

3

Técnicas de Fusão de Dados Aplicadas a Medição

3.1

Introdução

A fusão de dados de múltiplos sensores tem sido muito estudada para aplicações de sistemas que empregam um número de sensores com características iguais ou diferentes. As principais áreas de aplicação são: militar (C³I – *Command, Control, Communication and Information*), robótica (robôs inteligentes), medicina e segurança pública (análise e reconhecimento de imagens), indústria (detecção de falhas, monitoramento e controle de processos) [LUO, 2002].

A fusão de dados de múltiplos sensores envolve a combinação de informação coletada por diversas fontes (sensores), para derivar informações úteis que não são fornecidas pelos sensores individuais, principalmente quando a confiabilidade e a exatidão são consideradas. Portanto, o procedimento de medição baseado em técnicas de fusão de dados de sensores permite que o sistema seja mais tolerante a falhas e com capacidade para prover novas informações, que nenhum sensor poderia fornecer de maneira individual [FACELI, 2004].

Diversas aplicações de fusão de dados têm se concentrado em resolver o problema do refinamento dos valores produzidos através de processamento de dados. Entretanto, especificamente para instrumentação, a fusão de dados é utilizada fundamentalmente para reduzir potencialmente as incertezas associadas às medições, aumentando a confiança dos resultados medidos, por meio da combinação dos dados individuais [KELLY, 1999].

Em sistemas de medição, cada tipo de sensor apresenta sua própria limitação, com características de precisão e custo associado. Em muitos casos, algumas características de sensores necessárias ou desejáveis para uma aplicação em particular, como exatidão e confiabilidade, são limitadas por restrições, tais como o custo ou a complexidade do sensor. Levando-se em consideração estas restrições, a informação pode ser obtida a partir da combinação de um conjunto simples de sensores de baixo custo, provendo (com menor custo) informação mais completa, precisa e confiável acerca do ambiente. Deste modo, aplicações de sistemas de múltiplos sensores se sustentam nas limitações fundamentais inerentes à descrição de um ambiente baseado em sistemas com um único sensor, tais como:

- Um único sensor provê informação parcial sobre o seu ambiente de operação, tornando-se insuficiente para interpretar completamente os eventos e limitado na resolução de ambigüidades;

- As observações feitas por um único sensor podem ser incertas e ocasionalmente espúrias ou incorretas. Sistemas com um único sensor não têm a capacidade de reduzir esta incerteza intrínseca ou medidas errôneas;
- Os diferentes sensores fornecem tipos diferentes de informação apropriados para diferentes circunstâncias e tarefas;
- Sistemas que utilizam um único sensor não são robustos e podem falhar com consequências drásticas em sistemas críticos.

Pode-se dizer que o propósito dos métodos de fusão consiste em utilizar ativamente a diversidade da informação disponível a partir de múltiplos sensores para superar as limitações inerentes ao uso de um único sensor [FACELI, 2004].

Neste capítulo, tem-se como objetivo revisar algumas das principais técnicas de fusão de dados aplicadas à área de instrumentação, assim como descrever uma estrutura de integração de informações que faz uso de técnicas de fusão de dados.

3.2

Fusão de Dados Aplicada a Medição: Estado da Arte

A adoção de uma abordagem unificada para o projeto de sistemas de decisão baseado na técnica de fusão de dados tem um impacto significativo no que diz respeito ao desempenho de tais sistemas. Assim, aplicações baseadas em processos inteligentes de fusão de dados devem conter:

- Representação padronizada dos dados a serem combinados;
- Metodologias iterativas (com base na engenharia de *software* e *hardware*) para o desenvolvimento de sistemas inteligentes de fusão de dados;
- Formas de extração do conhecimento a partir dos dados qualitativos e quantitativos com base em modelos computacionais;
- Métodos que suportem múltiplos modelos de integração de dados;
- Mecanismos destinados à captura dos dados baseados em uma análise racional do sistema;
- Processos para a validação dos resultados obtidos.

Os sistemas de integração de múltiplas fontes de informação iniciam-se com a atuação de um mecanismo de tradução que realize o mapeamento de dados brutos coletados de um sistema qualquer para o formato padronizado de representação escolhido (homogeneização). Em seguida, os dados traduzidos são interpretados e combinados por mecanismos inteligentes de fusão de dados, normalmente implementados através de algoritmos de otimização que têm como objetivo a combinação de dados. Finalmente, um mecanismo inteligente de diagnóstico

atua, fornecendo a informação de melhor qualidade possível (a mais provável e crível) sobre o estado do sistema [DO COUTTO, 2007].

Para o propósito de integração de múltiplos sensores, uma estrutura de integração é apresentada na Figura 3.1, que contém os principais módulos ou blocos funcionais que envolvem os problemas mais gerais de uma arquitetura para fusão de múltiplos sensores, tais como [FACELI, 2001], [LUO, 2002]:

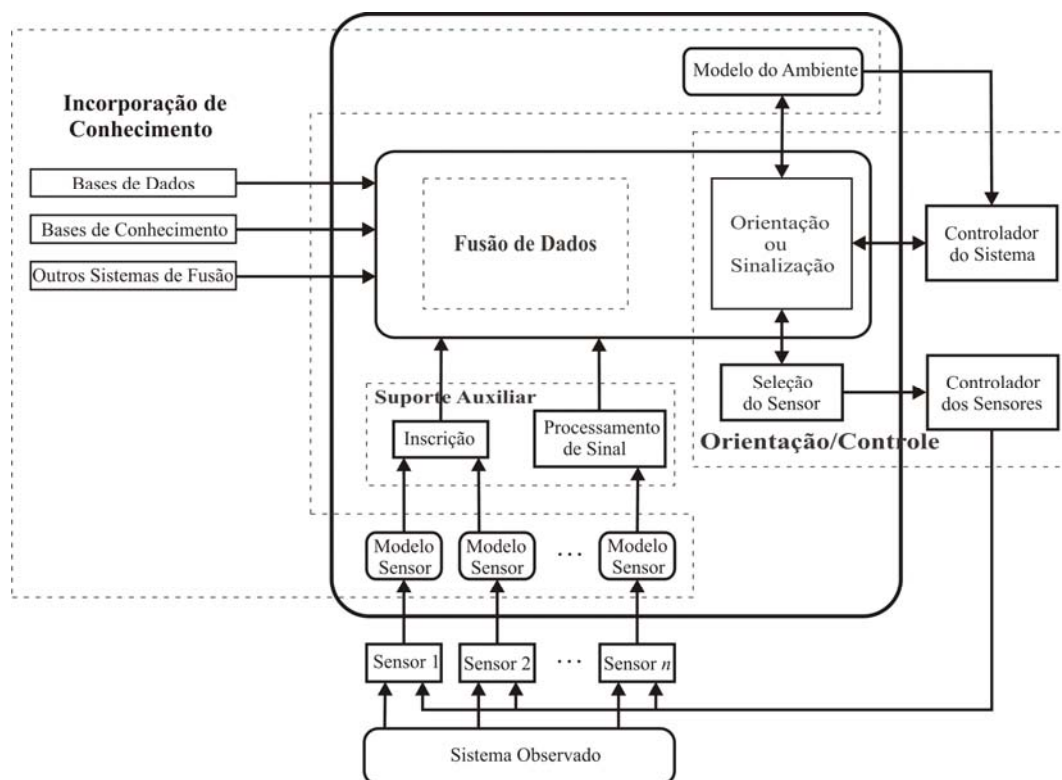


Figura 3.1. Diagrama funcional de integração de múltiplos sensores.

- **Módulo de Suporte Auxiliar:** módulo que requer elevado poder de processamento para realizar operações de alinhamento e de pré-processamento de dados, e gerenciamento da base de dados.
- **Módulo de Orientação/Controle:** inclui os sub-módulos de seleção dos sensores, orientação/sinalização, controlador do sistema e controlador dos sensores. Esses sub-módulos são elementos essenciais em algoritmos de fusão fortemente acoplados, em que a operação de um sub-módulo sensorial é afetada pela saída de outro módulo.
- **Módulo de Incorporação de Conhecimento:** composto pelos modelos dos sensores, o modelo do ambiente e entradas externas, tais como bases de dados, bases de conhecimento e outros sistemas de fusão. Os modelos dos sensores fornecem uma medida da qualidade da saída dos sensores e representam suas incertezas associadas. O

modelo do ambiente armazena informações a respeito do estado do ambiente ao redor do sistema, contendo informações obtidas previamente ou recém adquiridas.

- **Módulo de Fusão de Dados:** composto pelos algoritmos que combinam as informações dependendo do nível de abstração, por exemplo: 1º nível de sinal, 2º nível de *pixel*, 3º nível de atributos, e 4º nível simbólico.

Além dessa estrutura de integração de múltiplos sensores, outras propostas são descritas na literatura [McKERROW, 1996], [ZHANG, 1999]. Contudo, independentemente da estrutura de integração empregada, existem vários métodos e algoritmos que podem ser utilizados no processo de fusão. O fundamental é que o uso de uma maior quantidade de informação pode resultar em uma resposta com maior confiabilidade e exatidão.

3.3

Classificação de Algoritmos de Fusão de Dados

Algoritmos de fusão de dados podem ser classificados em função das características dos algoritmos considerados. Dependendo do tipo de informação a ser fusionada, podem ser divididos em algoritmos de fusão de dados de informação redundante e de informação complementar. Algoritmos de fusão de dados que utilizam informação redundante reduzem a incerteza e acrescentam robustez ao sistema pela fusão de informações oriundas de vários sensores de um mesmo tipo obtidas por um período de tempo. Em algoritmos de fusão de dados que utilizam informações complementares estas são oriundas de diferentes tipos de sensores. O algoritmo é capaz de discriminar situações ambíguas para um único sensor. Algoritmos de fusão de dados podem também ser classificados de acordo com o nível de abstração em que a fusão de dados ocorre. No caso de imagens visuais, por exemplo, a fusão pode ser realizada por sinais, pixels, atributos ou níveis simbólicos [CIRES, 1995], [GEE, 2003].

De forma a classificar os algoritmos de fusão de dados para diferentes níveis de abstração e considerando-se a natureza das informações (redundante e complementar), ilustra-se na Tabela 3.1 uma classificação para os modelos de fusão de dados, adotando-se a ordem crescente do nível de abstração (complexidade) encontrado nos dados submetidos ao processo de fusão. Para cada nível pode ser aplicado um algoritmo em particular para realizar a fusão de dados para múltiplos sensores [DO COUTTO, 2007].

Tabela 3.1. Classificação de algoritmos de fusão de dados em níveis crescentes de complexidade.

Níveis	Características dos dados	Algoritmos de Fusão de Dados
1°	Fusão de dados quantitativos redundantes	Métodos estatísticos para combinação de informações redundantes – <i>Média Ponderada; Filtro de Kalman; Princípio da Máxima Probabilidade; conjuntos fuzzy.</i>
2°	Fusão de dados quantitativos de natureza complementar	Métodos estatísticos para combinação de informações complementares - <i>Teorema de Bayes; Princípio do Máximo a Posteriori; Métodos Inteligentes e Híbridos baseados na extração e ponderação do conhecimento quantitativo.</i>
3°	Fusão de dados qualitativos	Métodos Inteligentes e Híbridos baseados na extração e ponderação do conhecimento qualitativo – <i>Regras de Dempster-Shafer; Eng. do Conhecimento; Funções de Crédito; Variáveis Nebulosas; Intervalos de Confiança.</i>
4°	Fusão de dados quantitativos e qualitativos	Metodologias Altamente Especializadas – <i>Lógica Neural/Nebulosa; Heurísticas Especializadas Híbridas Inteligentes.</i>

3.3.1

Nível 1 – Fusão de Dados Quantitativos Redundantes

Neste caso, os algoritmos de fusão de dados são implementados em nível local, que objetivam aumentar a qualidade de informação transportada pelo dado quantitativo coletado [DO COUTTO, 2007]. A modelagem combinacional neste nível pode estar baseada nas técnicas de média ponderada, princípio de máxima probabilidade, teoria dos conjuntos *fuzzy*, entre outros. Uma descrição mais detalhada destes algoritmos é efetuada na seção 3.4.

3.3.2

Nível 2 – Fusão de Dados Quantitativos Complementares

O uso de mecanismos integradores de dados de 2° nível vem se consolidando com diversas aplicações práticas, devido à grande capacidade de se monitorar, controlar e automatizar sistemas complexos de modo integrado. Tal processo de fusão de dados concorre para o aumento da qualidade da informação pós-processada, obtida a partir da combinação de informações complementares advindas de diversos mecanismos de integração de dados de 1° nível, atuantes no sistema em estudo. Levando em conta as características individuais dos dados heterogêneos coletados, algoritmos de fusão de dados podem ser utilizados baseados em processos estatísticos associados à aplicação do Teorema de Bayes. Estes algoritmos de fusão incluem classificadores estatísticos que medem a probabilidade da informação agregada como um vetor de parâmetros. Estes parâmetros são regidos por probabilidades condicionais. Assim,

utiliza-se a Regra de Bayes para obter a composição da probabilidade *a posteriori*. Aplica-se então uma regra de decisão, tal como o Princípio do Máximo *a Posteriori*, para selecionar a combinação de informações mais apropriada [DO COUTTO, 2007].

3.3.3

Nível 3 – Fusão de Dados Qualitativos

Genericamente, dados de natureza qualitativa podem ser combinados com base no conhecimento extraído de fontes de informações subjetivas (e.g., consultas a especialistas, agentes informantes, sistemas de múltiplos sensores especializados). Vários fatores interferem no processo de coleta de informação, tais como grau de confiabilidade do canal de comunicação (biológico ou tecnológico), convicção, ponto de vista, credibilidade, formação religiosa e cultural e influências externas de outros agentes. Desta forma, faz-se necessário utilizar ferramentas heurísticas inteligentes híbridas de apoio à decisão, destinadas à extração e combinação do conhecimento proveniente de diversas fontes imprecisas de informação.

Na área de inteligência artificial duas técnicas vêm se consagrando como heurísticas destinadas à elaboração de modelos quantitativos para a extração/combinação do conhecimento de fontes qualitativas de informação imprecisa: sistemas especialistas e modelos de credibilidade baseados na modelagem de sistemas multi-agentes inteligentes. Cada uma dessas técnicas tem seu domínio de aplicação bem definido, relacionado ao tipo de mapeamento funcional que se deseja utilizar para extrair dos dados qualitativos disponíveis o conhecimento a eles agregado [DO COUTTO, 2007].

3.3.4

Nível 4 – Fusão de Dados Quantitativos e Qualitativos

Processos de integração de dados numéricos (informação de natureza quantitativa) com dados lingüísticos (informação de natureza qualitativa) compõem o que aqui se denomina fusão de dados de 4º nível. Atuando como sistemas integradores de mais alto nível, podem apresentar como elementos de entrada dados provenientes de sistemas de fusão de 1º, 2º e 3º níveis [DO COUTTO, 2007].

Na Figura 3.2 apresenta-se um diagrama esquemático da arquitetura de um sistema de integração de informações de 4º nível. Um mecanismo integrador de informações heterogêneas recebe, como entrada, dados provenientes de uma base de dados composta por:

- Dados quantitativos redundantes, procedente de mecanismos de fusão de 1º nível;
- Dados quantitativos complementares, procedente de mecanismos de fusão de 2º nível;

- Dados qualitativos, procedente de mecanismos de fusão de 3º nível.

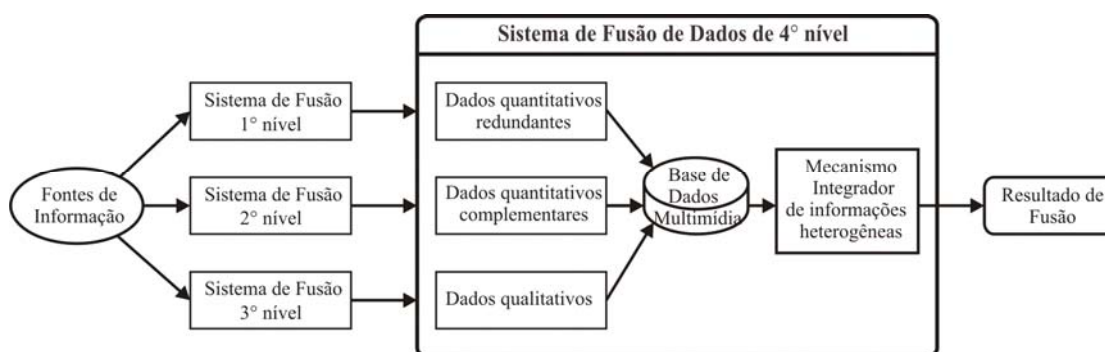


Figura 3.2. Diagrama de um sistema genérico de fusão de dados de 4º nível.

Esse mecanismo de fusão de informações heterogêneas promove um processo de agregação e combinação das informações de entrada de uma maneira otimizada, de modo a gerar como saída um conjunto de informações que apresente maior credibilidade, confiabilidade e nível de incerteza minimizado [DO COUTTO, 2007].

3.4

Algoritmos de fusão de dados quantitativos redundantes aplicados a instrumentação e medição

Visando a projetar um sistema de instrumentação que utiliza o procedimento de medição baseado em algoritmos de fusão de dados quantitativos redundantes, em seguida são apresentados quatro algoritmos que serão utilizados neste trabalho. O primeiro algoritmo (a) *estimação de máxima probabilidade* considera que as variáveis de informação são influenciadas unicamente por incertezas aleatórias, isto é, elas são representadas por funções de densidade de probabilidades. O segundo e terceiro algoritmos são (b) *fusão de dados utilizando relações de compatibilidade* e (c) *fusão de dados utilizando operadores OWA (Ordered Weighted Average)*. O quarto algoritmo, (d) *agregação parcial* considera a agregação parcial das fontes de informação utilizando os procedimentos de fusão combinada de relações de compatibilidade e operadores OWA.

3.4.1

Fusão de dados utilizando o algoritmo de estimação de máxima probabilidade (MLE)

O método de fusão baseado no algoritmo de estimação de máxima probabilidade consiste em analisar a função de probabilidade (*likelihood function*) de um conjunto de

amostras redundantes quantitativas e estimar os parâmetros que maximizem esta função. Este algoritmo está baseado na combinação da evidência contida no sinal com o conhecimento *a priori* da distribuição de densidade de probabilidade do processo. O estimador de máxima probabilidade toma como entrada o conjunto de observações das múltiplas fontes de informação e, usando um modelo probabilístico do processo (e.g., Gaussiano), estima os parâmetros desconhecidos. A exatidão da estimação depende da informação disponível e da eficiência do estimador [MAYBECK, 1997], [DHILLON, 1999] e [VASEGHI, 2000].

Para este propósito, considera-se um conjunto de amostras de tamanho n , definidas por x_1, \dots, x_n , com funções de densidade de probabilidade independentes. A função de probabilidade (*likelihood function*) do conjunto é dado por [MAYBECK, 1997] [DHILLON, 1999]:

$$K = f_1(\bar{x})f_2(\bar{x})\dots f_n(\bar{x}) \quad (3.1)$$

em que f_1, \dots, f_n são as funções de densidade de probabilidade e \bar{x} é o melhor valor estimado que maximiza K .

Considera-se que as funções de densidade de probabilidade têm uma distribuição normal com parâmetros conhecidos, dada por:

$$f(x) = \frac{1}{u\sqrt{2\pi}} e^{-[(x-\bar{x})^2/2u^2]} \quad (3.2)$$

em que u é o desvio padrão de x .

Substituindo (3.2) em (3.1), tem-se:

$$K = \prod_{i=1}^n \frac{1}{u_i\sqrt{2\pi}} e^{-[(\bar{x}-x_i)^2/2u_i^2]} \quad (3.3)$$

Tomando o logaritmo de (3.3):

$$\ln K = \ln\left[\frac{1}{\sqrt{2\pi}^n \prod_{i=1}^n u_i}\right] - \sum_{i=1}^n \frac{(\bar{x} - x_i)^2}{2u_i^2} \quad (3.4)$$

O valor de \bar{x} que maximiza (3.4) é conhecido como a estimação de máxima probabilidade, e pode ser calculado por:

$$\frac{\partial \ln K}{\partial \bar{x}} = 0 \quad (3.5)$$

Resolvendo (3.5) para \bar{x} , tem-se:

$$\bar{x} = \frac{1}{\Gamma u_1^2} x_1 + \frac{1}{\Gamma u_2^2} x_2 + \dots + \frac{1}{\Gamma u_n^2} x_n \quad (3.6)$$

$$\text{com } \Gamma = \left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \dots + \frac{1}{u_n^2} \right).$$

A equação (3.6) define um procedimento de fusão de dados para estimar o valor de fusão $x_F = \bar{x}$, a partir dos n valores medidos x_1, x_2, \dots, x_n , levando em consideração os valores dos desvios padrão u_1, u_2, \dots, u_n de cada medida, respectivamente.

Para exemplificar este procedimento, considera-se que uma grandeza x é medida utilizando-se três sensores diferentes, em que o primeiro sensor indica o valor da medida de $0,9x$ com uma incerteza σ , o segundo, o valor de $1,2x$ com uma incerteza 2σ e o terceiro, o valor de $1,15x$ com uma incerteza $1,5\sigma$. Utilizando o procedimento de fusão de máxima probabilidade (3.6), tem-se:

$$\Gamma = \frac{1}{\sigma^2} + \frac{1}{(2\sigma)^2} + \frac{1}{(1,5\sigma)^2} = \frac{1,6944}{\sigma^2} \quad (3.7)$$

$$x_F = \frac{0,9x}{\Gamma\sigma^2} + \frac{1,2x}{\Gamma(2\sigma)^2} + \frac{1,15x}{\Gamma(1,5\sigma)^2} = 1,01x \quad (3.8)$$

Observa-se que o resultado final de fusão $x_F = 1,01x$ deve apresentar uma menor incerteza do que as medições dos sensores individuais, questão que será abordada nas seções posteriores.

3.4.2

Fusão de dados utilizando relações de compatibilidade

A segunda estratégia de fusão de dados utiliza a teoria de conjuntos *fuzzy* para representar as informações individuais. Este algoritmo consiste em construir mapeamentos individuais para cada informação e integrá-los usando uma estimacão ponderada. Para este propósito, empregam-se relações de compatibilidade entre os elementos que participam do processo de fusão. Para um sistema de medição em particular, são realizadas n medições definidas por x_1, x_2, \dots, x_n . O procedimento de fusão de dados combina estas informações através de uma função de fusão definida por F , determinando-se o valor, \bar{x} , que é a melhor estimacão obtida pelo procedimento de fusão:

$$\bar{x} = F(x_1, x_2, \dots, x_n) \quad (3.9)$$

A combinação de informações por meio de F depende dos elementos a serem fundidos, isto é, deve existir uma compatibilidade entre as fontes de informação para se aplicar o procedimento de fusão. Portanto, define-se uma relação de compatibilidade r , a qual reduz o grau de pertinência associado a pares que são discrepantes e favorece o grau de pertinência

entre elementos compatíveis [DUBOIS, 1998] e [YAGER, 1996]. Para duas medições x_1 e x_2 , representadas por meio de números *fuzzy* X_1 e X_2 (como visto na Seção 2.5), a relação de compatibilidade dada por $r(X_1, X_2)$ depende da distância entre elas, tendo um comportamento decrescente quando a distância $|X_1 - X_2|$ aumenta. Consequentemente, define-se um conjunto *fuzzy* simétrico M , tal que $M(X) = M(-X)$, $M(0) = 1$ e:

$$r(X_1, X_2) = M(X_1 - X_2) \quad (3.10)$$

Assim, o mapeamento dos números *fuzzy* X_1 e X_2 através de r é dado por:

$$\{X_1\}_r = X_1 \oplus M \quad (3.11)$$

$$\{X_2\}_r = X_2 \oplus M \quad (3.12)$$

em que $\{.\}_r$ define o mapeamento e \oplus é a soma *fuzzy*. Generalizando a relação de compatibilidade r , para n medições, tem-se:

$$\mathbf{R}(X_1, \dots, X_n) = \min_{ij} r(X_i, X_j) \quad (3.13)$$

em que \mathbf{R} é a mínima compatibilidade entre todos os pares $r(X_i, X_j)$. O mapeamento do número *fuzzy* X_i através de \mathbf{R} é dado por:

$$\{X_i\}_{\mathbf{R}} = X_i \cap \left[\bigcap_{j \neq i} X_j \oplus \mathbf{R} \right] \quad (3.14)$$

em que $\{X_i\}_{\mathbf{R}}$ representa o mapeamento de X_i através de \mathbf{R} , relacionando a compatibilidade com as medições X_j , e \cap é o operador de intersecção [VILLANUEVA, 2007b].

Para o cálculo da relação de compatibilidade \mathbf{R} , dada em (3.13), utiliza-se uma medida de distância generalizada para números *fuzzy*, em que o resultado da distância entre dois números *fuzzy* é outro número *fuzzy* [VOXMAN, 1998].

Para dois números *fuzzy* X_1 e X_2 , com representação α -cut da forma $[X_{L1}^\alpha, X_{R1}^\alpha]$ e $[X_{L2}^\alpha, X_{R2}^\alpha]$, respectivamente, a distância entre X_1 e X_2 é definida por $M^\alpha = [M_L^\alpha, M_R^\alpha]$ e dada por:

$$M_L^{(\alpha)} = \begin{cases} \max \{ X_{L2}^{(\alpha)} - X_{R1}^{(\alpha)}, 0 \}; & \frac{X_{L1}^{(\alpha)} + X_{R1}^{(\alpha)}}{2} \leq \frac{X_{L2}^{(\alpha)} + X_{R2}^{(\alpha)}}{2} \\ \max \{ X_{L1}^{(\alpha)} - X_{R2}^{(\alpha)}, 0 \}; & \frac{X_{L2}^{(\alpha)} + X_{R2}^{(\alpha)}}{2} \leq \frac{X_{L1}^{(\alpha)} + X_{R1}^{(\alpha)}}{2} \end{cases} \quad (3.15)$$

$$M_R^{(\alpha)} = \max \{ X_{R1}^{(\alpha)} - X_{L2}^{(\alpha)}, X_{R2}^{(\alpha)} - X_{L1}^{(\alpha)} \} \quad (3.16)$$

Para exemplificar este procedimento, considera-se a fusão de três números *fuzzy* definidos por: $X_1 = [1; 6; 11]$, $X_2 = [3; 8; 13]$ e $X_3 = [10; 15; 20]$. Inicialmente, determina-se a relação de compatibilidade \mathbf{R} através de (3.13), em que os valores de $r(X_i, X_j)$ são determinados através de (3.10).

$$\mathbf{R}(X_1, X_2, X_3) = \min\{r(X_1, X_2), r(X_1, X_3), r(X_2, X_3)\} \quad (3.17)$$

$$\mathbf{R}(X_1, X_2, X_3) = \min\{2,84; 7,04; 9\} \quad (3.18)$$

$$\mathbf{R}(X_1, X_2, X_3) = 2,84 \quad (3.19)$$

Portanto, a relação de compatibilidade é definido por um conjunto *fuzzy* simétrico centrado em zero dado por $\mathbf{R} = [-2,84; 0; 2,84]$. Em seguida realiza-se o mapeamento, por meio de (3.14):

$$\{X_1\}_R = X_1 \cap [X_2 \oplus \mathbf{R}] \cap [X_3 \oplus \mathbf{R}] \quad (3.20)$$

$$\{X_2\}_R = X_2 \cap [X_1 \oplus \mathbf{R}] \cap [X_3 \oplus \mathbf{R}] \quad (3.21)$$

$$\{X_3\}_R = X_3 \cap [X_1 \oplus \mathbf{R}] \cap [X_2 \oplus \mathbf{R}] \quad (3.22)$$

Definindo-se a função de fusão F_R como a média aritmética das medições, tem-se:

$$F_R = \frac{\{X_1\}_R + \{X_2\}_R + \{X_3\}_R}{3} \quad (3.23)$$

Na Figura 3.3 apresentam-se os números *fuzzy* X_1 , X_2 e X_3 e o resultado da fusão F_R utilizando a relação de compatibilidade \mathbf{R} .

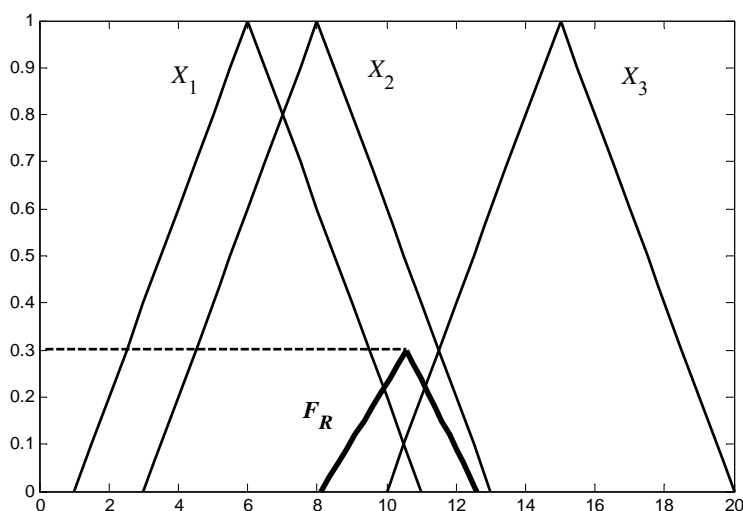


Figura 3.3. Fusão dos números *fuzzy* X_1 , X_2 e X_3 utilizando relações de compatibilidade.

Dois aspectos importantes devem ser observados na função de pertinência do resultado de fusão F_R : o suporte e altura (maior grau de pertinência). O suporte de F_R , definido pelo intervalo $[8,1; 10,57; 12,6]$, indica o resultado de fusão dos três números *fuzzy*. A altura final de F_R indica que o resultado de fusão foi fortemente afetado pela incompatibilidade dos números *fuzzy* fundidos, especialmente entre X_1 e X_3 , isto é, o grau de pertinência do resultado de fusão encontra-se relacionado à incompatibilidade dos elementos fundidos evidenciado ao

realizar o mapeamento de cada elemento de entrada, X_1 , X_2 e X_3 , com a relação de compatibilidade R , avaliados por meio de (3.20), (3.21) e (3.22).

Observa-se que, embora a função de fusão F_R seja definida como a média aritmética das entradas, ela leva em consideração as relações de compatibilidade entre as medições que participam no processo de fusão.

3.4.3

Fusão de dados utilizando operadores OWA

O conceito de operadores OWA (*Ordered Weighted Average*) baseia-se na agregação ou fusão de elementos em que o resultado da operação está contido no intervalo correspondente ao valor mínimo e máximo dos elementos. Os operadores OWA permitem a implementação de diferentes funções de fusão, atribuindo, através de pesos, ênfase aos elementos a serem fundidos. Assim, ao operador OWA de dimensão n é associado um vetor de pesos W , dado por [YAGER, 1996], [PEDRYCZ, 2007]:

$$W = [w_1, \dots, w_n]' \quad \text{com } w_i \in [0, 1] \quad \text{e} \quad \sum_{i=1}^n w_i = 1 \quad (3.24)$$

$$F(x_1, \dots, x_n) = y_1 w_1 + y_2 w_2 + \dots + y_n w_n \quad (3.25)$$

em que o vetor $\langle y_1, \dots, y_n \rangle$ é um ordenamento dos elementos do vetor $\langle x_1, \dots, x_n \rangle$ do maior ao menor valor.

Um aspecto fundamental destes tipos de operadores está na realização do procedimento de ordenamento. Deste modo, uma variável x_i não está associada a um peso w_i , mas um peso w_i está associado a uma posição em particular do vetor ordenado de $\langle x_1, \dots, x_n \rangle$.

Na função de fusão F , a escolha do vetor de pesos W permite a agregação ponderada das funções de pertinência dos números *fuzzy*. Os pesos do operador OWA são determinados como a solução de um problema de programação matemática [O'HAGAN, 1990], definido por:

$$\text{Minimizar } \sum_{i=1}^n w_i \ln w_i \quad (3.26)$$

$$\text{sujeito a: } \sum_{i=1}^n w_i h_n(i) = \varepsilon, \quad \text{com: } h_n(i) = \frac{n-i}{n-1} \quad \text{e} \quad \sum_{i=1}^n w_i = 1, w_i \geq 0 \quad (3.27)$$

em que $\varepsilon \in [0, 1]$ define o grau favorável de associação entre os números *fuzzy*.

Como uma extensão do procedimento de fusão usando relações de compatibilidade, em seguida apresenta-se um procedimento que combina as operações de relações de compatibilidade e operadores OWA.

Para um conjunto de medições $\{x_1, \dots, x_n\}$, cada elemento x_i , representado por um número *fuzzy* X_i é mapeado através da relação de compatibilidade $\{X_i\}_R$ definido em (3.14). Em seguida os números *fuzzy* $\{X_i\}_R$ são ordenados do maior ao menor, fornecendo o novo conjunto $\{X_j\}_R$, em que a cada valor de $\{X_j\}_R$ é associado um peso de ponderação w_j , dado por:

$$\{X_j\}_W = \{X_j\}_R \cdot w_j \quad (3.28)$$

em que $\{X_j\}_W$ representa o mapeamento de $\{X_j\}_R$ através do peso de ponderação w_j .

Para exemplificar este procedimento, considera-se a fusão de três números *fuzzy* definidos por: $X_1 = [1; 6; 11]$, $X_2 = [3; 8; 13]$ e $X_3 = [10; 15; 20]$, com relação de compatibilidade $R = [-2,84; 0; 2,84]$. Em seguida realiza-se o mapeamento, por meio de (3.14):

$$\{X_1\}_R = X_1 \cap [X_2 \oplus R] \cap [X_3 \oplus R] \quad (3.29)$$

$$\{X_2\}_R = X_2 \cap [X_1 \oplus R] \cap [X_3 \oplus R] \quad (3.30)$$

$$\{X_3\}_R = X_3 \cap [X_1 \oplus R] \cap [X_2 \oplus R] \quad (3.31)$$

Na Figura 3.4 ilustram-se os resultados destes mapeamentos. A partir destes resultados, realiza-se um ordenamento de maior a menor valor dos conjuntos *fuzzy* $\{X_1\}_R$, $\{X_2\}_R$ e $\{X_3\}_R$. A avaliação deste ordenamento baseia-se na determinação do valor meio do suporte de cada conjunto *fuzzy*. Seguindo este critério, pode-se calcular e observar que $\{X_3\}_R > \{X_2\}_R > \{X_1\}_R$.

A partir deste resultado e seguindo o procedimento de fusão utilizando o operador OWA, estes três resultados são ordenados do maior ao menor e realiza-se o mapeamento com os pesos de ponderação, dados por:

$$\{X_1\}_W = \{X_3\}_R \cdot w_1 \quad (3.32)$$

$$\{X_2\}_W = \{X_2\}_R \cdot w_2 \quad (3.33)$$

$$\{X_3\}_W = \{X_1\}_R \cdot w_3 \quad (3.34)$$

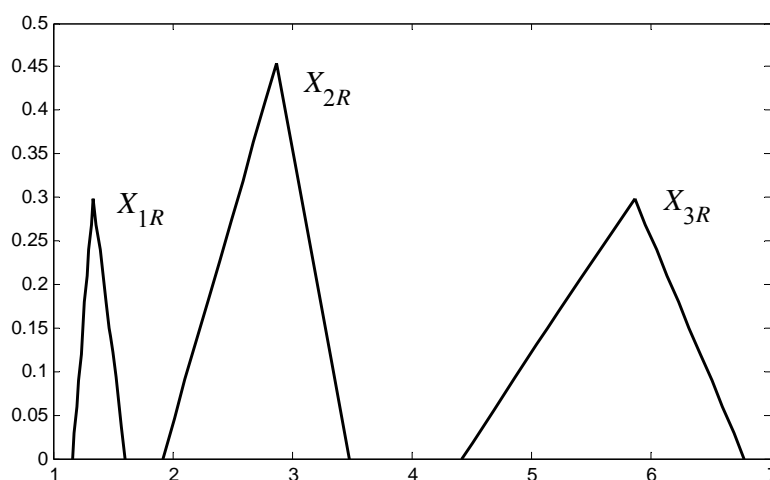


Figura 3.4. Mapeamento utilizando relações de compatibilidade: $\{X_1\}_R$, $\{X_2\}_R$ e $\{X_3\}_R$.

A partir de (3.26) e considerando um grau favorável de associação entre os números *fuzzy* $\varepsilon = 0,25$, são calculados os pesos de ponderação: $w_1 = 0,12$, $w_2 = 0,26$ e $w_3 = 0,62$. Finalmente a função de fusão baseada nas relações de compatibilidade e operadores OWA é dada por:

$$F_{ROWA} = \{X_1\}_W + \{X_2\}_W + \{X_3\}_W \quad (3.35)$$

Na Figura 3.5 apresentam-se os números *fuzzy* X_1 , X_2 e X_3 e o resultado da fusão F_{ROWA} . Observa-se a influência do ordenamento do operador OWA no resultado final de fusão: a função de pertinência final encontra-se localizada próxima das fontes de entradas compatíveis (X_1 e X_2) e afastada da entrada em conflito (X_3).

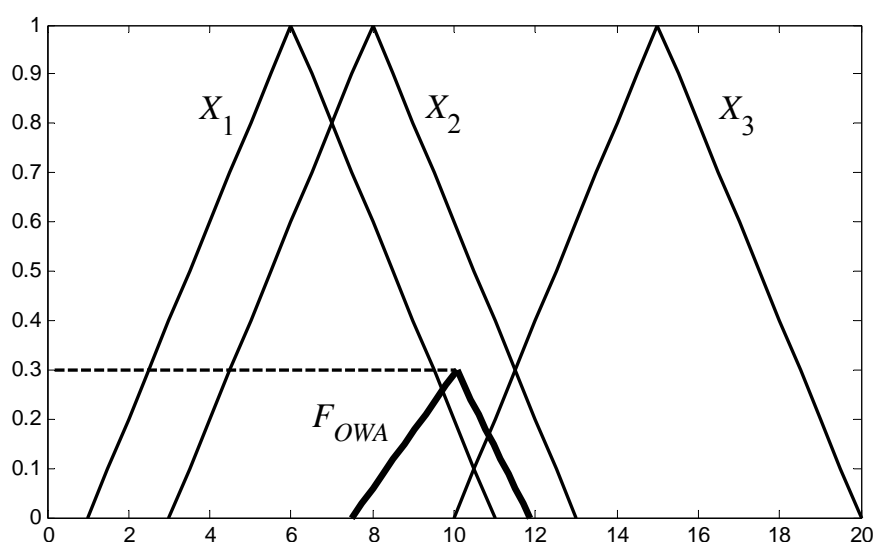


Figura 3.5. Fusão dos números *fuzzy* X_1 , X_2 e X_3 utilizando relações de compatibilidade e operadores OWA.

3.4.4

Fusão de dados baseada na Agregação Parcial das Informações

Nos dois métodos de fusão apresentados anteriormente (relações de compatibilidade e operadores OWA), o resultado da fusão requer a combinação de todas as fontes de entrada. Entretanto, existem situações em que as informações de entrada são defeituosas, pois elas podem se encontrar fora da faixa permissível de variação, com ruído ou operação incorreta. Deste modo, define-se um procedimento de fusão que se baseia na agregação parcial das fontes, isto é, utilizando-se um subconjunto das fontes de informação disponíveis [YAGER, 1996].

Para este propósito, para um conjunto de informações A define-se uma medida de confiança $S: A \rightarrow [0, 1]$, tal que:

- $S(\phi) = 0$;
- $S(A) = 1$;
- $\forall (x, y) \in A$, se $X \subset Y$ então $S(X) \leq S(Y)$.

Considera-se que o conjunto de informações de entrada é definido por $A = \{A_1, A_2, \dots, A_n\}$ e um subconjunto de A é definido por $T = \{A_1, A_2, \dots, A_m\}$ com $m \leq n$. Para este subconjunto T , a medida de confiança $S(T)$ representará o grau de confiabilidade que o usuário atribui aos subconjuntos a fim de calcular o resultado final de fusão.

Um caso especial da medida S é definido por:

$$\forall T \in 2^A, S(T) = 1 \quad (3.36)$$

A equação (3.36) expressa que cada subconjunto T é um grupo representativo para o cálculo do resultado de fusão.

Deste modo, para cada subconjunto T , podem-se calcular os resultados parciais de fusão *fuzzy* utilizando o procedimento de fusão baseado nas relações de compatibilidade e operadores OWA como descrito nas seções anteriores.

A partir dos resultados parciais de fusão, realiza-se a agregação final utilizando a medida de confiança S , determinando a fusão final $F_{ROWAparcial}$:

$$F_{ROWAparcial} = \max_{T \in 2^A} \{F_{ROWA}(T) \cap S(T)\} \quad (3.37)$$

em que o operador max corresponde à união *fuzzy*.

Portanto, observa-se que a medida *fuzzy* S pode ser facilmente incluída no processo de fusão, permitindo obter uma saída agregada considerando um subconjunto das fontes de informação.

Para exemplificar este procedimento, considera-se a fusão de três números *fuzzy* definidos por: $X_1 = [1; 6; 11]$, $X_2 = [3; 8; 13]$ e $X_3 = [10; 15; 20]$, com relação de compatibilidade $R = [-2,840; 0; 2,84]$ e com grau de associação entre os números *fuzzy* $\varepsilon = 0,25$. Considerando-se o procedimento de agregação parcial, define-se uma medida de confiança S dada por:

$$\forall T \in A, S(T) = \begin{cases} 0 & \text{se } |T| < 2 \text{ (}|T| = 0 \text{ ou } |T| = 1) \\ 1 & \text{se } |T| \geq 2 \text{ (}|T| = 2 \text{ ou } |T| = 3) \end{cases} \quad (3.38)$$

Esta medida atribui uma confiança total para qualquer subconjunto de pelo menos duas entradas *fuzzy*, isto é, em que pelo menos duas entradas estão corretas.

Em seguida são determinadas as agregações parciais utilizando a relação de compatibilidade e operadores OWA dos subconjuntos: $F_{ROWA1} = \{X_1, X_2\}$, $F_{ROWA2} = \{X_1, X_3\}$, $F_{ROWA3} = \{X_2, X_3\}$ e $F_{ROWA4} = \{X_1, X_2, X_3\}$.

Na Figura 3.6 ilustra-se o resultado final de fusão, $F_{ROWA\text{parcial}}$, obtida a partir da agregação dos resultados parciais utilizando (3.37).

$$F_{ROWA\text{parcial}} = F_{ROWA1} \cup F_{ROWA2} \cup F_{ROWA3} \cup F_{ROWA4} \quad (3.39)$$

Observa-se neste resultado que o valor do máximo grau de pertinência é muito maior quando comparado com o procedimento de fusão utilizando somente relações de compatibilidade e operadores OWA (Figura 3.5), pois neste procedimento considera-se que pelo menos duas entradas são compatíveis (X_1 e X_2). Observa-se também que a fonte de entrada em conflito (X_3) contribui para a função de pertinência do resultado final de fusão (partes baixas), pois este procedimento, embora diminua a contribuição da entrada conflitante, não a elimina no procedimento de fusão.

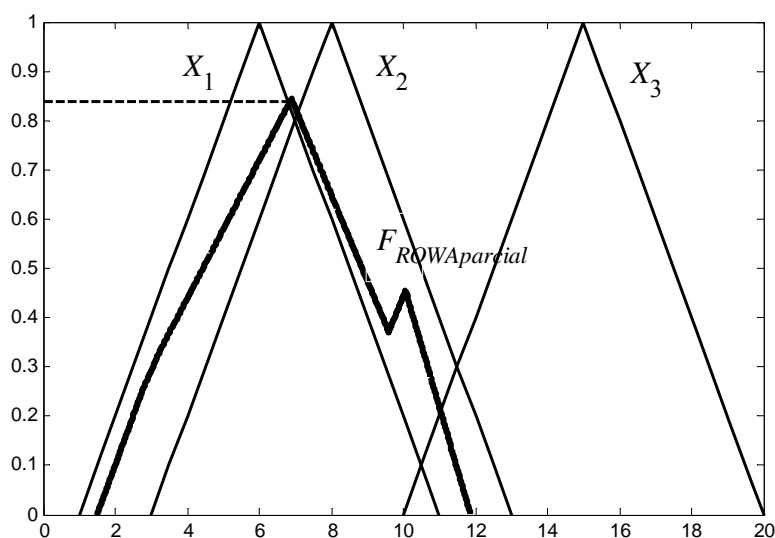


Figura 3.6. Fusão dos números *fuzzy* X_1 , X_2 e X_3 utilizando relações de compatibilidade, operadores OWA e agregação parcial.

3.5

Análise de Incertezas dos Resultados de Fusão

Conforme o guia GUM (*Guide to the expression of Uncertainty in Measurement*), os componentes de incertezas de medidas são categorizados, de acordo com o método de avaliação, em incertezas do tipo A e do tipo B. No A, os componentes das incertezas são avaliados a partir das distribuições estatísticas dos resultados de uma série de medidas e podem ser caracterizados através do desvio padrão experimental. Na avaliação de incertezas do tipo B, as incertezas podem ser avaliadas por métodos não convencionais, como, por exemplo,

utilizando a teoria de conjuntos *fuzzy*, quando o intervalo de confiança acerca dos resultados de medição é conhecido [GUM, 1993], [FERRERO, 2006].

Estuda-se, nesta seção, a avaliação das incertezas dos resultados finais de fusão quando aplicados os métodos de fusão quantitativos redundantes, baseados na estimação de máxima probabilidade (Tipo A) e considerando o resultado final uma função de pertinência *fuzzy* como resultado do processo de fusão utilizando relações de compatibilidade, operador OWA e agregação parcial (Tipo B).

3.5.1

Análise de Incertezas Tipo A

Considerando-se o resultado de fusão x_F utilizando a técnica de estimação de máxima probabilidade (Eq. 3.6), é realizada a análise da propagação de incertezas associadas às grandezas medidas x_1, x_2, \dots, x_n , [TAYLOR, 1997], para determinar a incerteza final do processo de fusão u_{x_F} .

$$x_F = \frac{1}{\Gamma u_{x_1}^2} x_1 + \frac{1}{\Gamma u_{x_2}^2} x_2 + \dots + \frac{1}{\Gamma u_{x_n}^2} x_n \quad \text{com} \quad \Gamma = \left(\frac{1}{u_{x_1}^2} + \frac{1}{u_{x_2}^2} + \dots + \frac{1}{u_{x_n}^2} \right) \quad (3.40)$$

$$u_{x_F}^2 = \left(\frac{\partial x_F}{\partial x_1} u_{x_1} \right)^2 + \left(\frac{\partial x_F}{\partial x_2} u_{x_2} \right)^2 + \dots + \left(\frac{\partial x_F}{\partial x_n} u_{x_n} \right)^2 \quad (3.41)$$

$$u_{x_F}^2 = \left(\frac{1}{\Gamma u_{x_1}^2} u_{x_1} \right)^2 + \left(\frac{1}{\Gamma u_{x_2}^2} u_{x_2} \right)^2 + \dots + \left(\frac{1}{\Gamma u_{x_n}^2} u_{x_n} \right)^2 \quad (3.42)$$

$$u_{x_F}^2 = \frac{1}{\Gamma^2} \left(\frac{1}{u_{x_1}^2} + \frac{1}{u_{x_2}^2} + \dots + \frac{1}{u_{x_n}^2} \right) \quad (3.43)$$

$$u_{x_F}^2 = \frac{1}{\Gamma} = \frac{1}{u_{x_1}^2} + \frac{1}{u_{x_2}^2} + \dots + \frac{1}{u_{x_n}^2} \quad (3.44)$$

Finalmente, a incerteza final u_{x_F} associada a x_F é calculada como:

$$\frac{1}{u_{x_F}^2} = \frac{1}{u_{x_1}^2} + \frac{1}{u_{x_2}^2} + \dots + \frac{1}{u_{x_n}^2} \quad (3.45)$$

Pode-se observar que a incerteza final u_{x_F} é menor do que as incertezas independentes $u_{x_1}, u_{x_2}, \dots, u_{x_n}$, isto é, o resultado da fusão apresenta a menor incerteza possível.

Para exemplificar este procedimento, considera-se que uma grandeza é medida utilizando-se dois sensores diferentes, em que o primeiro sensor tem uma incerteza σ e o segundo, uma incerteza 2σ . Utilizando o procedimento de fusão de máxima probabilidade, tem-se uma incerteza final de fusão de $0,89\sigma$, com um nível de confiança de 68%, como ilustrado na Figura 3.7.

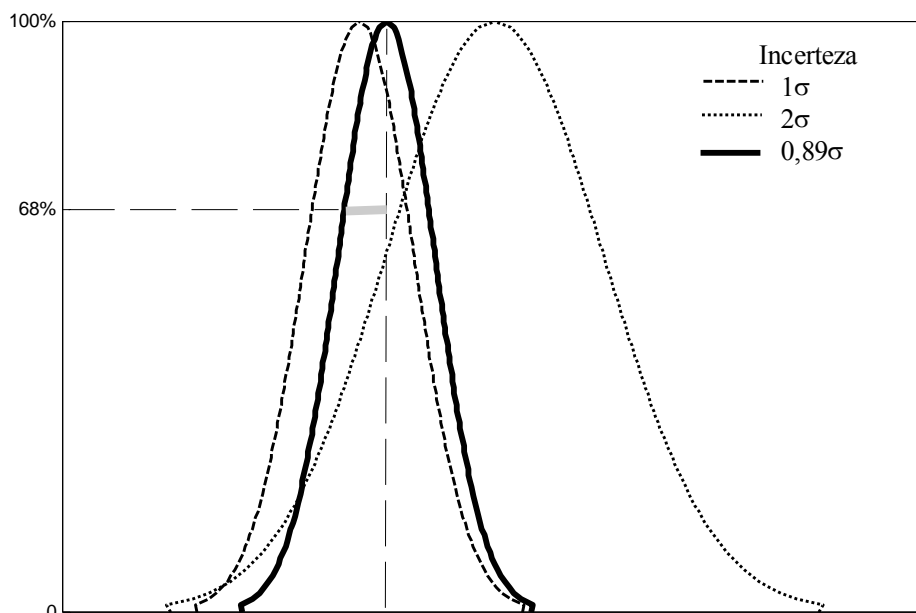


Figura 3.7. Diminuição da incerteza final ($0,89\sigma$) utilizando o procedimento de fusão de estimação de máxima probabilidade (MLE).

3.5.2

Análise de Incertezas Tipo B

De acordo com o guia para a expressão de incertezas (GUM), uma incerteza do tipo B é frequentemente dada em forma de intervalos correspondentes a níveis de confiança. Portanto, para representar as incertezas da medição, pode-se utilizar o cálculo intervalar baseado na teoria de conjuntos *fuzzy*. Esta estratégia é uma boa alternativa para a avaliação convencional de incertezas, sendo compatível com o GUM [GUM, 1993], [MAURIS, 2000] e [MAURIS, 2001].

Em um sistema de instrumentação baseado na teoria de conjuntos *fuzzy*, o resultado final da medição é expresso por uma função de pertinência derivada da realização de operações aritméticas básicas (soma, diferença, multiplicação e divisão), interseção e união. A forma desta função de pertinência, $\mu_A(x)$, pode ser dos tipos normal, subnormal, convexo e não-convexo, como ilustrado na Figura 3.8.

A fim de caracterizar o resultado da medição por um único valor e sua incerteza associada, devem ser definidos parâmetros de avaliação da função de pertinência final. Para este propósito, deve-se determinar o valor médio e o desvio padrão de uma forma geral, isto é, considerando-se a forma da função de pertinência (normal, subnormal, convexo, não-convexo) [LALLA, 2008].

Para uma função de pertinência definida por A , o valor médio pode ser definido por [BODJANOVA, 2006], [MENDEL, 2007]:

$$\bar{x} = \frac{\sum_{i=1}^N x_i \mu_A(x_i)}{\sum_{i=1}^N \mu_A(x_i)} \quad (3.46)$$

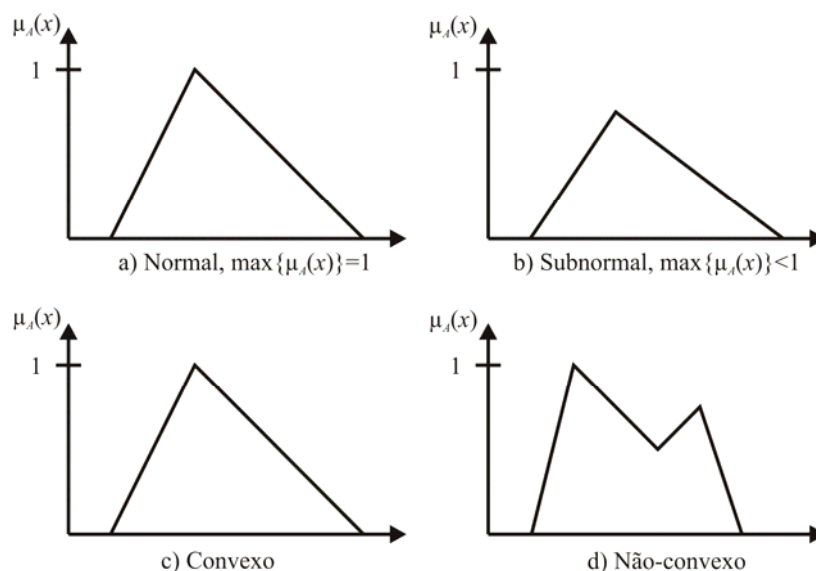


Figura 3.8. Funções de pertinência normal, subnormal, convexo e não convexo.

A variância é definida como o valor esperado do desvio quadrático entre o valor médio e os elementos do suporte de A , definido por:

$$\text{var} = \frac{1}{N} \sum_{i=1}^N [\bar{x} - x_i]^2 \mu_A(x_i) \quad (3.47)$$

Finalmente, o desvio padrão de A é definido por:

$$\sigma = \text{var}^{1/2} \quad (3.48)$$

A relação entre o valor médio (centróide) e o desvio padrão de A determina o intervalo de confiança, definido por: $[\bar{x} - \sigma ; \bar{x} + \sigma]$, como ilustrado na Figura 3.9.

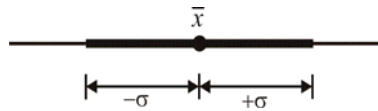


Figura 3.9. Intervalo de confiança da medição.

3.6

Conclusão

Entre as principais conclusões deste capítulo estão:

- A teoria de conjuntos *fuzzy* pode ser utilizada para operar com medidas de incertezas de uma maneira compatível com o guia de expressão e incertezas (GUM). Esta abordagem é particularmente interessante para medições com incertezas, em que não se têm as distribuições de probabilidade.
- Foram apresentadas técnicas de fusão de dados aplicadas à área de instrumentação, especificamente: (a) estimação de máxima probabilidade, (b) relações de compatibilidade, (c) operadores OWA (*Ordered Weighted Average*) e (d) fusão *fuzzy* baseada nas relações de compatibilidade e operadores OWA com agregação parcial.