Fast Computation of Centroids for Constant-Width Interval-Valued Fuzzy Sets

Jerry M. Mendel and Hongwei Wu
Signal and Image Processing Institute
Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089-2564, USA

Vladik Kreinovich and Gang Xiang
NASA Pan-American Center for
Earth and Environmental Studies
University of Texas at El Paso
El Paso, TX 79968, USA, vladik@utep.edu

1. Practical Need for Fuzzy Sets: A Brief Reminder

- *Fact:* in many application areas, top-quality best experts (medical doctors, pilots, etc.) provide the best recommendations.
- *Desirable:* incorporate the knowledge of these top experts into automated (expert) systems.
- *Problem:* usually, experts cannot describe their knowledge in precise terms.
- *Example:* size is "approximately 1.5, with an error about 0.1".
- *Fuzzy logic:* designed by Zadeh to formalize such statements.
- *How:* a statement like "approximately 1.5" is described by a *membership function*:
  - to each real value x around 1.5,
  - we assign a degree $\mu(x)$ to which this value x fits the expert's description.

2. Where Do Membership Degrees Come From: A Brief Reminder

- *Fact:* membership degrees $\mu(x)$ describe the degree of certainty of an expert.
- *Conclusion:* it is natural to ask the expert to provide these degrees.
- *Interpolation:*
  - an expert provides the degree of certainty for several values x;
  - then we use interpolation and extrapolation to estimate the values of $\mu(x)$ for all other x.
- *Important case:* we ask the expert to provide us with:
  - the range $[\underline{x}, \overline{x}]$ of possible values – so that $\mu(x) = 0$ for all $x \notin [\underline{x}, \overline{x}]$;
  - the "most probable" value $\widetilde{x}$ of the quantity x – for which $\mu(\widetilde{x}) = 1$.

  We use linear interpolation on $[\underline{x}, \widetilde{x}]$ and $[\widetilde{x}, \overline{x}]$.
- *Resulting* $\mu(x)$ is triangular:
  - 0 outside the interval $[\underline{x}, \overline{x}]$;
  - its graph on this interval consists of two straight line segments:
    * rising from 0 to 1, and then
    * going from 1 to 0.

3. Applications of Fuzzy Sets: In Brief

- *Expert systems:* we transform fuzzy rules and fuzzy inputs into fuzzy recommendations.
- *Intelligent control:* the objective is to generate a single (crisp) control value u_c.
- *Need for defuzzification:*
  - we first combine the membership functions of the inputs into a membership function $\mu(u)$ for the desired control u;
  - then, we must transform $\mu(u)$ into a single value u_c.
- *Centroid defuzzification:* minimizing the mean square difference between the actual (unknown) optimal control u and u_c:

$$\int \mu(u) \cdot (u - u_c)^2 \, du \to \min_{u_c}.$$

- *Resulting algorithm:* differentiate by u_c and equate to 0:

$$u_c = \frac{\int u \cdot \mu(u) \, du}{\int \mu(u) \, du}.$$

- *Comment:* it is called centroid because it is the center of mass of the region under the graph of $\mu(u)$.

4. Limitations of Traditional Fuzzy Sets and a Practical Need for Interval-Valued Fuzzy Sets

- *Assumption:* an expert can provide us with the degrees $\mu(x)$ that describe how values x fit his/her statement.
- *Fact:* this expert is unable to provide the crisp values of the desired quantities.
- *Corollary:* the expert is usually unable to describe his or her degree of certainty by an exact number.
- *In practice:* an expert can provide us with:
  - an *approximate* degree $\widetilde{\mu}(x)$, and
  - a bound $\Delta(x)$ describing his uncertainty.
- *Enter interval:* in effect, we have an *interval* of possible values of certainty:

$$[\underline{\mu}(x), \overline{\mu}(x)] = [\widetilde{\mu}(x) - \Delta(x), \widetilde{\mu}(x) + \Delta(x)].$$

- *Comment:* interval-valued membership functions
  - also provides us with the expert's estimates of how accurately this expert estimates his/her own uncertainty;
  - are, thus, a more adequate description of expert's uncertainty.

5. Important Case: Constant-Width Interval-Valued Fuzzy Sets

- *Fact:* extracting the valued $\mu(x)$ from an expert is difficult.
- *Corollary:* extracting the uncertainty bounds on these degrees is even more difficult.
- *Conclusion:* in practice, often, at best, we can ask for an overall estimate Δ of the expert's uncertainty:
  - instead of different bounds $\Delta(x)$ for different values x,
  - we get the same bound Δ for all x.
- *Constant-width set:* $\overline{\mu}(x) - \underline{\mu}(x) = 2\Delta \equiv \text{const}$.

6. Increase in Computational Complexity: One of the Main Reasons Preventing Us from Wider Use of Interval-Valued Fuzzy Sets

+ Interval-valued fuzzy sets have been successfully applied to many practical problems.

– Replacing traditional fuzzy sets with interval-valued fuzzy sets often leads to an increase in computational complexity:

  – to describe a traditional fuzzy set, we store a *single* value $\mu(x)$ for each x;

  – to describe an interval-valued fuzzy set, we need *two* values for every x:

    * either the endpoints $\underline{\mu}(x)$ and $\overline{\mu}(x)$,
    * or midpoint $\widetilde{\mu}(x)$ and half-width $\Delta(x)$.

- *Fact:* for efficient data processing algorithms, the computation time grows as a square $O(n^2)$ or a cube $O(n^3)$ of the size n of the input.

- *Corollary:* the doubling of the input size leads to a 4 or 8 times increase in computation time.

- *This is crucial:* the increase in computation time is especially crucial in real-time control applications.

7. Centroid Defuzzification: The Most Time-Consuming Part of Intelligent Control

- *Reminder:*
$$u_c = \frac{\int u \cdot \mu(u)\,du}{\int \mu(u)\,du}.$$
- *Computation time:* if we have n values $\mu(x_i)$ of $\mu(x)$, we need a linear time $O(n)$.
- *In practice:* we often know the analytical expression for $\mu(u)$ (e.g., triangular) and we can integrate analytically.
- *Result:* we get explicit analytical expressions for u_c – and thus, avoid time-consuming integrations altogether.
- *Interval-valued case:* in effect, we may have different $\mu(x) \in [\underline{\mu}(x), \overline{\mu}(x)]$.
- *Fact:* different $\mu(x)$ lead, in general, to different u_c.
- *Objective:* produce an *interval* $[\underline{u}_c, \overline{u}_c]$ of possible values of u_c.
- *Situation:* for each of n values x, we consider m possible values
$$\mu(x) \in [\underline{\mu}(x), \overline{\mu}(x)].$$
- *Straightforward approach:* we need to check m^n different membership functions.
- *Problem:* for large n, this number exceeds the number of particles in the Universe.

8. Sophisticated Algorithms Make Centroids of Interval-Valued Fuzzy Sets Computable – But Still More Computationally Intensive Than for Traditional Fuzzy Sets

- *Known formulas:*

$$\underline{u}_c = \min_\ell \frac{\underline{N}(\ell)}{\underline{D}(\ell)}, \overline{u}_c = \max_\ell \frac{\overline{N}(\ell)}{\overline{D}(\ell)},$$

where

$$\underline{D}(\ell) \stackrel{\text{def}}{=} \int_{\underline{u}}^{\ell} \overline{\mu}(u)\,du + \int_{\ell}^{\overline{u}} \underline{\mu}(u)\,du; \quad \underline{N}(\ell) \stackrel{\text{def}}{=} \int_{\underline{u}}^{\ell} u \cdot \overline{\mu}(u)\,du + \int_{\ell}^{\overline{u}} u \cdot \underline{\mu}(u)\,du;$$

$$\overline{D}(\ell) \stackrel{\text{def}}{=} \int_{\underline{u}}^{\ell} \underline{\mu}(u)\,du + \int_{\ell}^{\overline{u}} \overline{\mu}(u)\,du; \quad \overline{N}(\ell) \stackrel{\text{def}}{=} \int_{\underline{u}}^{\ell} u \cdot \underline{\mu}(u)\,du + \int_{\ell}^{\overline{u}} u \cdot \overline{\mu}(u)\,du.$$

- *Algorithm:*
  - compute these ratios for different values ℓ, and
  - to find the ratio which is the smallest (for $\underline{u}_c$) or the largest (for $\overline{u}_c$).
- *Computation time:* $O(n)$ steps for each of n values ℓ, i.e., $O(n^2)$.

\+ This is drastically faster than m^n steps – and actually doable.

– It is still much slower than $O(n)$ computational steps that we need for a traditional fuzzy set.

9. There Are Techniques for Computing the Centroid somewhat Faster – but in General, These Faster Computations Still Require Time $O(n^2) \gg O(n)$.

- *Known fact:* the min and the max of a function are attained when the derivative is 0.
- *Corollary:*
  - for $\underline{u}_c$, we get $\underline{D}(\ell) \cdot \ell = \underline{N}(\ell)$;
  - for $\overline{u}_c$, we get $\overline{D}(\ell) \cdot \ell = \overline{N}(\ell)$.
- *Conclusion:* we can find:
  - $\underline{u}_c$ as the value ℓ for which $\ell = \dfrac{\underline{N}(\ell)}{\underline{D}(\ell)}$;
  - $\overline{u}_c$ as the value ℓ for which $\ell = \dfrac{\overline{N}(\ell)}{\overline{D}(\ell)}$.
- *Advantage:*
  - once we fine ℓ, we can stop;
  - so, on average, we need to check $n/2$ values ℓ;
  - for optimization, in general, we need to check all n values ℓ.
- *Remaining problem:* this is twice faster, but we still need $O(n^2)$ steps.

10. Our Result: Fast Computation of Centroids for Arbitrarily Shaped Constant-Width Interval-Valued Fuzzy Sets

- *Known fact:* for a constant-width triangular membership function, we can use explicit formulas to compute $\underline{u}$ and $\overline{u}$.
- *New result:* extend to arbitrary constant-width interval-valued fuzzy sets.
- *Main idea:* $\overline{\mu}(u) = \underline{\mu}(u) + 2\Delta$, hence

$$\underline{D}(\ell) = \int_{\underline{u}}^{\overline{u}} \underline{\mu}(u)\,du + \int_{\underline{u}}^{\ell} 2\Delta\,du = D_0 + 2\Delta \cdot (\ell - \underline{u}),$$

where $D_0 \stackrel{\text{def}}{=} \int_{\underline{u}}^{\overline{u}} \underline{\mu}(u)\,du$. Similarly,

$$\underline{N}(\ell) = \int_{\underline{u}}^{\overline{u}} u \cdot \underline{\mu}(u)\,du + \int_{\underline{u}}^{\ell} u \cdot 2\Delta\,du = N_0 + \Delta \cdot (\ell^2 - \underline{u}^2),$$

where $N_0 \stackrel{\text{def}}{=} \int_{\underline{u}}^{\overline{u}} u \cdot \underline{\mu}(u)\,du$.

- *Corollary:* the condition $\underline{D}(\ell) \cdot \ell = \underline{N}(\ell)$ is a quadratic equation, hence $\underline{u}_c = \ell = \dfrac{-(D_0 - 2\Delta \cdot \underline{u}) \pm \sqrt{\underline{d}}}{2\Delta}$, where $\underline{d} = D_0^2 + 4\Delta \cdot (N_0 - D_0 \cdot \underline{u})$.
- Similarly, $\overline{u}_c = \ell = \dfrac{(D_0 + 2\Delta \cdot \overline{u}) \pm \sqrt{\overline{d}}}{2\Delta}$, where $\overline{d} \stackrel{\text{def}}{=} D_0^2 - 4\Delta \cdot (N_0 - D_0 \cdot \overline{u})$.
- *Algorithm:* linear time $O(n)$.

Practical Need for . . .
Where Do . . .
Applications of Fuzzy . . .
Limitations of . . .
Important Case: . . .
Increase in . . .
Centroid . . .
Sophisticated . . .
There Are Techniques . . .
Our Result: Fast . . .
Discussion: Can We . . .
Acknowledgments
Title Page
Page 11 of 13
Go Back
Full Screen
Close
Quit

11. Discussion: Can We Extend Our Algorithms to a More General Case?

- *Linear width:* for a linear (or piece-wise linear) $\Delta(u)$, we get *cubic* equations for $\underline{u}_c$ and $\overline{u}_c$.

+ For cubic equations, there are also explicit formulas.

– These formulas are not so numerically simple and useful.

- *Case of narrow intervals:* when $\Delta(u) \ll \widetilde{\mu}(u)$, we can *linearize*, i.e., ignore quadratic and higher order terms in $\Delta(u)$.

- *Resulting algorithm:*
  - first, compute $\widetilde{u} = \dfrac{\widetilde{N}}{\widetilde{D}}$, where $\widetilde{D} \stackrel{\text{def}}{=} \int \widetilde{\mu}(u)\,du$, $\widetilde{N} \stackrel{\text{def}}{=} \int u \cdot \widetilde{\mu}(u)\,du$;
  - then, compute $\delta u = \dfrac{\delta N}{\widetilde{D}}$, where

$$\delta N \stackrel{\text{def}}{=} \int_{\tilde{u}}^{\overline{u}} u\cdot\Delta(u)\,du - \int_{\underline{u}}^{\tilde{u}} u\cdot\Delta(u)\,du - \tilde{u}\cdot\left(\int_{\tilde{u}}^{\overline{u}} \Delta(u)\,du - \int_{\underline{u}}^{\tilde{u}} \Delta(u)\,du\right);$$

  - the resulting centroid interval is $[\widetilde{u} - \delta u, \widetilde{u} + \delta u]$.

- *Complexity:* $O(n)$, but with three intervals ($\widetilde{D}$, $\widetilde{N}$, and δN) instead of two (D and N).

12. Acknowledgments

This work was supported in part:

- by NSF grant EAR-0225670,
- by NASA under cooperative agreement NCC5-209,
- by NIH grant 3T34GM008048-20S1,
- by Army Research Lab grant DATM-05-02-C-0046,
- by Star Award from the University of Texas System,
- and by Texas Department of Transportation grant No. 0-5453.