Multi-sensory Machine Diagnosis on Security Printing Machines with Two-Layer Conflict Solving

Voth, Karl[1]; Glock, Stefan[2]; Mönks, Uwe[1]; Türke, Thomas[3]; Lohweg, Volker[1]
[1]inIT-Institute Industrial IT, Ostwestfalen-Lippe University of Applied Science, Liebigstr. 87, 32657 Lemgo
[2]Friedrich-Alexander University of Erlangen-Nuremberg, Institute for Electronics Engineering, 91058 Erlangen
[3]KBA-NotaSys SA, Avenue du Grey 55, CH-1018 Lausanne

Abstract

The *Two Layer Conflict Solving* approach is based on the *Evidence Theory* and uses conflict solving to fuse data. We introduce an extension of the TLCS approach with reference to highly complex machine conditioning applications. In this context, the sensors are grouped to attributes applying expert knowledge. The fusion of the fuzzyfied sensors' observations that are elements of one particular attribute is accomplished by the TLCS. Subsequently, the attributes' conditions are merged using an *Ordered Weighted Averaging Operator*. The introduced approach is evaluated on a wiping simulator.

1. Introduction

Sensor fusion has found more and more applications recently, ranging from fault diagnosis and machine conditioning [1], [2] to military defence [3]. Data fusion deals with data that is received from sensors, experts or human linguistic words, etc. Furthermore, a lot of such knowledge is cognitive and imprecise (incomplete) to some degree. To deal with uncertain knowledge, various research groups often use *Dempster-Shafer Theory* (DST) [4], [5] because it is capable of managing uncertainty due to its framework. DST acts as the pioneer in data fusion algorithms, which was proposed by Dempster and extended by Shafer subsequently. However, data that is received from sensors or from cognitive knowledge can lead to counter-intuitive results, if one of the sensors or a human returns a wrong measurement or cognitive unreliable knowledge. It should be noted that sensors can also be mentioned as experts and vice versa.

The inherent defect pointed out by Zadeh [6] brings criticism as well as many other invented alternatives. In this paper, a *Two Layer Conflict Solving* (TLCS) data fusion approach, which was described in [8], is extended in a way that it can be applied to highly complex systems by dividing sensors into several groups and combining them afterwards, using an *Ordered Weighting Averaging* (OWA) operator.

In the first section we introduce the DST and the TLCS approach. The second section deals with the *Attribute Fusion Layer* as an extension to the TLCS approach. Finally, we evaluate the introduced approaches in a wiping simulator, and conclude with a summary and an outlook.

1.1 Dempster-Shafer Theory

Serving as a seminal fusion approach, the *Dempster-Shafer Theory* (DST) stirs up many discussions and studies in data fusion. DST is a mathematical theory of evidence, which combines independent sources of information [4], [5]. By combination of evidence sources obtained from sensors (experts), more reliable and convincing fusion results are expected. First, a finite frame of discernment, that forms a set Ω, is defined, $\Omega = \{\Psi_1, \Psi_2, \Psi_3, \ldots, \Psi_n\}$. A power set $\Theta = 2^{\Omega}$ includes all the possible combinations of propositions Ψ. Propositions are regarded to be mutually exclusive and exhaustive. A function *m*: $2^{\Omega} \rightarrow [0,1]$ is called a mass function, also known as *Basic Probability (Belief) Assignment* (*BPA, BBA*) with

$$m(\varnothing) = 0, \sum_{A \subseteq \Theta} m(A) = 1. \tag{1}$$

If there is no element in the BBA, then the mass is zero. On the other hand, as Θ is a power set composed of all the subsets of Ω, the sum of all the masses must be equal to one. Furthermore, the focal element (mass is larger than zero) is defined as:

$$\{(A, m(A)) \mid A \subseteq \Theta,\ m(A) > 0\}. \tag{2}$$

Belief (*Bel*) and plausibility function (*Pl*) are essential concepts in DST which are used in decision-making:

$$Bel(A) = \sum_{B \subseteq A} m(B) \text{ and } Pl(A) = \sum_{B \cap A \neq \varnothing} m(B). \tag{3}$$

Bel is called *lower bound probability*, while *Pl* is the *upper bound probability*, for the reason that *Bel* is the "must-be"-probability and on the other hand *Pl* is the "might-be"-probability. Therefore, *Pl* includes more mass than *Bel*, which is illustrated in Eq. 4:

$$Bel(A) \leq Pl(A). \tag{4}$$

After obtaining the above-mentioned concepts, we are able to use DST to fuse independent data sources by applying

$$\oplus_{i=1}^{n} m_i(A) = K \sum_{A_1 \cap .. \cap A_n = A} \prod_{i=1}^{n} m_i(A). \tag{5}$$

The term $\sum_{A_1 \cap .. \cap A_n = A} \prod_{i=1}^{n} m_i(A)$ aggregates the consonant opinions (non-conflicting parts) from sensors which is subsequently multiplied with *conflicting factor* $K = (1-k_c)^{-1}$, where

$$k_c = \sum_{A_1 \cap .. \cap A_n = \varnothing} \prod_{i=k}^{n} m_i(A_k). \tag{6}$$

The variable $m_i(A_k)$ denotes the mass of propositions from sensor i. According to the definition of k_c, it defines the empty intersection of the propositions of all sensors. Therefore, it is also called *conflicting coefficient*.

1.2 Two-Layer Conflict Solving

Because of the counter-intuitive results of DST [6], [7] and other alternatives have limited assistance as remedies, a *Two-Layer Conflict Solving* data fusion approach is suggested, which includes two layers to combine pieces of evidence. The conflict is solved to some degree during combination – hence the name conflict solving. The first layer resolves the conflict to some extent, and the second continues to solve it. By this, it achieves more stable results. The TLCS approach is depicted in Fig. 1.

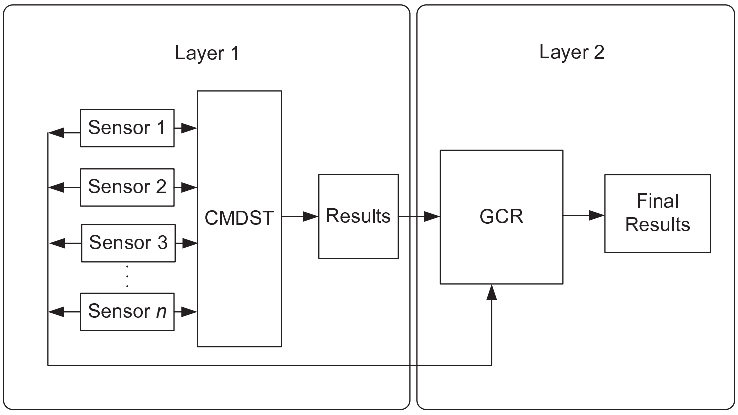

Figure 1: Two-Layer Conflict Solving System [8].

Psychologically, as clearly stated in [10], '*Decision making has been traditionally studied at three levels: individual, group and organizational.*' (cf. [11]). This shows that a decision is made on three layers, in which conflict is unavoidable to be considered and solved: The individual level is the basic element that holds conflict; group level has a larger range, which includes conflict, while organizational level is the largest. In such a way, people believe that conflict can be solved optimally, although it is impossible to totally eliminate its negative impacts. Therefore, a TLCS data fusion algorithm is suggested, studied and extended in this work. The approach is also applicable, when several groups of sensors are considered in a larger system. Layer 1 is regarded as working at the individual level because *Conflict-Modified-DST* (CMDST) [8] is an approach that combines every two sensors' data so that a conflict is considered and solved between two individuals. After receiving the results from the previous layer, layer 2 collects all sensors' original information and fuses it with the combined results from CMDST. Here the conflict is further resolved at a group level.

The *Balanced TLCS* approach (*BalTLCS*) is based on the normalised version of CMDST and is based on the TLCS. It was introduced in [9]. The *normalised conflicting factor* ${}^{N}K_{cm}$ is defined as follows:

$${}^{N}K_{cm} = \frac{B_c(n)}{B_c(n) - k_{cm} + \varepsilon} < \infty, \tag{7}$$

where $\forall \varepsilon > 0, \forall n > 1, n \in \mathbb{N}, k_{cm} = 0 \exists L \in \mathbb{R} : L = (1 - {}^{N}K_{cm}) < \varepsilon$. It is calculated by the binomial coefficient $B_c(n) = \binom{n}{2}$, the number of possible combinations of *n* sensors, and the modified conflicting factor

$$k_{\mathrm{cm}} = \sum_{\substack{A_1 \cap A_2 = \varnothing, A_1 \cap A_3 = \varnothing, \cdots, \\ A_i \cap A_k = \varnothing, \cdots, A_{n-1} \cap A_n = \varnothing}} \prod_{i=k}^{n} m_i(A_k), \tag{8}$$

calculated between every two sensors instead of all sensors together.

Therefore, as $0 \le k_{cm} \le B_c(n)$, the normalized conflicting factor ranges between

$$\frac{B_c(n)}{B_c(n)+\varepsilon} \le {}^{N}K_{cm} \le B_c(n)\cdot\varepsilon^{-1} < \infty \,. \tag{9}$$

The non-conflicting part is determined and coupled with the conflicting part as

$$^{\mathrm{N}}\,\mathrm{CMDST}(A) = \oplus_{i=1}^{n} m_i(A) = \frac{^{N}K_{cm}}{B_c(n)} \sum_{\substack{A_1\cap A_2=A,\ldots\\ A_1\cap A_3=A,\ldots\\ A_{n-1}\cap A_n=A}} \prod_{i=1}^{n} m_i(A)\,. \tag{10}$$

The *Balanced Group Conflict Redistribution* (*BalGCR*) approach combines sensors' propositions in a group manner that implies that in this case all sensors will participate additively in this procedure.
The intention is to utilize the inverse of the normalized conflict factor as a control parameter. If no conflict has occurred, mainly the $^{\mathrm{N}}$CMDST fusion result should contribute to the overall result. If the conflict is high, then all different information sources have to be taken into account. None of the information sources is allowed to dominate another, and none of them is allowed to play a part in the overall result with more than $1/n$.
Furthermore, in a strong conflict case it is not intended to shift the information content (hypotheses) to the universal set Θ, defining ambiguity or ignorance, because a conflict has to be solved in any case. As the set must be complete regarding the sources, all information sources (hypotheses) which can appear must be covered in a set. The only situation where hypotheses can be transferred to the universal set Θ occurs when a source delivers no reliable data.
Formally, *BalGCR* fusion consists of two parts. The first one describes the non-conflicting part:

$$m_{nc}(A) = {}^{N}K_{cm}^{-1}\cdot {}^{\mathrm{N}}\,\mathrm{CMDST}(A) = \frac{1}{B_c(n)} \sum_{\substack{A_1\cap A_2=A,\ldots\\ A_1\cap A_3=A,\ldots\\ A_{n-1}\cap A_n=A}} \prod_{i=1}^{n} m_i(A)\,. \tag{11}$$

The second part characterizes the conflicting part. Eq. 11 tends to zero in heavy conflicts $\left(k_{cm} \to B_c(n)\right)$. Therefore, it is proposed here that all masses supporting a certain hypothesis have to be averaged to fulfil the above mentioned statements. Furthermore, the average value has to be controlled, e. g., by the normalised conflicting factor or coefficient. It is proposed to use Eq. 12 for the conflicting part:

$$m_c(A) = \frac{k_{cm}}{B_c(n)}\cdot\frac{1}{n}\sum_{i=1}^{n} m_i(A)\,. \tag{12}$$

It can be recognized in Eq. 12 that, in the case of maximum conflict, the average value of all sensory hypotheses is determined. In the case of minimum conflict, $m_c(A)$ tends to zero. Both parts (Eq. 11 and Eq. 12) are additively connected:

$$m(A) = m_c(A) + m_{nc}(A) = \frac{k_{cm}}{B_c(n)}\cdot\frac{1}{n}\sum_{i=1}^{n} m_i(A) + \frac{1}{B_c(n)} \sum_{\substack{A_1\cap A_2=A,\\ A_1\cap A_3=A,\\ A_2\cap A_3=A,\ldots\\ A_{n-1}\cap A_n=A}} \prod_{i=1}^{n} m_i(A)\,. \tag{13}$$

2. Attribute Fusion Layer

In this section we introduce an extension of the TLCS fusion approach. Experts in the DST approach and its derivates require independent information sources observing the same attribute. It can be the weather, a certain part in a machine, or the holistic machine condition.
In order to improve the estimation of the machine's condition we assume, that information sources are specialized in observing specified machine parts or attributes. Our aim is to define machines' attributes, define sources that observe these attributes, estimate the attributes' condition using the TLCS approach, and finally merge the attributes' conditions using an attribute fusion layer.

2.1 Aggregation Operators

Various relations between different criteria have to be considered in many applications in engineering and science. Relations are intimately connected with classification, pattern recognition, and control. In general, relations represent the mapping of sets. In the case of crisp relations, there are only two degrees of relationship between the elements of the sets: "completely related" (i.e., being an element of the set); and "not related" (i.e., being not an element of the set). Contrary to crisp relations, fuzzy relations have an infinite number of relationships between the boundaries "completely related" and "not related". These relationships are represented by so-called fuzzy membership functions. Furthermore, different fuzzy subsets can be combined into one overall set using so-called aggregation operators.
In this paper we concentrate on averaging rules (averaging operators) [13]. Fig. 2 illustrates aggregation operators with their linguistic quantifiers.
Ordered Weighted Averaging (*OWA*) Operators represent a distinct family of averaging operators and were introduced by Yager in 1988 [14]. These operators allow aggregations between the pure "AND" and the pure "OR" (cf. Fig. 2). The dispersion of an OWA operator characterizes the degree to which the information of individual arguments is used for the aggregation process. Since, "AND" and "OR" are evaluated by the operators "min" and "max", respectively, the andness of an averaging operator is defined as distance to the max, relative to the distance between the min and the max.

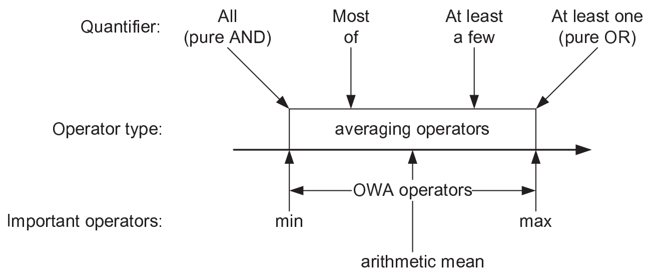

Figure 2: Different rules of aggregation, with their associated operators [13].

2.2 Attribute Fusion

With reference to highly complex machine conditioning applications, such as the monitoring of an intaglio machine's wiping process (it will be discussed in the following section), the TLCS approach can be improved by segmenting the information sources into groups. In each group, the experts decide about the health of special attributes in the machine. The extended version of the DST-based approach is illustrated in Fig. 3.

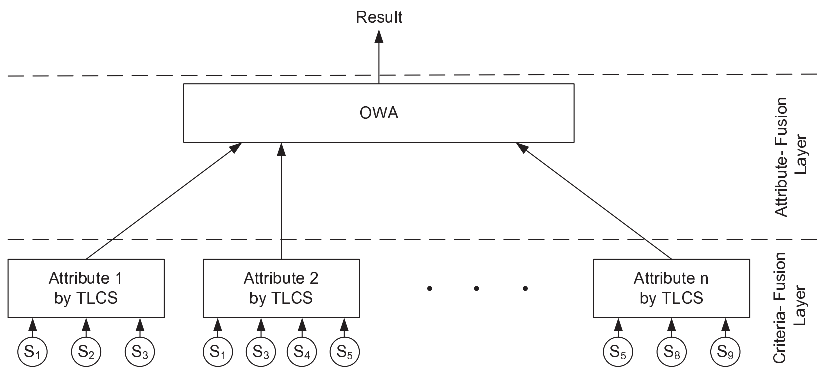

Figure 3: Combining TLCS-results using an OWA operator.

The attribute fusion is accomplished using an OWA operator. Thus, the andness degree can be chosen depending on the causal relation between the attributes:

- If the attributes describe similar effects in a machine, there will be a high redundancy in their information. A low andness degree has to be chosen. Errors are only recognized when all attributes are classified as "bad".
- If the attributes describe different effects in a machine, there will be a low redundancy in their information. A high andness degree has to be chosen. Errors are recognized even when only one attribute is classified as "bad".

This approach leads to a more precise estimation of a machine's health status.

3. Experiments

Nowadays, security prints such as banknotes are highly sophisticated and exhibit many security features. For banknote security, a printing technique that is of major importance is intaglio. This printing method uses a metal plate with engraved characters and structures. During the printing process the engraved structures are filled with ink and pressed under huge pressure directly onto the paper. As a result, a tactile relief and fine lines are formed, unique to the intaglio printing process and almost impossible to reproduce through any commercial printing method. In the intaglio printing machine the wiping unit is the most observed part. It is responsible for removing surplus ink around the engravings. Even small parameter manipulations cause wiping errors (cf. Fig. 4).

Figure 4: Wiping Error (on the Right).

A multi sensory machine conditioning system is observing the wiping unit and estimating its health. Because tests on the intaglio printing machine are expensive and time consuming, a demonstrator was build that simulates the wiping process.
Basically, the simulator consists of two cylinders that rotate in opposite directions. Their principle set-up is illustrated in Fig. 5. The upper small cylinder is coated with a synthetic rubber. This cylinder emulates the wiping cylinder. The large cylinder emulates the plate cylinder of the intaglio printing machine. Similar to the plate cylinder of the intaglio printing machine, it exhibits three gaps, necessary to clamp the printing plates onto the plate cylinder.

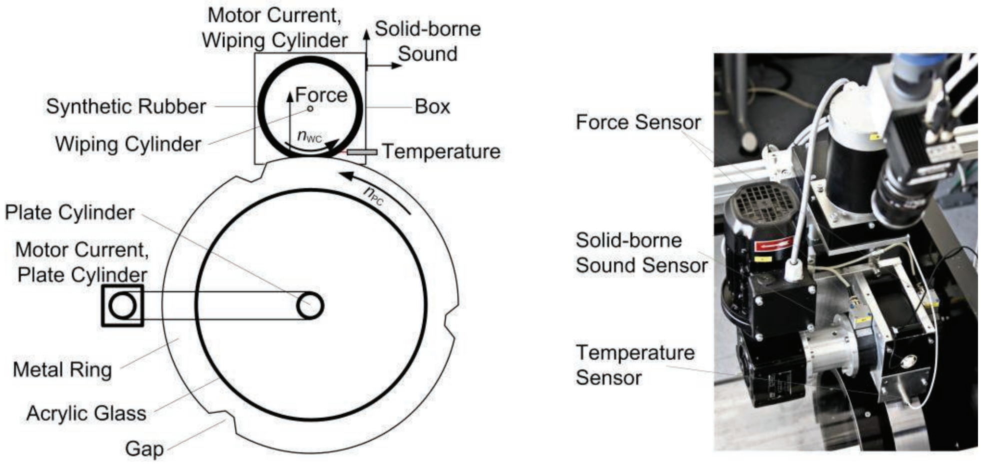

Figure 5: Wiping Simulator.

The sensors, mounted on the simulator, provide following data:

- The motor current of the plate cylinder (MCPC) is mainly influenced by the pressure and the friction between the wiping and the plate cylinder.
- The motor current of the wiping cylinder (MCWC) is mainly influenced by the pressure and the friction between the wiping and the plate cylinder.
- The temperature of the synthetic rubber (T) is mainly influenced by the pressure between the wiping and the plate cylinder.
- The solid-borne sound of the wiping unit is split into the following two signals applying discrete filtering:
  - the sound of the motors (MS) impel the wiping and the plate cylinder, and
  - the sound intensity (SI) is mainly influenced by the contact pressure between both cylinders.
- The contact force of the wiping cylinder (CFWC).

Sensors with similar causal relationships are grouped to the same attribute. The causality relationships are determined using expert knowledge. In this context, a specific machine attribute is related to one particular malfunction of the wiping process. The following two attributes are monitored:

- A1: Lubrication Solvent: In the wiping unit, the ink acts also as a lubrication solvent. In the simulator this effect is monitored by the sensors MCWC, MCPC and MS.
- A2: Contact Pressure: A misalignment of the contact pressure among both the cylinders may cause malfunctions. In the simulator it is observed mainly by the sensors CFWC, SI, and T.

In order to investigate the TLCS-approach and its extension, two experiments were accomplished on the wiping simulation unit. In the first one, a „good” production status was adjusted and kept constant. This experiment was used to learn the inspection system.
In the second experiment parameters were changed and monitored. Fig. 6 shows the fusion results of both, the BalTLCS approach, using all sensors in one group, and the BalTLCS, dividing sensor information into the above mentioned two attributes. Because of a low causal relation, a high andness degree was chosen, and a pessimistic attribute fusion was accomplished.
The result $\pi(x)$ can be interpreted as a gradually health status and ranges between zero (“bad”) and one (“good”).
During the second experiment

- the speed of the drives was changed,
- the pressure between both cylinders was reduced, and
- the lubrication solvent was decreased slowly.

As seen in Fig. 6, both the approaches recognize change of the machine condition. The extended approach (blue line) estimates the overall condition more sensitive compared to the BalTLCS approach (red line), that was introduced in [9]. The sensors are focused on their specialized attribute and provide their opinion mainly in their field of work. Thus, no more irrelevant opinions account.

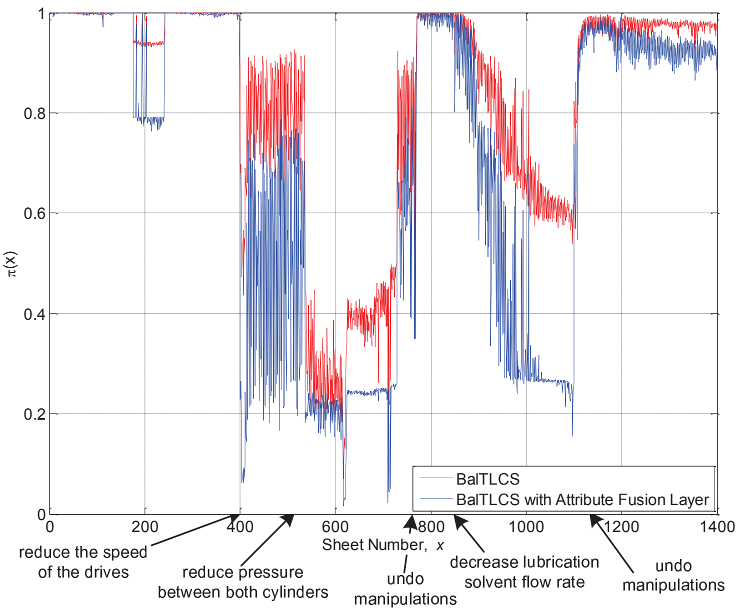

Figure 6: Evaluation of TLCS Approaches on a Wiping Simulator.

4. Conclusion and Outlook

In this paper we introduced an approach for condition monitoring of highly complex tasks and tested it on a wiping simulator of an intaglio printing machine. The inspection system acquires information from several sources, uses different physical quantities, integrates expert knowledge and perception, and generates intuitive results. We used the TLCS approach, which is based on the *Evidence Theory* and uses conflict solving to fuse data. We also introduced an extension of the TLCS approach with reference to highly complex machine conditioning applications. In this context, the sensors were grouped to attributes by applying expert knowledge. The fusion of the fuzzyfied sensors' observations that are elements of one particular attribute is accomplished by TLCS. Subsequently, the attributes' conditions are merged using an *Ordered Weighted Averaging* operator.

So far we assumed that the machine works on a constant working point. The next step is to develop an adaptive learning algorithm, which is expanding its knowledge and also forgetting irrelevant details. Furthermore, we intend to extend our approach to sensor's condition and reliability monitoring.

References

[1] X. F. Fan and M. J. Zuo, "Fault diagnosis of machines based on D-S evidence theory. Part 1: D-S evidence theory and its improvement," Pattern Recognition Letters, Vol. 27, Issue 5, pp. 366-376, 2006.

[2] X. F. Fan and M. J. Zuo, "Fault diagnosis of machines based on D-S evidence theory. Part 2: D-S evidence theory and its improvement," Pattern Recognition Letters, Vol. 27, Issue 5, pp. 377-385, 2006.

[3] J. Schubert, "An information fusion demonstrator for tactical intelligence processing in network-based defense," Int. J. Information Fusion, Vol. 8, Issue 1, 2007.

[4] A. P. Dempster, "Upper and Lower Probabilities induced by a multi-valued mapping," Annals of Mathematical Statistics, 38:325-339, 1967.

[5] G. Shafer, "A Mathematical Theory of Evidence", Princeton University Press, 1976.

[6] L. A. Zadeh, "A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination," AI Magazine, 7:85-90, 1986.

[7] F. Smarandache and J. Dezert, "Advances and Applications on DSmT for Information Fusion," Collected Works 1 (2), American Research Press, Rehboth, 2004 (2006).

[8] R. Li and V. Lohweg, "A Novel Data Fusion Approach using Two-Layer Conflict Solving," IAPR Workshop on Cognitive Information Processing, Santorini, Greece, 2008.

[9] V. Lohweg and U. Mönks, "Sensor Fusion by Two-Layer Conflict Solving", The 2nd International Workshop on Cognitive Information Processing, Italy, 2010.

[10] R. Lipshitz, G. Klein, J. Orasanu, et al., "Taking Stock of Naturalistic Decision Making," Journal of Behavioral Decision Making, 14:331-352, 2001.

[11] S. A. Sunita, "Order effects and memory for evidence in individual versus group decision making in auditing," Journal of Behavioral Decision Making, Vol. 12, Issue 1, pp. 71 – 88, 1999.

[12] G. J. Klir and Y. Yuan, "Fuzzy Set and Fuzzy Logic: Theory and Applications", Prentice-Hall Inc., New York, 1995.

[13] H. L. Larsen, "Importance Weighted OWA Aggregation of Multicriteria Queries", 18th International Conference of the North American Fuzzy Information Processing Society, 740–744, 1999.

[14] R. R. Yager, "On Ordered Weighted Averaging Aggregation Operators in Multi-criteria Decision Making", IEEE Transactions on Systems, Man and Cybernetics 18, 183–190, 1988.