If the border probability density functions are Gaussian $\mathcal{N}(\hat{x}_A, \sigma_A)$, $\mathcal{N}(\hat{x}_B, \sigma_B)$, $\mathcal{N}(\hat{x}_C, \sigma_C)$, $\mathcal{N}(\hat{x}_D, \sigma_D)$, the resulting probability will be as follows:

$$\begin{aligned} & P(x_U \in \mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D}) = \\ = \; & \frac{\left(1 + \operatorname{erf}\left(\frac{x_U - \hat{x}_A}{\sqrt{2}\,\sigma_A}\right)\right)\left(1 - \operatorname{erf}\left(\frac{x_U - \hat{x}_B}{\sqrt{2}\,\sigma_B}\right)\right)\left(1 + \operatorname{erf}\left(\frac{x_U - \hat{x}_C}{\sqrt{2}\,\sigma_C}\right)\right)\left(1 - \operatorname{erf}\left(\frac{x_U - \hat{x}_D}{\sqrt{2}\,\sigma_D}\right)\right)}{4\left(1 + \operatorname{erf}\left(\frac{\hat{x}_B - \hat{x}_A}{\sqrt{2\,(\sigma_A^2 + \sigma_B^2)}}\right)\right)\left(1 + \operatorname{erf}\left(\frac{\hat{x}_D - \hat{x}_C}{\sqrt{2\,(\sigma_C^2 + \sigma_D^2)}}\right)\right)} \,. \end{aligned} \tag{5}$$

3.2.2 Probabilistic fuzzy disjunction

Probabilistic fuzzy disjunction can be easily deduced from conjunction operation (4) with aid of elementary theorem of probability theory

$$P(K \vee L) = P(K) + P(L) - P(K \wedge L) \tag{6}$$

which holds for any random events K, L.

$$P(x_U \in \mathcal{F}_{A,B} \vee \mathcal{F}_{C,D}) = P(x_U \in \mathcal{F}_{A,B}) + P(x_U \in \mathcal{F}_{C,D}) - P(x_U \in \mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D}) \tag{7}$$

Evaluation of (7) under assumption of normal distribution can be simply done with substitutions (2), (5).

3.2.3 Probabilistic fuzzy complement

Membership function of probabilistic fuzzy complement can be obtained from (1) as easy as fuzzy conjunction from (4). The same theorem (6) can be used with special option

$L = \neg K$. Due to this option, $P(K \vee \neg K) = 1$ and $P(K \wedge \neg K) = 0$ for any random event K. Theorem (6) then claims

$$1 = P(K) + P(\neg K) \ .$$

Thus the following equalities hold for fuzzy complement $\neg \mathcal{F}_{A,B}$.

$$P(x_U \in \neg \mathcal{F}_{A,B}) = P(x_U \notin \mathcal{F}_{A,B}) = 1 - P(x_U \in \mathcal{F}_{A,B}) \tag{8}$$

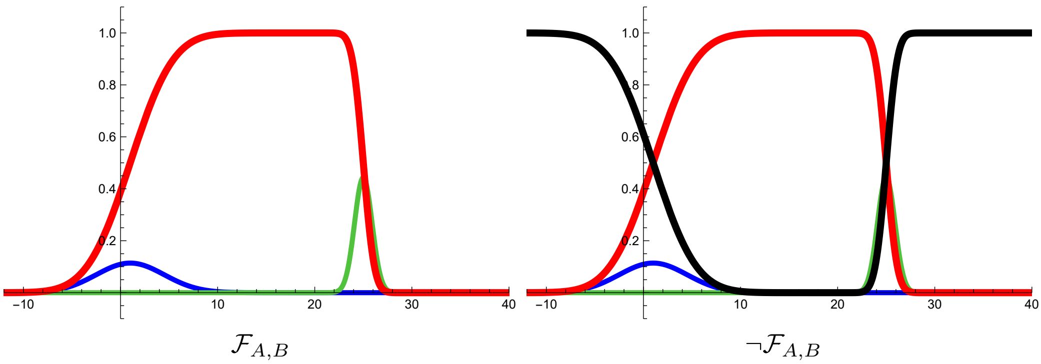

Figure 2: Membership function of probabilistic fuzzy set $\mathcal{F}_{A,B}$ – drawn by red line. Border probability density functions f_A (blue line) and f_B (green line) are shown on the left and right hand side. The right part of the figure shows also membership function of probabilistic fuzzy complement $\neg \mathcal{F}_{A,B}$ (black line).

3.2.4 Membership functions of the fuzzy set operations

Membership functions of the above introduced fuzzy set operations are given by

$$\begin{aligned}
\mu_{\mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D}}(\xi) &:= P(\xi \in \mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D}) \ , \\
\mu_{\mathcal{F}_{A,B} \vee \mathcal{F}_{C,D}}(\xi) &:= P(\xi \in \mathcal{F}_{A,B} \vee \mathcal{F}_{C,D}) = \mu_{\mathcal{F}_{A,B}}(\xi) + \mu_{\mathcal{F}_{C,D}}(\xi) - P(\xi \in \mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D}) \ , \\
\mu_{\neg \mathcal{F}_{A,B}}(\xi) &:= P(\xi \notin \mathcal{F}_{A,B}) = 1 - \mu_{\mathcal{F}_{A,B}}(\xi) \ .
\end{aligned} \tag{9}$$

These membership functions are defined by means of the source membership functions $\mu_{\mathcal{F}_{A,B}}$, $\mu_{\mathcal{F}_{C,D}}$ except fuzzy conjunction. Therefore, proper definiton of fuzzy conjunction has to be accomplished by the following two–step procedure.

1. extract border density functions f_A, f_B from $\mu_{\mathcal{F}_{A,B}}$ and f_C, f_D from $\mu_{\mathcal{F}_{C,D}}$,
2. evaluate $P(x_U \in \mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D})$ with the aid of (4) .

The first step of this procedure may not be satisfactorily achievable since the membership functions $\mu_{\mathcal{F}_{A,B}}$, $\mu_{\mathcal{F}_{C,D}}$ can be given in other ways than by (1). Extraction of border density functions from an arbitrary membership function is subject of further research.

Illustration of the fuzzy set operations gives Figure 3 and Figure 2.

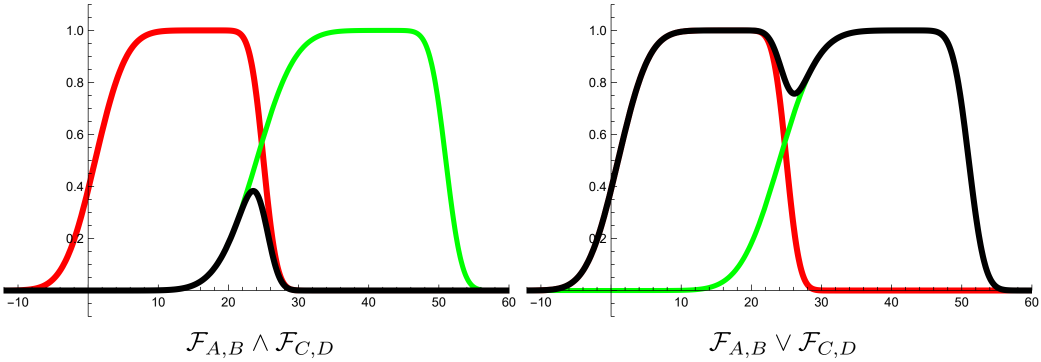

Figure 3: Membership functions of probabilistic fuzzy conjunction $\mathcal{F}_{A,B} \wedge \mathcal{F}_{C,D}$ (left) and disjunction $\mathcal{F}_{A,B} \vee \mathcal{F}_{C,D}$ (right) are drawn by black line. The original fuzzy sets are shown by red and light green colors.

4 Solution — 2D case

4.1 2D membership function

Two–dimensional generalization of formula (1) becomes much more complicated if a polygonal boundary is considered. Closed polygon with n vertices is given. Coordinates of i-th vertex are given in form of column vector $\mathbf{x}_i \in \mathbb{R}^2$, $i \in \{1, 2, \ldots, n\}$. Sides of the polygon must not intersect each other if they are not adjacent. Moreover, for the sake of simplicity, the polygon is supposed to be convex. Coordinates of the all vertices creates $2n$-dimensional vector

$$\mathbf{x} := [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] \ .$$

Area of the given polygon is

$$S(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^{n} \det([\mathbf{x}_i, \mathbf{x}_{i+1}]) \ , \tag{10}$$

where $[\mathbf{x}_i, \mathbf{x}_{i+1}]$ is 2×2 real matrix for each $i \in \{1, 2, \ldots, n\}$,

$$\mathbf{x}_{n+1} := \mathbf{x}_1 \ .$$

Note that $S(\mathbf{x}) \geq 0$ if the polygon is oriented counter–clockwise.

Let each vertex be a random point in $\mathbb{R}^2$. Then the all random vertices form random vector

$$\mathbf{X} := [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n]$$

that creates random boundary of a polygon $\mathcal{F}_n$. Such imprecise polygon can or cannot cover a fixed point $\mathbf{x}_U \in \mathbb{R}^2$. Probability that the point $\mathbf{x}_U$ lies inside the imprecise polygon is

$$P(\mathbf{x}_U \in \mathcal{F}_n) = P(\mathbf{X} \in \mathcal{U}(\mathbf{x}_U) \,|\, (S(\mathbf{X}) > 0) \wedge \mathcal{C}(\mathbf{X})) \ . \tag{11}$$

Statement $\mathcal{C}(\mathbf{x})$ claims that polygon with vertices $\mathbf{x}$ is convex.

Set $\mathcal{U}(\mathbf{x}_U)$ stands for the all possible convex polygons that include point $\mathbf{x}_U$. It is defined by

$$\mathcal{U}(\mathbf{x}_U) := \{\mathbf{x} \in \mathbb{R}^{2n} \mid \left(\bigwedge_{i=1}^{n} \kappa(\mathbf{x}_i, \mathbf{x}_{i+1}, \mathbf{x}_U) > 0 \right) \wedge \mathcal{C}(\mathbf{x})\} \ ,$$

where function κ expresses perpendicular oriented distance of the point $\mathbf{x}_U$ from line that passes through oriented tuple of points $\mathbf{a}, \mathbf{b}$.

$$\kappa : \mathbb{R}^6 \to \mathbb{R} \ : \ [\mathbf{a}, \mathbf{b}, \mathbf{x}_U] \mapsto (\mathbf{a} - \mathbf{x}_U) \cdot \frac{(\mathbf{b} - \mathbf{a})^{\perp}}{||\mathbf{b} - \mathbf{a}||} \tag{12}$$

Symbol $^{\perp}$ stands for perpendicularity.

$$^{\perp} : \mathbb{R}^2 \to \mathbb{R}^2 \ : \ [x, y] \mapsto [x, y]^{\perp} := [-y, x] \ .$$

If the vertices of the polygonal boundary have probability distribution with density function f, then

$$P(\mathbf{x}_U \in \mathcal{F}_n) = \frac{\underbrace{\int\int\cdots\int\int}_{\mathcal{U}(\mathbf{x}_U)} f(\mathbf{x})\, d\mathbf{x}}{P((S(\mathbf{X}) > 0) \wedge \mathcal{C}(\mathbf{X}))} \ . \tag{13}$$

The multiple integral is $2n$-tuple.

$$P((S(\mathbf{X}) > 0) \wedge \mathcal{C}(\mathbf{X})) = \underbrace{\int\int\cdots\int\int}_{(S(\mathbf{x})>0)\wedge\mathcal{C}(\mathbf{x})} f(\mathbf{x})\, d\mathbf{x} \ .$$

4.2 2D fuzzy set operations

4.2.1 Probabilistic 2D fuzzy conjunction

Conjunction of two 2D fuzzy sets, say $\mathcal{F}_n$, $\mathcal{G}_m$ can be defined similarly as in 1D case (4) by probability

$$P(\mathbf{x}_U \in \mathcal{F}_n \wedge \mathcal{G}_m) = \frac{\underbrace{\int\int\cdots\int\int}_{\mathcal{U}(\mathbf{x}_U)\wedge\mathcal{V}(\mathbf{x}_U)} f(\mathbf{x})\, g(\mathbf{y})\, d\mathbf{x}\, d\mathbf{y}}{P(S(\mathbf{x}) > 0 \wedge S(\mathbf{y}) > 0 \wedge \mathcal{C}(\mathbf{x}) \wedge \mathcal{C}(\mathbf{y}))} \ , \tag{14}$$

where $\mathbf{y}$ is $2m$-dimensional vector which contains coordinates of m vertices of a convex polygon. Corresponding random vector $\mathbf{Y}$ has probability density function g.

$$\mathcal{V}(\mathbf{x}_U) := \{\mathbf{y} \in \mathbb{R}^{2m} \mid \left(\bigwedge_{i=1}^{m} \kappa(\mathbf{y}_i, \mathbf{y}_{i+1}, \mathbf{x}_U) > 0 \right) \wedge \mathcal{C}(\mathbf{y})\} \ ,$$

where function κ was introduced in (12).

$$P(S(\mathbf{X}) > 0 \wedge S(\mathbf{Y}) > 0 \wedge \mathcal{C}(\mathbf{X}) \wedge \mathcal{C}(\mathbf{Y})) = \underbrace{\int\!\!\int \cdots\cdots \int\!\!\int}_{S(\mathbf{x})>0 \wedge S(\mathbf{y})>0 \wedge \mathcal{C}(\mathbf{x}) \wedge \mathcal{C}(\mathbf{y})} f(\mathbf{x})\, g(\mathbf{y})\, d\mathbf{x}\, d\mathbf{y} \; .$$

The above multiple integrals are $2(n+m)$-tuple.

Suitable approximation have to be applied to evaluate the multiple integration for normal densities f, g.

4.2.2 Probabilistic 2D fuzzy disjunction

$$P(x_U \in \mathcal{F}_n \vee \mathcal{G}_m) = P(x_U \in \mathcal{F}_n) + P(x_U \in \mathcal{G}_m) - P(x_U \in \mathcal{F}_n \wedge \mathcal{G}_m) \; . \tag{15}$$

4.2.3 Probabilistic 2D fuzzy complement

$$P(x_U \in \neg\mathcal{F}_n) = P(x_U \notin \mathcal{F}_n) = 1 - P(x_U \in \mathcal{F}_n) \; . \tag{16}$$

4.2.4 Membership functions of the 2D fuzzy set operations

Membership functions of the 2D fuzzy set operations are given similarly as in (9) by

$$\begin{aligned} \mu_{\mathcal{F}_n \wedge \mathcal{G}_m}(\xi) &:= P(\xi \in \mathcal{F}_n \wedge \mathcal{G}_m) \; , \\ \mu_{\mathcal{F}_n \vee \mathcal{G}_m}(\xi) &:= P(\xi \in \mathcal{F}_n \vee \mathcal{G}_m) = \mu_{\mathcal{F}_n}(\xi) + \mu_{\mathcal{G}_m}(\xi) - P(\xi \in \mathcal{F}_n \wedge \mathcal{G}_m) \; , \\ \mu_{\neg\mathcal{F}_n}(\xi) &:= P(\xi \notin \mathcal{F}_n) = 1 - \mu_{\mathcal{F}_n}(\xi) \; . \end{aligned} \tag{17}$$

Problem of extracting probability density functions f, g from membership functions $\mu_{\mathcal{F}_n}$, $\mu_{\mathcal{G}_m}$ is much more arduous in 2D than in 1D case. Solution of this problem has crucial importance in real examples, namely in cartography and material analysis. These examples will be addressed in the next section.

5 Applications

Probabilistic fuzzy regions can be found all around. Therefore, the designed approach has many practical applications, namely classification of land cover, cadastral mapping, material quality analysis, interferometric monitoring of bridges.

5.1 Cartography and material analysis

One of the most frequent task of digital cartography is classification of land cover. Simplest case of the classification is recognition of certain region of interest against other type of earth surface. The region of interest has to be localized by determination of points on its boundary and input into geographic information system (GIS). Precision of these points can be inferred from so called probability map which is by–product of classification procedure. The resulting region of interest then will be obtained as an imprecise region similar to Figure 1. Spatial operations that are necessary component of every GIS can then be realized by the fuzzy operation designed in this contribution. Similar problems with imprecise boundary occures in cadastral systems where parcels have polygonal shapes.

Another application area has been emerged in material quality analysis. For example, shape and size changes of microscopic grains in cocnrete blocks under radiation exposure are important for security assessment of nuclear powerplants. Size of these grains can be precisely estimated with the aid of the probabilistic representation designed in this contribution.

5.2 Radar interferometry

Radar interferometry is very effective method for determination of descends and rises of a bridge under traffic load. Magnitude and direction of movements of the bridge body can be very precisely (up to 0.01 mm) measured by two ground based radars (GB-RAR), see Figure 4. Unfortunately, part of the bridge body (so called range–bin) that corresponds to the measurement cannot be determined precisely. Gray–scale rectangles on Figure 4 represent response of radar rays from different range–bins. Regions of the range–bins are imprecise, so that their conjunctions are imprecise as well. Probabilistic quantification of these imprecisions which is offered in the presented fuzzy approach can improve interpolation quality of the bridge movement in an arbitrary point.

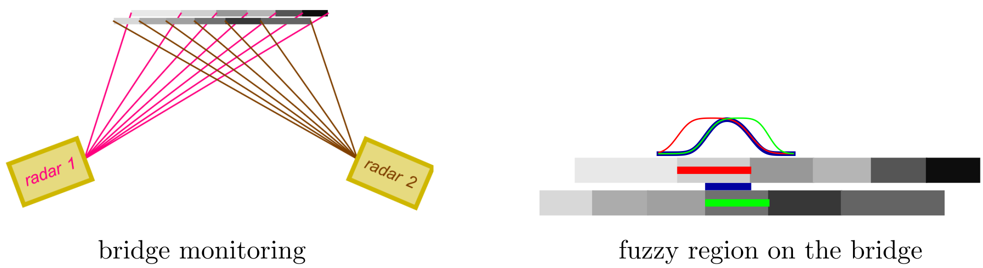

Figure 4: **Bridge monitoring** (left) by simultaneous measurement of two interferometric radars. Gray–scale rectangles represent response of radar rays from different range–bins. **Fuzzy region on the bridge** (right) shows conjunction of two overlapping range–bins (red and green). The conjunction is marked by dark blue color.

6 Conclusion

Representation of a fuzzy set of polygonal shape was designed in such a way that probability of membership of any point in the set is derived from imprecision of the polygonal boundary. This representation enables a non-traditional definition of fuzzy set operations (conjunction, disjunction, complement) that produces other fuzzy sets which gain the same probabilistic interpretation. Several real-world applications of the presented approach were addressed.

Acknowledgement

The research was supported by the Ministry of Interior of the Czech Republic in the framework of grant project "The advanced technology of rapid determination of bridges deformation by radar interferometry and its use in diagnostics", No. VI20192022167.

References

S. D. Bruin. Querying probabilistic land cover data using fuzzy set theory. *International Journal of Geographical Information Science*, 14(4):359–372, 2000.

E. Cunha and B. Martins. Using one-class classifiers and multiple kernel learning for defining imprecise geographic regions. *International Journal of Geographical Information Science*, 28(11):2220–2241, 2014.

A. Schmitz and A. Morris. Modelling and manipulating fuzzy regions: strategies to define the topological relations between fuzzy regions. *Control and Cybernetics*, 35:245–261, 2006.

N. D. Singpurwalla and Jane M. Booker. Membership Functions and Probability Measures of Fuzzy Sets. *Journal of the American Statistical Association*, 99(467):867–877, 2004.

R. Viertl. *Statistical Methods for Non-Precise Data.* CRC Press, 1996.

L. A. Zadeh. Probability measures of fuzzy events. *Journal of Mathematical Analysis and Applications*, 23:421–427, 1968.

L. A. Zadeh. Discussion: Probability Theory and Fuzzy Logic Are Complementary Rather than Competitive. *Technometrics*, 37:271–276, 1995.