

Technischer Bericht

Auftraggeber: SFB 627 (Nexus)

Titel: **Metriken für Qualitätskennzahlen zur
Degradierung von skalaren Sensordaten**

Verfasser: Uwe-Philipp Käppeler

Version: 1.0

Datum: 30.11.2009

CR-Klassifikation: H.2.1

Inhaltsverzeichnis

Kurzfassung.....	2
1. Einleitung.....	3
1.1. Domänen der Degradierung von skalaren Sensordaten.....	3
1.2. Sensorbezogene Metriken für Datenqualität.....	3
1.3. Degradierung von skalaren Kontext Informationen.....	4
2. Metriken zu Qualitätskennzahlen.....	6
2.1. Metriken für einseitig unendliche Wertebereiche von Degradierung.....	6
2.2. Metriken für abgeschlossen definierte Wertebereiche von Degradierung.....	7
2.3. Metriken für Degradierung bezüglich abgeschlossen definierten Gültigkeitsbereichen.....	8
3. Qualitätskennzahlmetriken für einzelne Domänen der Degradierung.....	10
3.1. Domäne Empirie.....	10
3.2. Domäne Querempfindlichkeit.....	11
3.3. Domäne Signalverarbeitung.....	11
3.4. Domäne Zuverlässigkeit.....	11
3.5. Domäne Temporale Aspekte.....	12
4. Überwachung der Datenqualität für einzelne Anwendungen.....	13
Referenzen.....	14

Kurzfassung

Dieser Bericht definiert die Bestimmung von Qualitätskennzahlen zur Degradierung von skalaren Sensordaten. Es werden die Metriken zur Degradierung und deren Semantik im Detail beschrieben. Dieser Bericht baut auf den in [Käppeler 2008] eingeführten Domänen der Degradierung auf.

1. Einleitung

Aus [Käppeler 2008]: Die Sensordaten Integration nach [Käppeler et. al. 2005] und [Käppeler et. al. 2005b] in die Umgebungsmodelle der Nexus Plattform [Nicklas et. al. 2001] soll um Qualitätsaspekte erweitert werden. Dazu sollen neben Inkonsistenzen zwischen einzelnen Sensordaten oder Kontextinformationen Eigenschaften von Sensoren beschrieben werden, welche einen Grad der Degradierung [Dudkowski et. al. 2006] für deren Messdaten, angeben. Diese Informationen werden innerhalb der Metadaten angegeben, welche zu jedem Attribut oder Nexus-Objekt innerhalb der Umgebungsmodelle optional gespeichert werden können [Hönle et. al. 2005].

Darauf aufbauend werden in diesem Bericht Metriken definiert, die Angaben zur Degradierung von Sensordaten auf Qualitätskennzahlen abbilden. Durch die Metrik werden Sensordaten aus verschiedenen Quellen vergleichbar. Die Qualitätskennzahlen sind zunächst unabhängig von einzelnen Anwendungen und können auf die physikalischen Eigenschaften der eingesetzten Sensorik bezogen werden. Dadurch ist es einerseits für beliebige Anwendungen, Auswertungen oder Fusionsalgorithmen möglich, die Qualität der bereitgestellten Sensordaten zu bewerten und zu überwachen. Andererseits sind die beschriebenen Metriken auch auf vorgegebene Qualitätsanforderungen einzelner Anwendungen anwendbar.

Der Bericht wendet sich an Entwickler von kontextbezogenen Anwendungen, die die Nexus Plattform benutzen, als auch an Entwickler, die Dienste der Föderation implementieren.

1.1. Domänen der Degradierung von skalaren Sensordaten

Durch die Definition von Domänen bezüglich Qualitätskriterien können verschiedene Ursachen für Degradierung getrennt voneinander behandelt oder miteinander verknüpft werden. Die fünf Domänen, aus denen die Qualitätsangaben zu skalaren Sensordaten bestehen, sind: **Empirie, Querempfindlichkeit, Signalverarbeitung, Zuverlässigkeit und Temporale Aspekte**.

Jede Qualitätskennzahl bezieht sich auf eine vorgegebene Anforderung an Qualität [IEC2371], es kann im offenen System aber verschiedene Anforderungen an Datenqualität geben. Daher soll es in der Nexus Plattform einzelnen Anwendungen oder Anwendungsentwicklern ermöglicht werden, eigene Metriken für Qualitätskriterien anzugeben. Dazu können die Schlüsselworte der in [Käppeler 2008] definierten Syntax verwendet und zu Qualitätsmetriken verknüpft werden. Die jeweilige Semantik der einzelnen Domänen und eine Abbildung auf Qualitätskennzahlen werden in den folgenden Kapiteln beschrieben.

1.2. Sensorbezogene Metriken für Datenqualität

Um Sensordaten generell mit Qualitätsangaben zu versehen, können Messdaten von einem SensorContextServer [Käppeler et. al. 2005b] mit Qualitätsangaben behaftet werden, die sich zunächst auf die physikalischen Eigenschaften des eingesetzten Sensors oder auf entsprechende Vorgaben des Herstellers beziehen. Damit ist eine Beschreibung der Datenqualität möglich, die unabhängig ist von verschiedenen Auswertungen der Daten in unterschiedlichen Anwendungen.

Wie wird das Qualitätsmass für die Nexus Plattform definiert? Gibt es eine Anforderung an die Qualität, so sollte das Einhalten der Anforderung in der Mitte des Intervalls der Qualitätskennzahlen angegeben werden können. Auf diese Weise ist sowohl das Unterschreiten als auch das Übertreffen der Qualitätsanforderung von der exakten Einhaltung der geforderten Qualität unterscheidbar. Dadurch können einseitig unendliche Intervalle wie

$[0; \infty)$ nicht für die Kennzahl Wertebereiche verwendet werden. Das beidseitig unendliche Intervall $(-\infty; \infty)$ wäre denkbar, die Verarbeitung beliebig großer Zahlen ist für technische Systeme jedoch problematisch. Wird der Wertebereich auf das abgeschlossene Intervall $[0; 1]$ oder das halboffene Intervall $(0; 1]$ festgelegt, so bleiben numerische Probleme der digitalen Datenverarbeitung bestehen, es können jedoch standardisierte Formen der approximativen Darstellung für Reelle Zahlen auf unterschiedlichen Datenverarbeitenden Systemen vorausgesetzt werden um den Wertebereich abzubilden. Hierbei steht die 0 für minimale Qualität oder maximale Degradierung, die 1 beschreibt die maximal erreichbare Qualität. Durch diese Ausrichtung der Semantik der Qualitätskennzahl zur Degradierung ist eine einfachere Kombination mit weiteren Qualitätsmerkmalen wie Angaben zur Konsistenz oder einem Vertrauenswert möglich. Der Nachteil dabei ist, dass zur Berechnung der Qualitätskennzahl meist eine Fehlerkennzahl wie zum Beispiel die Angabe einer Standardabweichung invertiert werden muss. Die beiden Grenzwerte 0 und 1 des Wertebereichs für Qualitätsmaße sind in der Praxis nicht immer erreichbar.

Damit Kennzahlen zu unterschiedlichen Qualitätsangaben vergleichbar und die Entwicklung einer einzelnen Kennzahl semantisch sinnvoll überprüfbar sind, muss die Funktion zur Berechnung der Kennzahl stetig sein. Damit eine Entwicklung der Qualität in unterschiedlichen Bereichen vergleichbar ist, muss die Funktion zur Berechnung ebenfalls auch stetig ableitbar sein.

1.3. Degradierung von skalaren Kontext Informationen

Für die oben aufgeführte, zweite Klasse von Anwendungen und Diensten ist eine einzige, allgemeine Qualitätsangabe zur Degradierung von Kontext ausreichend. Die in diesen Metadaten angegebenen Informationen können oft von den Herstellern der Sensoren, die für die Generierung der Daten verwendet wurden, übernommen werden. Dafür wurden in der Nexus Plattform die Metadaten zu Fehlerangaben definiert. Es können hierbei Angaben zu einem absoluten, relativen oder reduzierten Fehler oder eine Standardabweichung angegeben werden. Die folgenden Erklärungen dieser Fehlerbeschreibungen sind übernommen aus [Hoffmann 2007], [Dietrich 1991], [DIN94], [DIN95] und [DIN99]. Vorab sollen die in Tabelle 1 aufgeführten Begriffsklärungen und Kurzbeschreibungen die wichtigsten, in der Literatur aber nicht immer identisch verwendeten Begriffe zu Messunsicherheiten zusammenfassen. Ausführliche Erklärungen sind in den jeweils angegebenen Quellen zu finden.

Begriff	Kurzbeschreibung	Quellen
Messabweichung, Messfehler	Abweichung vom wahren Wert beim Messen	[Hoffmann 2007]
Wahrer Wert	Unbekannte Größe, die gemessen wird	[DIN95], [DIN99], [Hoffmann 2007]
Richtiger Wert	Vorgegebene, bekannte Größe, die gemessen wird	
Messunsicherheit	Intervall um Schätzgröße oder Messwert	[Dietrich 1991], [DIN95], [Hoffmann 2007]

Tabelle 1: Begriffsklärungen zum Thema Messunsicherheit

1.3.1. Absoluter Fehler

Der absolute Fehler f_{abs} , definiert in Gleichung (1.1), entspricht einer maximal möglichen Messabweichung des gemessenen Wertes x_m vom wahren Wert x_w . Die zugehörige Einheit muss derjenigen entsprechen, die dem beobachteten Phänomen selbst entspricht und wird nicht explizit angegeben.

$$(1.1) \quad f_{abs} = |x_w - x_m|$$

1.3.2. Relativer Fehler

Im Gegensatz zum absoluten Fehler kann mit dem relativen Fehler ein Wert für eine Messabweichung angegeben werden, deren maximal mögliche Größe abhängig ist vom gemessenen Wert selbst. Der relative Fehler f_{rel} ist abhängig vom gemessenen Wert x_m und dem wahren Wert und kann in Prozent angegeben, wie in (1.2) angegeben.

$$(1.2) \quad f_{rel} = \frac{|x_w - x_m|}{x_m} \cdot 100\%$$

Im Gegensatz zum absoluten Fehler, welcher eine (in der Modellierung nicht explizit angegebene, aber dem Messwert entsprechende) Einheit besitzt, ist der relative Fehler grundsätzlich ohne Einheit.

1.3.3. Reduzierter Fehler

Eine weitere Möglichkeit, einen Fehler anzugeben, ist der reduzierte Fehler nach Gleichung (1.3). Hier wird die Abweichung vom wahren Wert nicht in Relation zum Messwert angegeben sondern als Anteil des Messbereichs eines Sensors bzw. des Wertebereichs eines Attributs. Für den reduzierten Fehler f_{red} ergibt sich in Abhängigkeit der unteren und oberen Grenzen x_{max} und x_{min} des Wertebereichs:

$$(1.3) \quad f_{red} = \frac{x_w - x_m}{x_{max} - x_{min}}$$

1.3.4. Standardabweichung eines Sensors

Mit der Angabe einer Standardabweichung ist es möglich, die stochastische Streuung eines Sensors als Normalverteilung anzugeben. Die empirische Standardabweichung σ kann durch n Vergleichsmessungen ermittelt werden:

$$(1.4) \quad \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_w - x_{m,i})^2}$$

In der Praxis wird für den unbekannten, wahren Wert x_w das arithmetische Mittel aller Vergleichsmessungen verwendet. Für die Standardabweichung gilt unter der Annahme einer Normalverteilung, dass sich der mit stochastischem Fehler behaftete Messwert in einem bestimmten Intervall um den wahren Wert beziehungsweise der wahre Wert mit einer gewissen Wahrscheinlichkeit innerhalb eines Intervalls um den gemessenen Wert befindet. Beispiele zu den Angaben einzelner Fehlerarten und zugehörige AWM-Elemente sind in [Käppeler 2008] zu finden.

2. Metriken zu Qualitätskennzahlen

2.1. Metriken für einseitig unendliche Wertebereiche von Degradierung

Die Abbildung eines Fehlers f oder einer Degradierung definiert auf dem rechtsseitig unendlichen Intervall $[0; \infty)$ auf den Wertebereich der Qualitätskennzahl $(0; 1]$, wie sie u.a. für die Bewertung einer Standardabweichung in [Benkmann et. al. 2008] verwendet wird, ist durch die Exponentialfunktion in Gleichung (2.1) gegeben. Es wird der Streckungsfaktor γ aus Gleichung (2.2) verwendet, um den Vorgabewert v , der hier der Qualitätsanforderung entspricht, auf den Kennzahlen-Zielwert η zu normieren, der in unserem Fall mit 0,5 in der Mitte des Wertebereich-Intervalls der Qualitätskennzahl angegeben wird.

$$(2.1) \quad q_1(f) = e^{-\gamma f}$$

$$(2.2) \quad \gamma(s, \eta) = -\frac{\ln(\eta)}{v}$$

Sind keine Anforderungen von Anwendungen gegeben, so wird die physikalische Eigenschaft eines Sensors, wie sie vom Hersteller beschrieben ist, als Vorgabewert auf den entsprechenden Qualitätszielwert 0,5 normiert. Wird von einer Anwendung eine Vorgabe für die einzuhaltende Qualität gegeben, so wird genau dieser Wert mit der Qualitätskennzahl 0,5 bewertet. Beispielhaft für verschiedene Qualitätskennzahlen ist die Funktion zu deren Bestimmung in Abbildung 1 dargestellt.

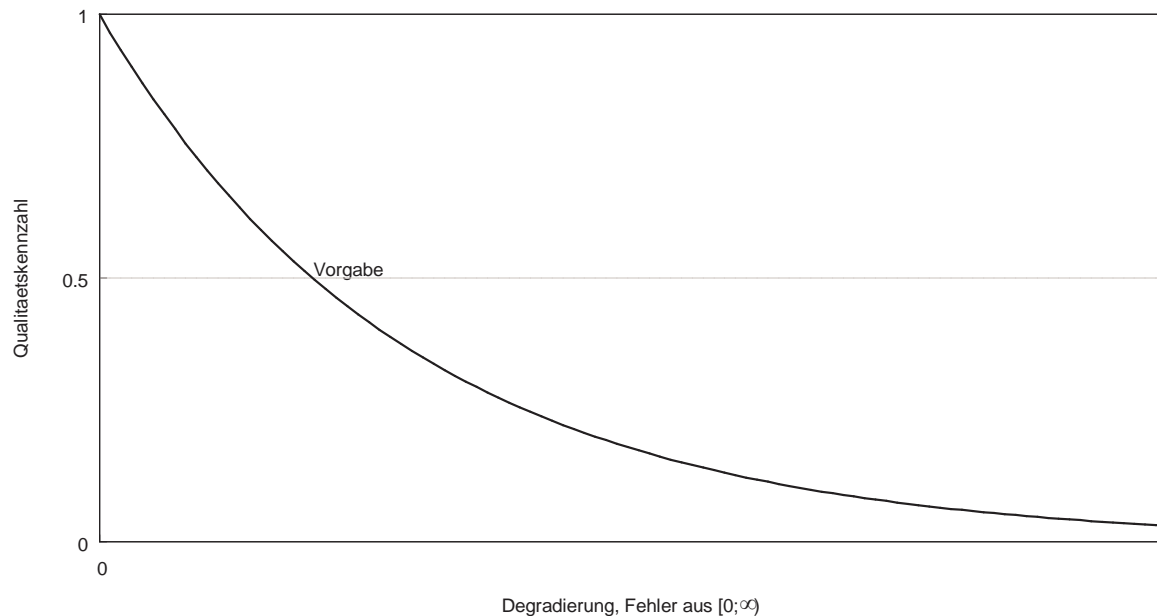


Abbildung 1: Stetig ableitbare Funktion nach Gleichung (2.1) zur Bestimmung einer Qualitätskennzahl, die für einen vorgegebenen Wert auf den Zielwert 0,5 normiert wird.

2.2. Metriken für abgeschlossen definierte Wertebereiche von Degradierung

Ist ein Fehler f oder eine Degradierung definiert auf einem abgeschlossenen Intervall $[f_{\min}; f_{\max}]$, so kann deren lineare Abbildung auf eine Qualitätskennzahl q_2 durch Gleichung (2.3) definiert werden.

$$(2.3) \quad q_2 = \begin{cases} 1 - \frac{f}{f_{\max} - f_{\min}} & , \quad \text{für } f < f_{\max} - f_{\min} \\ 0 & , \quad \text{für } f \geq f_{\max} - f_{\min} \end{cases}$$

Soll auch hier die Kennzahl zusätzlich für einen Vorgabewert $f=v$ auf den Kennzahl-Zielwert η abgebildet werden, so wären für die Abbildung auf eine Qualitätskennzahl Linearinterpolationen zwischen Minimum und Vorgabe und auch zwischen Vorgabe und Maximum als denkbar, wie in Abbildung 2 im Vergleich zu einer stetig ableitbaren Interpolation dargestellt. Der Nachteil hierbei ist aber die fehlende stetige Ableitbarkeit am Punkt, der der Qualitätsvorgabe entspricht, was bedeutet, dass gleiche Änderungen der Kennzahl (direkt) vor und nach dem vorgegebenen Wert von unterschiedlichen Änderungen der Degradierung herrühren und damit eine Überwachung und Prognose der Entwicklung dieser Qualität über die Zeit komplexer wird.

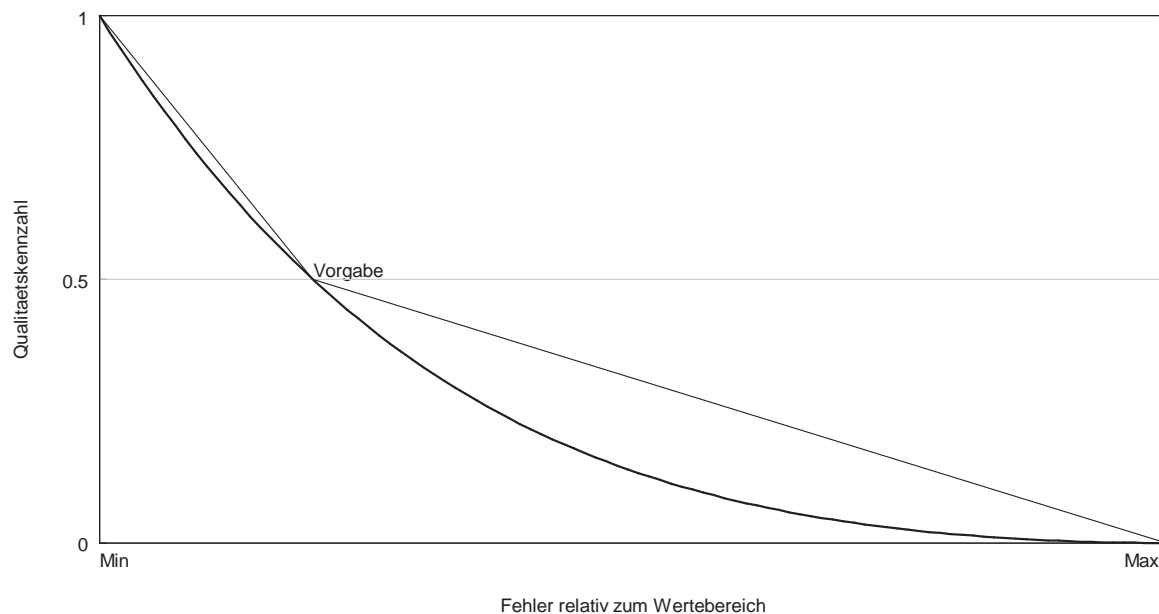


Abbildung 2: interpolierte Funktion zur Bestimmung einer Qualitätskennzahl in Abhängigkeit einer Degradierung, welche auf einem abgeschlossenen Intervall $[Min; Max]$ definiert ist. Vergleich von Linearinterpolation und stetig ableitbarer Interpolation.

Für die stetig ableitbare Abbildung eines Fehlers auf die Qualitätskennzahl kann die stetig ableitbare Thiele Interpolation angewendet werden [Thiele 1909], [Abramowitz et. al. 1964]. Der Thiele Kettenbruch wird beschrieben durch Gleichung (2.4) nach [Hildebrand 1974], [Lether 1999].

$$(2.4) \quad f(x) = a_0 + \sum_{i=1}^n \frac{x - x_{i-1}}{|a_i|}$$

Die Abbildung auf eine Qualitätskennzahl q_2 im Wertebereich $[m_1; m_2]$ unter der Bedingung, dass der vorgegebene Qualitätswert $v \in (m_1; m_2)$ auf einen Zielwert η abgebildet wird: $q_3(v) = \eta$, wird durch die drei Punkte (Min; 1), $(v; \eta)$ und (Max; 0) definiert und ist entsprechend durch Gleichung (2.5) gegeben. Abbildung 3 zeigt beispielhaft die Qualitätskennzahlenfunktionen für $m_1 = \text{Min} = 0$, $m_2 = \text{Max} = 100$ und $\eta = 0,5$ mit verschiedenen Degradierungs-Vorgabewerten $v \in \{5; 25; 50; 75; 99; 99,999999\}$.

$$(2.5) \quad q_3(x) = 1 + \frac{x - m_1}{\frac{m_1 - v}{1 - \eta} + \frac{x - v}{\frac{m_1 - m_2}{\frac{m_1 - v}{1 - \eta} - \frac{v - m_2}{\eta}} + \eta - 1}}$$

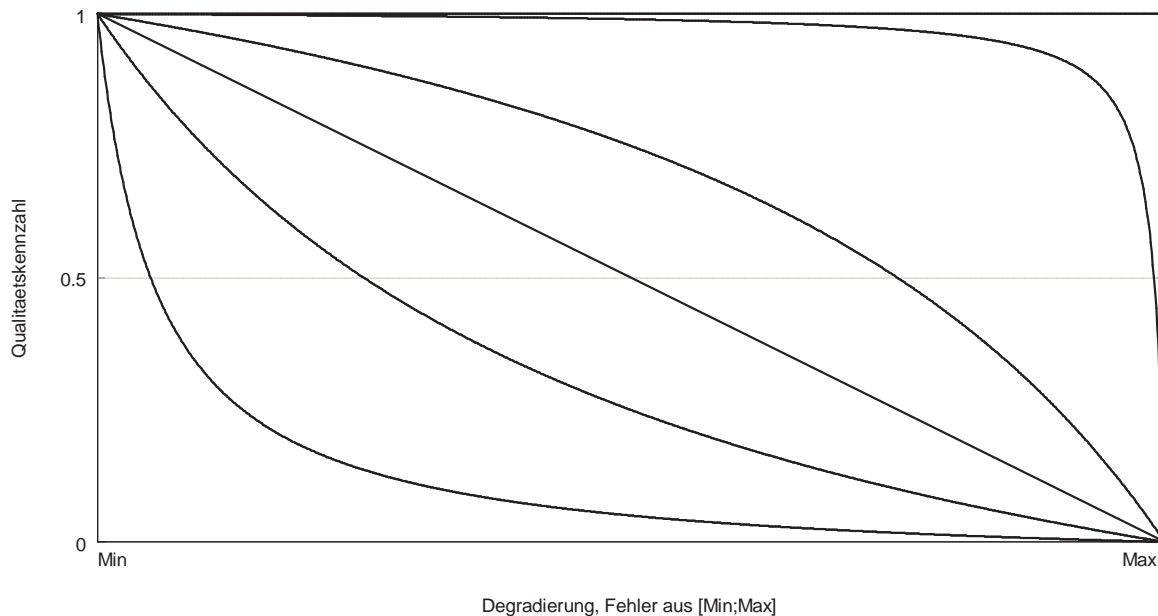


Abbildung 3: Qualitätskennzahlenfunktionen zum Fehler aus [Min; Max] für verschiedene Vorgabewerte

2.3. Metriken für Degradierung bezüglich abgeschlossen definierten Gültigkeitsbereichen

Für eine Qualitätskennzahl q bezüglich eines Gültigkeitsbereiches kann binär angegeben werden, ob der Gültigkeitsbereich eingehalten wird, womit gilt $q=1$, oder nicht, also $q=0$. Um aber eine Entwicklung dieser Qualität beobachten und prognostizieren zu können, kann eine Bestimmung der Qualitätskennzahl q bezogen auf einen Bereich angegeben werden, wie es in Abbildung 4 dargestellt ist am Beispiel des Luftfeuchtigkeit Sensors HIH4000 von Honeywell, der laut Datenblatt des Herstellers einen Arbeitsbereich von -45°C bis 80°C besitzt.

Für die Kennzahl werden die Grenzen des Arbeitsbereiches, welcher durch das Minimum m_1 und das Maximum m_2 definiert ist, jeweils auf die Kennzahl 0,5 abgebildet entsprechend der Semantik, dass ab diesem Wert vorgegebene Qualitäten eingehalten werden. Die Mitte des Arbeitsbereiches wird auf das Maximum 1 der Kennzahl abgebildet. Außerhalb des Arbeitsbereiches soll die Kennzahl sich asymptotisch der 0 nähern. Dies ist in der Praxis oft irrelevant, da auch die Sensorik zur Bestimmung des Wertes einer Umgebungsbedingung nur einen begrenzten Wertebereich besitzt, theoretisch jedoch lassen sich auf diese Weise noch beliebig schlechte Werte außerhalb des Arbeitsbereichs untereinander Vergleichen.

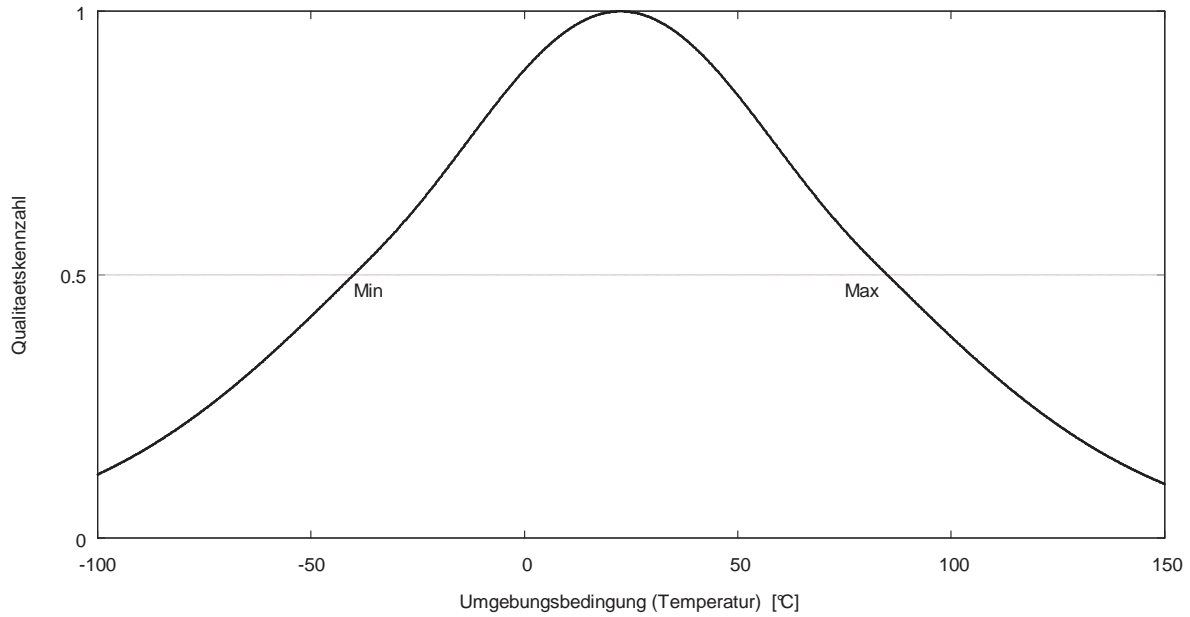


Abbildung 4: Qualitätskennzahl bezüglich eines Gültigkeitsbereichs, Beispiel-Arbeitsbereich [-45°C; 80°C]

Die Beschreibung einer Umgebungsbedingung aus dem Wertebereich $(-\infty; \infty)$ wird auf eine Qualitätskennzahl mit Wertebereich $(-\infty; \infty)$ abgebildet, wobei für den Intervall des Arbeitsbereichs $[m_1; m_2]$ eine Abbildung auf den Wertebereich $[0,5; 1]$ erwünscht ist. Die Funktion zur Bestimmung der Qualitätskennzahl wurde zweifach stetig ableitbar zusammengesetzt aus einer Exponentialfunktion außerhalb des Arbeitsbereichs und einem Polynom innerhalb des Arbeitsbereichs. Es wird ein Polynom 6. Grades benötigt, um die gewünschte zweifach stetige Ableitbarkeit zu erfüllen. Aus den Vorgaben ergeben sich die Koeffizienten a_0 bis a_6 des Polynoms wie in Gleichung (2.6) in Abhängigkeit der Grenzen des Arbeitsbereichs m_1 und m_2 angegeben.

$$\begin{aligned}
 (2.6) \quad a_0 &= \frac{m_1 - 8m_1^5 m_2 + 25m_1^4 m_2^2 - 60m_1^3 m_2^3 + 25m_1^2 m_2^4 - 8m_1 m_2^5 + m_2^6}{2(15m_1^4 m_2^2 - 20m_1^3 m_2^3 - 6m_1^5 m_2 + 15m_1^2 m_2^4 - 6m_1 m_2^5 + m_1^6 + m_2^6)} \\
 a_1 &= \frac{m_1^5 - 5m_1^4 m_2 + 40m_1^3 m_2^2 + 40m_1^2 m_2^3 - 5m_1 m_2^4 + m_2^5}{15m_1^4 m_2^2 - 20m_1^3 m_2^3 - 6m_1^5 m_2 + 15m_1^2 m_2^4 - 6m_1 m_2^5 + m_1^6 + m_2^6} \\
 a_2 &= \frac{30m_1^3 m_2 + 120m_1^2 m_2^2 + 20m_1 m_2^3}{15m_1^4 m_2^2 - 20m_1^3 m_2^3 - 6m_1^5 m_2 + 15m_1^2 m_2^4 - 6m_1 m_2^5 + m_1^6 + m_2^6} \\
 a_3 &= \frac{10m_1^3 + 110m_1^2 m_2 + 110m_1 m_2^2 + 10m_2^3}{15m_1^4 m_2^2 - 20m_1^3 m_2^3 - 6m_1^5 m_2 + 15m_1^2 m_2^4 - 6m_1 m_2^5 + m_1^6 + m_2^6} \\
 a_5 &= \frac{36m_1 + 36m_2}{15m_1^4 m_2^2 - 20m_1^3 m_2^3 - 6m_1^5 m_2 + 15m_1^2 m_2^4 - 6m_1 m_2^5 + m_1^6 + m_2^6} \\
 a_6 &= \frac{12}{15m_1^4 m_2^2 - 20m_1^3 m_2^3 - 6m_1^5 m_2 + 15m_1^2 m_2^4 - 6m_1 m_2^5 + m_1^6 + m_2^6}
 \end{aligned}$$

Aus den gewünschten Eigenschaften ergibt sich Gleichung (2.7) zur Bestimmung der Qualitätskennzahl q_4 , wobei die Koeffizienten a_0 bis a_6 vereinfacht zusammengefasst sind.

$$(2.7) \quad q_4(m_1, m_2, x) = \begin{cases} \begin{aligned} & - (10m_1^4 m_2 x - 220m_1 m_2^2 x^3 + 60m_1^3 m_2 x^2 + 240m_1^2 m_2^2 x^2 + 220m_1 m_2 x^4 \\ & + 60m_1 m_2^3 x^2 - 80m_1^3 m_2^2 x + 10m_1 m_2^4 x - 80m_1^2 m_2^3 x - 220m_1^2 m_2 x^3 + 24x^2 \\ & - m_1^6 - m_2^6 + 8m_1^5 m_2 - 25m_1^4 m_2^2 + 60m_1^3 m_2^3 - 25m_1^2 m_2^4 + 8m_1 m_2^5 - 2m_1^5 x \\ & - 2m_2^5 x - 20m_1^3 x^3 - 20m_2^3 x^3 + 70m_1^2 x^4 + 70m_2^2 x^4 - 72m_1 x^5 - 72m_2 x^5) / \\ & 2(15m_1^4 m_2^2 - 20m_1^3 m_2^3 - 6m_1^5 m_2 + 15m_1^2 m_2^4 - 6m_1 m_2^5 + m_1^6 + m_2^6), \\ & \text{für } m_1 < x < m_2 \end{aligned} \\ \frac{1}{2} e^{\left(\frac{2(x-m_1)(x-m_2)}{(m_2-m_1)^2} \right)}, \quad \text{sonst} \end{cases}$$

3. Qualitätskennzahlmetriken für einzelne Domänen der Degradierung

Im Folgenden werden die Domänen einzelner Qualitätsangaben beschrieben, Metriken zur sensorbezogenen Qualitätsbestimmung angegeben und die Semantik der Intervallgrenzen beschrieben.

3.1. Domäne Empirie

Es werden in dieser Domäne die stochastischen oder zufälligen Fehler [Hoffmann 2007], [Käppeler 2008] zusammengefasst, die eine Messung beeinflussen. Dazu können die in Kapitel 1.3 beschriebenen, absoluten, relativen oder reduzierten Fehler oder die Standardabweichung verwendet werden, um ein Qualitätsmaß anzugeben. In Bezug zu physikalischen Eigenschaften des Sensors selbst und auf das einheitliche Intervall des Qualitätsmaßes $[0; 1]$ beschränkt, werden die verschiedenen Fehler des Sensors aus dem Wertebereich $[0; \infty)$ auf eine Qualitätszahl normiert.

Dabei können der absolute oder relative Fehler per Exponentialfunktion aus Gleichung (1.1) abgebildet werden. Ein reduzierter Fehler aus einem abgeschlossenen Intervall kann durch Gleichung (1.3) abgebildet werden. Der Vorgabewert, der dabei auf den Zielwert 0,5 abgebildet wird, kann den Herstellerangaben zum maximal auftretenden absoluten oder relativen Fehler entnommen werden. Der tatsächlich auftretende absolute, relative oder reduzierte Fehler lässt sich jedoch schwer schätzen. Dadurch ist eine Qualitätszahl bezogen auf den Sensor selbst wenig sinnvoll. Lediglich bei vorhandener Vorgabe der Qualitätsanforderung einer Anwendung bezüglich eines absoluten, relativen oder reduzierten Fehlers, welche auf den Zielwert abgebildet wird, kann eine davon abhängige Qualitätskennzahl für den jeweils verwendeten Sensor und dessen vom Hersteller angegebenen Fehler berechnet werden. Dies ist für die Anwendung dann von Bedeutung, sofern nicht immer derselbe Sensor zur Beobachtung identischer Phänomene verwendet wird, was bei einigen kontextbezogenen Anwendungen, die auf mobilen Geräten eingesetzt werden, durchaus denkbar ist.

Die Angabe einer Standardabweichung des Sensors lässt sich auch ohne Vorgaben einer Anwendung auf eine Qualitätskennzahl sinnvoll abbilden, da die aktuelle Standardabweichung empirisch durch Vergleichsmessungen ermittelt werden kann und dann in Relation zur vom Hersteller angegebenen Standardabweichung auf eine Kennzahl

abgebildet werden kann. Damit ist der Zustand eines Sensors überwachbar, wie in [Benkmann et. al. 2008] gezeigt wird.

Die Bestimmung einer Qualitätskennzahl, die basierend auf der geschätzten Standardabweichung ermittelt wird, kann nach [Käppeler et. al. 08] durchgeführt werden. Dabei werden Standardabweichungen mittels Abgleich der Messungen verschiedener Sensoren geschätzt und die Qualitätskennzahl nach Gleichung (2.1) auf den Intervall $[0;1]$ normiert, wobei der Normierungsfaktor γ dafür sorgt, dass die normierte Kennzahl für eine vom Hersteller vorgegebene Standardabweichung den Zielwert η erhält, der in unserem Fall mit 0,5 vorgegeben wird.

3.2. Domäne Querempfindlichkeit

Als Querempfindlichkeit werden physikalische Einflüsse auf einen Sensor, also seine Umgebungsbedingungen, und deren Einfluss auf die Messgenauigkeit beschrieben. Beispiele für einzelne Umgebungsbedingungen sind unter anderem Temperatur, Druck, Luftfeuchte, oder Helligkeit. Für einzelne Umgebungsbedingungen können Arbeitsbereiche eines Sensors vorgegeben werden, für die jeweils ein maximaler relativer oder absoluter Fehler bekannt ist. Wird vom Hersteller eines Sensors wie üblich nur ein einziger Arbeitsbereich in Bezug auf einzelne Umgebungsbedingung vorgegeben (ohne resultierende Fehler anzugeben), so kann für die Qualitätskennzahl q_{quer} binär angegeben werden, ob der Arbeitsbereich eingehalten werden kann, womit gilt $q_{quer}=1$, oder nicht, also $q_{quer}=0$. Um jedoch eine Entwicklung der Qualität bezüglich Querempfindlichkeit beobachten und prognostizieren zu können, kann eine Bestimmung der Qualitätskennzahl q_{quer} bezogen auf einen Bereich angegeben werden, wie es in 2.3 beschrieben ist.

3.3. Domäne Signalverarbeitung

Diese Domäne beschreibt verschiedene Einflüsse der Signalverarbeitung auf die Qualität des Messwertes. Mittels Sensorik werden Beobachtungen zu elektrischen Signalen umgewandelt. Dadurch entsteht grundsätzlich ein Informationsverlust der mittels eines relativen oder absoluten Fehlers angegeben wird. Es handelt sich bei diesen Angaben um einen Anteil am gesamten Messfehler.

Entsprechende Qualitätskennzahlen geben entweder konstante Anteile des Messfehlers der Domäne Empirie an oder den Einfluss von Umgebungsbedingungen auf einzelne Signalverarbeitungsschritte analog zu den Angaben aus der Domäne Querempfindlichkeit. Die Metriken zur Qualitätskennzahl sind entsprechend zu übernehmen.

3.4. Domäne Zuverlässigkeit

Die Zuverlässigkeit eines Sensors kann zum Beispiel empirisch ermittelt werden, indem nach [Käppeler et. al. 2008] Messdaten verschiedener Sensoren abgeglichen und dadurch bewertet werden oder nach [Benkmann et. al. 2008] ein Fuzzy Clustering auf die Messdaten angewendet wird um Ausreiser zu erkennen. Entsprechend kann eine Zuverlässigkeit beziehungsweise Ausfallrate eines Sensors oder messenden Systems ermittelt werden, die per Exponentialfunktion auf eine Qualitätskennzahl abgebildet und wie in Abbildung 1 dargestellt normiert und auf eine Herstellerangabe oder auf eine Garantie eines Betreibers bezogen werden kann.

Aus nötigen oder durchgeführten Wartungsarbeiten lässt sich ebenfalls eine Qualitätskennzahl ableiten. Dabei wird gemäß Gleichung (2.1) die seit der letzten Wartung verstrichene Zeit x in Relation zur Herstellerangabe v oder zum Wartungsplan angegeben, wobei wiederum die Vorgegebene Zeitspanne auf die Kennzahl $\eta=0,5$ festgesetzt wird. Analog dazu lässt sich auch

eine Abnutzung von Sensoren als Qualitätskennzahl beschreiben. Werden in Sensoren gewisse Betriebsmittel benötigt wie Indikator-Chemikalien [Römpp et. al. 99], Energie, Schmierfett, Kühlmittel oder mechanische Teile, die sich bei der Durchführung von Messungen abnutzen [Lucke et. al. 2009], so wird aus der bisherigen Betriebsdauer x in Relation zur Lebensdauer des Bauteils v eine Qualitätskennzahl definiert. Zu beachten ist, dass die Kennzahl bezüglich einer „Vorgabe“ einer Qualität bei Verbrauchsmaterial oder Energievorräten nicht semantisch sinnvoll für den Wert 0,5 anwendbar ist. Trotzdem muss der anteilige Vorrat vom Intervall $[1; 0]$ auf den Wertebereich $[1; 0,5]$ normiert werden, damit die entsprechende Kennzahl sinnvoll mit weiteren Qualitätskennzahlen verknüpft werden kann deren Zielwert ebenfalls mit $\eta=0,5$ definiert ist.

3.5. Domäne Temporale Aspekte

Zu dieser Domäne werden unterschiedliche temporale Aspekte zusammengefasst, die sich auf die Qualität von Messdaten auswirken.

Die Hysterese beschreibt den Einfluss bisheriger Messungen auf einen aktuellen Messwert. Dies ist oft bei gefilterten Messwerten eine Ursache für Messabweichungen. Zum Beispiel kann die Hysterese von den ermittelten Varianzen eines eingeschwungenen Kalmanfilters abhängen und in Abhängigkeit dieser Varianzen beschrieben werden.

Eine Einschwingzeit beschreibt die Dauer, die zwischen einer Änderung der Eingangsgröße eines Messgerätes und der daraus resultierenden Änderung des Messwertes (innerhalb vorgegebener Grenzen) liegt. Als Beispiel sei

Die Präzision, auch Wiederholgenauigkeit genannt, beschreibt indirekt die Qualität eines Messverfahrens. Der Wert gibt die relative oder absolute Abweichung voneinander unabhängiger Messungen untereinander an. Dies wird ermittelt, indem zum Beispiel dieselbe Größe mit einem Messgerät mehrmals erfasst wird.

Die Abtastrate, Samplingrate oder Samplerate gibt ein minimales oder konstantes Zeitintervall an, das zwischen zwei Messungen eines Sensors liegt. Besonders bei der Beobachtung dynamischer Größen beeinflusst die Abtastrate das Ergebnis. Oft werden durch Interpolation oder rückgekoppelte Filterung Zustände zwischen den einzelnen Messungen geschätzt.

Eine Drift beschreibt eine Abweichung, die meist bei indirekten Messungen wie zum Beispiel einem Dead Reckoning zu berücksichtigen ist. Oft wird diese Drift durch eine Initialisierung des Messsystems reduziert. Die Drift wird angegeben in absolutem oder relativem Offset pro Messvorgang oder Zeiteinheit.

Für eine Auswertung von Sensordaten sind die temporalen Aspekte nur von Bedeutung, sofern dynamische Phänomene beobachtet werden. Die entsprechenden Auswirkungen sind jedoch stark Anwendungsabhängig und so wird an dieser Stelle darauf verzichtet, allgemein gültige Qualitätskennzahlen für Hysterese, Einschwingzeit, Präzision, Abtastrate oder Drift zu definieren. Für eine Hysterese, die durch eine Hysteresekurve angegeben wird, und für die weiteren temporalen Aspekte, die mit unterschiedlichen Einheiten angegeben werden können, ist es schwierig, eine allgemeine Beschreibung für eine Ungenauigkeit zu definieren. Meist werden Hysterese, Einschwingzeit, Präzision, Abtastrate oder Drift statisch vorgegeben und müssen nicht überwacht, aber als Ursache für einen absoluten oder relativen Fehler beachtet werden. Trotzdem soll ein SensorContextServer die Möglichkeit bieten, die Qualität eines Sensorwertes diesbezüglich zu überwachen, sofern er dazu Vorgaben einer Anwendung erhält, welche der internen Zustände von speziellen Schätzfiltern zum Beispiel in die Bewertung der aktuellen Datenqualität einfließen sollen.

Ein Temporaler Aspekt, der wiederum direkt auf eine Qualitätskennzahl abbildbar ist, ist die Aktualität. Hierfür wird die Differenz $x=t-t_s$ zwischen dem Zeitstempel t_s der Messung und

der aktuellen Zeitangabe t auf eine Qualitätszahl nach Gleichung (2.1) per Exponentialfunktion abgebildet, indem als Vorgabewert v , für den die Qualitätszahl 0,5 definiert wird, der Zeitintervall verwendet wird, der durch die maximale Abtastrate des Sensors gegeben ist. Auch von Anwendungen lassen sich durch einfache Angaben von einzuhaltenden Zeitintervallen eigene Anforderungen bezüglich dieses Qualitätsmerkmals vorgeben.

4. Überwachung der Datenqualität für einzelne Anwendungen

Ein SensorContextServer soll Qualitätsmetriken verwalten und die Überwachung der von der Anwendung geforderten Qualität übernehmen können. Damit sollen die kontextbezogenen Anwendungen, welche auch auf mobilen Geräten eingesetzt werden und dadurch nur über stark beschränkte Ressourcen an Energie und Rechenleistung verfügen können, entlastet werden. Eine entsprechende Implementierung und Festlegung der Syntax für allgemeine und durch die Anwendungen definierbare Berechnungen der Qualitätskennzahlen fehlen jedoch bisher. Die verfügbare Implementierung sieht nur eine Vorgabe für Qualitätskennzahlenmetriken durch den SensorContextServer selbst vor.

Das Konzept für den SensorContextServer sieht vor, dass Anwendungen ihre Anforderungen an die Qualität auf unterschiedlichen Arten vorgeben können.

Zum Einen soll es möglich sein, vollständige Metriken inklusive der Schlüsselworte aus [Kae08] und damit komplette Anweisungen für die Berechnung von Qualitätskennzahlen vorzugeben. Zusätzlich sollen auch die Verknüpfungen der Domänen zu einem Gesamtqualitätswert oder deren einzelnen Einflüsse zu je einem Wert angegeben werden können.

Eine weitere Möglichkeit sieht vor, dass Anwendungen nur Teile von Metriken wie Faktoren, Zeitvorgaben und Grenzen von Abweichungen (zum Beispiel durch Verzögerungen) vorgeben, die die hier angegebenen Metriken aus den Gleichungen (2.1) bis (2.6) vervollständigen.

Die dritte Möglichkeit sieht lediglich die Angabe von Gewichten für die einzelnen Qualitätsdomänen vor, um Schwerpunkte der Qualitätsanforderungen für einzelne Anwendungen festzulegen. Somit kann eine Anwendung für sich festlegen, ob zum Beispiel die Aktualität der verwendeten Kontextinformationen von Relevanz ist, ob Informationen zu Wartungen von Sensorik ausgewertet werden oder ob zum Beispiel Ausfälle von Sensoren über die Schätzung der Standardabweichung in den Qualitätswert einfließen und damit schon dadurch überwacht werden können, dass nur eine einzige Qualitätskennzahl ausgewertet wird.

Für die Zusammenfassung verschiedener Qualitätskennzahlen $q_1..q_n$ aus den einzelnen Domänen zu einer Gesamtqualitätskennzahl Q bezüglich der Degradierung können verschiedene Ansätze verfolgt werden. Sinnvolle Verknüpfungen sind die Auswahl des Minimums nach Gleichung (4.1). Mit der Überwachung des Minimums aller Qualitätskennzahlen kann eine Anwendung ganz einfach überprüfen, ob alle vorgegebenen Qualitätsanforderungen eingehalten werden. Die Anforderungen werden jeweils auf den Kennzahlwert 0,5 normiert. Für $Q_{\min} < 0,5$ gibt es mindestens eine Vorgabe an die Qualität der Kontextinformationen, die nicht eingehalten wird. Das arithmetische Mittel aller Qualitätskennzahlen dagegen ermöglicht die Überwachung der Entwicklung der Gesamtqualität von Kontextinformationen über einen zeitlichen Verlauf betrachtet oder das Arithmetische Mittel der Kennzahlen nach Gleichung (4.2). Mit der Überwachung des Minimums aller Qualitätskennzahlen kann eine Anwendung ganz einfach überprüfen, ob alle

vorgegebenen Qualitätsanforderungen eingehalten werden. Die Anforderungen werden jeweils auf den Kennzahlwert 0,5 normiert. Für $Q_{\min} < 0,5$ gibt es mindestens eine Vorgabe an die Qualität der Kontextinformationen, die nicht eingehalten wird. Das arithmetische Mittel aller Qualitätskennzahlen dagegen ermöglicht die Überwachung der Entwicklung der Gesamtqualität von Kontextinformationen über einen zeitlichen Verlauf betrachtet.

$$(4.1) \quad Q_{\min} = \min(q_1, \dots, q_n)$$

$$(4.2) \quad Q_{\text{avg}} = \sum_{i=1}^n \frac{q_i}{n}$$

Die Nexus Föderation muss in der Lage sein, die verschiedenen Metriken zu verwalten, neue Metriken von den Anwendungen aufnehmen können und in einem Katalog den Anwendungen zur Auswahl vorhalten. Ein Zugriff über die Semantik der einzelnen Metriken kann über die Schlüsselworte aus [Kae08] realisiert werden. Zusammengefasst werden folgende Typen von Metriken verwaltet:

Typ der Metrik	Beschreibung
Vollständige Anwendungsvorgaben	Metriken für einzelne Domänen und deren Verknüpfung werden von Anwendungen vollständig vorgegeben.
Templates	Generell einsetzbare Metriken werden von der Nexus Föderation bereitgestellt. Diese sind mit unterschiedlichen Parametrisierungen für verschiedene Anwendungsbereiche versehen, die von einzelnen Anwendungen vorgegeben werden können. Anwendungen können sich auch darauf beschränken, Gewichte für die Domänen zur Berechnung der Gesamtqualität vorzugeben.
Sensorbezogen	Einzelne Metriken zur Degradierung von Sensordaten beziehen sich lediglich auf physikalische Eigenschaften des Sensors selbst. Ein aktueller Qualitätswert kann nur in Relation zu Herstellerangaben betrachtet werden.

Tabelle 2: Zusammenfassung verschiedener Typen von Qualitätsmetriken

Referenzen

[Abramowitz et. al. 1964]

Abramowitz, Milton; Stegun, Irene: **Handbook of Mathematical Functions**, National Bureau of Standards Applied Mathematics Series No. 55, U.S. Govt. Printing Office, Washington, 1046 Seiten, 1964

[Benkmann et. al. 2008]

Benkmann, Ruben; Käppeler, Uwe-Philipp; Zweigle, Oliver; Lafrenz, Reinhard; Levi, Paul: **Resolving Inconsistencies using Multi-agent Sensor Systems**. In: Seabra Lopez, Luis (Hrsg); Silva, Filipe (Hrsg); Santos, Vitor (Hrsg): Proceedings of the 8th Conference on Autonomous Robot Systems and Competition: Robotica08; Aveiro, Portugal, 2. April, 2008

[Dietrich 1991]

Cornelius Frank Dietrich: **Uncertainty, Calibration and Probability, The Statistics of Scientific and Industrial Measurement**. The Adam Hilger Series on Measurement Science and Technology, Adam Hilger Imprint, ISBN 0-7503-0060-4, 2nd Edition, 1991

[DIN94]

Internationales Wörterbuch der Metrologie, Herausgeber: DIN Deutsches Institut für Normung e.V., Beuth GmbH Verlag, 2. Auflage, 1994

[DIN95]

Guide to the Expression of Uncertainty in Measurement. ISO/BIPM-Leitfaden (Internationales Büro für Maß und Gewicht). DIN 01319, 1995

[DIN99]

Leitfaden zur Angabe der Unsicherheit beim Messen, Maßgebliche deutsche Fassung des GUM „Guide to the Expression of Uncertainty in Measurement“ ist die Vornorm DIN V ENV 13005, aktuelle Ausgabe:1999-06, 1999

[Dudkowski et. al. 2006]

Dudkowski, Dominique; Käppeler, Uwe-Philipp; Nicklas, Daniela; Schwarz, Thomas; Siemoneit, Oliver; Volz, Steffen; Wieglering, Klaus; Zweigle, Oliver: **Konsistenz in Nexus**, Technischer Bericht, Sonderforschungsbereich Nexus, Universität Stuttgart, 2006

[Hildebrand 1974]

Hildebrand, Francis: **Introduction to Numerical Analysis**, McGraw-Hill, New York, 1974

[Hoffmann 2007]

Jörg Hoffmann: **Handbuch der Messtechnik**, Carl Hanser Verlag München, 3. Auflage, ISBN 978-3446407503, 2007

[Hönle et. al. 2005]

Hönle, Nicola; Käppeler, Uwe-Philipp; Nicklas, Daniela; Schwarz, Thomas: **Benefits Of Integrating Meta Data Into A Context Model**; In: Proceedings of 2nd IEEE PerCom Workshop on Context Modeling and Reasoning (CoMoRea) at 3rd IEEE International Conference on Pervasive Computing and Communication (PerCom'05); Hawaii, March 12, 2005.

[IEC2371]

International Electrotechnical Commission, IEC/ISO 2371, ANSI S2.38: **Description and Evaluation of Field Balancing Equipment**, 6 Seiten, B S I Standards, ISBN 978-0580078408, 1973

[Käppeler et. al. 2005]

Uwe-Philipp Käppeler, Dominique Dudkowski, Georg Kindermann, Darko Klinec: **Modellierung von Sensoren und Sensordaten in der Nexus-Plattform**; Technischer Bericht, Sonderforschungsbereich Nexus, Universität Stuttgart, 2005.

[Käppeler et. al. 2005b]

Uwe-Philipp Käppeler, Tobias Drosdol, Thomas Schwarz, Michael Scharf: **SensorContextServer und SensorClient in der Nexus-Plattform**; Technischer Bericht, Sonderforschungsbereich Nexus, Universität Stuttgart, 2005.

[Käppeler 2008]

Uwe-Philipp Käppeler: **Modellierung der Degradierung von skalaren Sensordaten in ContextServer und SensorContextServer**. Technischer Bericht, Sonderforschungsbereich Nexus, Universität Stuttgart, 2008

[Käppeler et. al. 2008]

Käppeler, Uwe-Philipp; Benkmann, Ruben; Zweigle, Oliver; Lafrenz, Reinhard; Levi, Paul: **Resolving Inconsistencies in Shared Context Models using Multiagent Systems**. In: Dillmann, Rüdiger (Hrsg); Burgard, Wolfram (Hrsg): Proceedings of the 10th International Conference on Intelligent Autonomous Systems: IAS-10; Baden Baden, Germany, July 23-25, 2008

[Lether 1999]

F.G. Lether: **Thiele Rational Interpolation for the Numerical Computation of the Reversible Randles-Sevcik Function in Electrochemistry**, Journal of Scientific Computing 14, Seiten 259-274, Springer Verlag, ISSN 0885-7474 (Print) 1573-7691 (Online), 1999

[Lucke et. al. 2009]

Dominik Lucke, Carmen Constantinescu, Engelbert Westkämper: **Context Data Model, the Backbone of a Smart Factory**. In: Tichkiewitch, Serge (Herausgeber); CIRP: Sustainable Development of Manufacturing Systems : 42nd CIRP Conference on Manufacturing Systems, Grenoble, June 3-5, 2009

[Nicklas et. al. 2001]

Daniela Nicklas, Matthias Grossmann, Thomas Schwarz, Steffen Volz, Bernhard Mitschang: **A model-based, open architecture for mobile, spatially aware applications**; in Proc. of Symp. on Spatial and Temporal Databases, 2001.

[Römpp et. al. 1999]

Hermann Römpp, Jürgen Falbe, Manfred Regitz: **Römpp Basislexikon Chemie**, Kompaktausgabe, 4 Bände, Thieme Verlag Stuttgart, deutsch, ISBN 978 3131157119, 1999

[Thiele 1909]

Thiele, Thorwald: **Interpolationsrechnung**, Teubner Verlag Leipzig, 175 Seiten, 1909