Hierarchical and conditional combination of belief functions induced by visual tracking

John Klein, Christèle Lecomte, Pierre Miché

Hierarchical and conditional combination of belief functions induced by visual tracking

John Klein[a,*], Christèle Lecomte[b], Pierre Miché[b]

[a] *LAGIS - UMR CNRS 8146, University of Lille1, France*
[b] *LITIS - EA 4108, University of Rouen, France*

Abstract

In visual tracking, sources of information are often disrupted and deliver imprecise or unreliable data leading to major data fusion issues. In the Dempster-Shafer framework, such issues can be addressed by attempting to design robust combination rules. Instead of introducing another rule, we propose to use existing ones as part of a hierarchical and conditional combination scheme. The sources are represented by mass functions which are analyzed and labelled regarding unreliability and imprecision. This conditional step divides the problem into specific sub-problems. In each of these sub-problems, the number of constraints is reduced and an appropriate rule is selected and applied. Two functions are thus obtained and analyzed, allowing another rule to be chosen for a second (and final) fusion level. This approach provides a fast and robust way to combine disrupted sources using contextual information brought by a particle filter. Our experiments demonstrate its efficiency on several visual tracking situations.

Key words: Dempster-Shafer Theory, Combination Rules, Visual tracking

1. Introduction

Visual Tracking (VT) consists in locating one or more objects throughout a video using visual information processing. Existing approaches can be improved in two main ways: by designing more precise models using machine learning techniques and/or by introducing a data fusion step that makes the observation/model matching more robust. In this article, we follow the latter path. VT raises a challenging data fusion problem as sources involved in the process can, from time to time, be highly imprecise or unreliable.

In terms of data fusion, the Dempster-Shafer theory (DST) [33] has gained popularity because it can process data that are not only uncertain but also imprecise. Using this framework, one usually aggregates different sources of information using a combination rule. The fusion process underlying a combination rule is regulated by properties. VT specific fusion requirements are mainly related to imprecision and unreliability. These requirements can be expressed in terms of combination rule properties. A rule possessing all the required properties can thus be expected to lead to better VT performances.

Unfortunately, this constraint satisfaction problem appears to have no ideal solution. Some rules can adapt their behaviours to either highly imprecise sources [44, 21] or to unreliability [3, 11] but the conjunction of these two kinds of sources is much more difficult to deal with. To overcome this difficulty, broader fusion schemes can be designed. In [39], Smets addresses conflict management using a conditional scheme that makes use of particular rules depending on assumption rejections or validations. Ha-Duong [12] presented a hierarchical scheme for the fusion of expert groups' opinions. Denoeux's cautious rule [6] is used for fusion within each group and the outputs are then fused using the disjunctive rule. Quost *et al.* [29, 30] also introduced a similar two-level fusion scheme for classifier combination. Sources are clustered according to

*Corresponding author
Email addresses: john.klein@univ-lille1.fr (John Klein), christele.lecomte@univ-rouen.fr (Christèle Lecomte), pierre.miche@univ-rouen.fr (Pierre Miché)

their pairwise dependencies. A within-cluster rule is then designed and a second between-cluster rule is applied to the outputs of the first one.
In this article, a hierarchical and conditional combination scheme (HCCS) is presented so as to address a VT-related fusion challenge: highly imprecise and unreliable source combination. First, a source analysis step identifies highly imprecise sources and unreliable sources, yielding groups of sources. Within each group, the fusion problem is less constrained and is solved using a single rule. Our approach is also hierarchical because a second fusion level is needed to aggregate the outputs of each group. These outputs are analysed as well, allowing a final rule selection whose application yields the fusion result.
The first section of this paper presents general facts about belief functions. The second section reviews combination rules and their properties, and discusses the interest of more refined fusion schemes. The third section focuses on the analysis of the VT problem and its implications on the HCCS proposed. The scheme is then presented in detail. Finally, the contribution of the scheme is demonstrated in the fourth part through a VT algorithm: evidential particle filtering. The experiments show that our method outperforms classical combination rules, and allows for more robust multiple-source object tracking.

2. Dempster-Shafer Theory: fundamental concepts

DST provides a formal framework for dealing with both imprecise and uncertain data. The finite set of mutually exclusive solutions is denoted by $\Omega = \{\omega_1, ..., \omega_K\}$ and is called the frame of discernment. The set of all subsets of Ω is denoted by 2^Ω. A source S collects pieces of evidence leading to the assignment of belief masses to some elements of 2^Ω. The mass of belief assigned to A by S is denoted $m[S](A)$. For the sake of simplicity, the notation $m[S_1]$ is replaced by m_1 hereafter. The function $m : 2^\Omega \to [0,1]$ is called **basic belief assignment** (**bba**) and is such that $\sum_{A\subseteq\Omega} m(A) = 1$. The set of all bbas is denoted $\mathfrak{B}^\Omega$. A set A such that $m(A) > 0$ is called a **focal element**. Two elements of 2^Ω represents hypotheses with noteworthy interpretations:

- $\emptyset$: the solution of the problem may not lie within Ω.
- Ω: the problem's solution lies in Ω but is undetermined.

The open-world assumption states that $m(\emptyset) > 0$ is possible. The closed-world assumption bans $\emptyset$ from any belief assignments. Under the closed-world assumption, the standard way of combining distinct[1] pieces of evidence m_1 and m_2 is Dempster's combination rule $\oplus$:

$$\forall A \neq \emptyset, m_\oplus(A) = \frac{1}{1-\kappa} \sum_{B,C|B\cap C=A} m_1(B)\, m_2(C), \quad (1)$$

$$\text{with } \kappa = \sum_{B,C|B\cap C=\emptyset} m_1(B)\, m_2(C). \quad (2)$$

The mass κ is also called the degree of conflict. The open-world counterpart of Dempster's rule is the conjunctive rule $\textcircled{\cap}$. The rule equation is the same as Dempster's without normalization factor and $m_{\textcircled{\cap}}(\emptyset) = \kappa$ (see Table 1).
A bba is denoted A^w if it has two focal elements: $A \neq \Omega$ and Ω, and if:

$$A^w(A) = 1 - w \text{ and } A^w(\Omega) = w. \quad (3)$$

Such bbas are called **simple bbas** (sbbas). By extension of this notation, the bba denoted Ω^0 stands for total ignorance ($\Omega^0(\Omega) = 1$); it is called the vacuous bba. The one for total conflict is $\emptyset^0$. A bba such that $m(\Omega) = 0$ is said to be **dogmatic**. It is said to be **normalized** if $m(\emptyset) = 0$.

[1] Pieces of evidence are distinct if the construction of beliefs according to one piece of evidence does not restrict the construction of beliefs using another piece of evidence.

There are other ways of representing beliefs, including the belief *bel*, implicability b, plausibility pl and commonality q functions. In view of some further developments in this article, a brief presentation of the conjunctive weight function w is also needed. Smets [37] has shown that a non-dogmatic bba can be decomposed into a conjunctive combination of generalized simple bbas (gsbbas). A gsbba $\mu : 2^\Omega \longrightarrow \mathbb{R}$ is a sbba whose focal element $A \subset \Omega$ is assigned $1 - w$ with $w \in [0, +\infty)$. Depending on the value of w, two cases can be distinguished:

- if $w \leq 1$: these gsbbas are sbbas.
- if $w > 1$: these gsbbas are not bbas and are called inverse sbbas (isbbas).

An isbba with focal element A is interpreted as a representation of the belief that there exists some reason *not to* believe in A. In other words, it constitutes a "debt" of belief, hence the notation $A^{1/w}$ for an isbba. Smets shows that any non-dogmatic bba m can be expressed as a conjunctive combination: $m = \bigcirc\!\!\!\!\!\cap_{A \subset \Omega} A^{w(A)}$ with $w : 2^\Omega \longrightarrow [0, +\infty)$ a function such that:

$$\forall A \subset \Omega, w(A) = \prod_{B \supseteq A} \left(\sum_{C \supseteq B} m(C) \right)^{-1^{|B|-|A|+1}} . \tag{4}$$

The function w can be obtained from m and conversely. Detailed definitions of belief representation functions can be found in [6]. Besides, if a source S of information is known to be unreliable, then it is possible to reduce its impact using an operation called discounting [33]. Discounting with discount rate $\alpha \in [0, 1]$ is defined as:

$$m[\alpha, S](X) = \begin{cases} (1-\alpha) m[S](X) \text{ if } X \neq \Omega, \\ (1-\alpha) m[S](X) + \alpha \text{ if } X = \Omega. \end{cases} \tag{5}$$

The higher α is, the stronger the discounting. Thanks to discounting, an unreliable source's bba is transformed into a function closer to the vacuous bba. Mercier *et al.* [23] presented a refined discounting, in which discount rates are computed for each subset and each source. The discounting is consequently more precise and efficient. It is, however, necessary to have enough information allowing subset-specific computation.
Partial ordering relationships can be defined on $\mathfrak{B}^\Omega$ according to bba information content. Two examples of partial orderings are:

- *pl*-ordering: $m_1 \sqsubseteq_{pl} m_2$ iff $pl_1(A) \leq pl_2(A)$, for all $A \subseteq \Omega$;
- q-ordering: $m_1 \sqsubseteq_q m_2$ iff $q_1(A) \leq q_2(A)$, for all $A \subseteq \Omega$.

If $m_1 \sqsubseteq_x m_2$, then m_1 is said to be x-more committed than m_2, meaning that m_1 is more informative than m_2 in solving the problem. Formal definitions of other partial orderings as well as their dependencies can be found in [6].

3. Information fusion in DST

This section is a short review of information fusion techniques in DST. Several combination rules, their properties and more complex fusion schemes are presented.

3.1. Combination rules

For rules without pre-defined symbols, the notation $\odot_{xx}$ refers to the combination rule named xx, and m_{xx} is understood as the bba resulting from the combination using $\odot_{xx}$. Combination rule equations are gathered in Table 1.
Both Dempster's rule and the conjunctive rule transfer beliefs to intersections of subsets. In contrast,

the disjunctive rule, introduced by Dubois and Prade [8] and denoted $\textcircled{\cup}$, transfers beliefs to unions of subsets. This rule is based on a different assumption on the reliability of sources[2]. Concerning this aspect, the conjunctive and disjunctive combinations are two extreme cases, and consequently some authors have proposed rules in between these two cases. Dubois and Prade [9] introduced another rule, denoted DPR, that combines sources conjunctively but reallocates κ disjunctively.

Smets [38] generalized conjunctive and disjunctive rules by introducing the family of α-junction rules. The coefficient $\alpha \in [0, 1]$ can be seen as a degree of conjunction or disjunction. The exclusive disjunctive rule, denoted $\underline{\textcircled{\cup}}$, belongs to this family of rules and transfers beliefs using the symmetric difference of subsets. The interest of this rule is mainly theoretical because it can be interpreted as a solution to a specific data fusion problem (see Section 3.2.2).

More recently, Florea *et al.* [11] introduced robust combination rules (RCRs). These rules average the conjunctive and disjunctive rules using conflict-dependent weights. In the rest of the paper, RCR refers to the robust rule with the weight functions recommended by the authors (see Table 1). This family of rules was extended by Martin *et al.* [21] resulting in the so-called mix rules. In this extension, the weights are functions of pairs of subsets and do not depend mandatorily on κ. In the rest of the paper, Martin's mix rule (MMR) refers to the mix rule with the weight functions recommended by the authors (see Table 1).

Note that MMR weights take subset cardinalities into account. The larger a cardinality, the more imprecise the hypothesis. Consequently, imprecision influences the result of MMR computation. Using cardinalities for weighting belief transfers was first suggested by Zhang [44] and applied to Dempster's rule. However, Zhang's rule (ZR) output bba must be renormalized after combination.

Many authors have tried to work on Dempster's rule basis by reallocating κ in different manners. Yager's rule (YR) [42] transfers it directly to the ignorance. Inagaki [13] designed a family of rules dealing with conflict reallocation. This family was extended by Lefevre *et al.* [19]. The main idea behind these rules is to distribute κ to some subsets according to an appropriate scheme.

Using non-linear functions, Dezert and Smarandache's PCR5 [35] redistributes conflict to the subsets from which it was generated. Martin *et al.* [36] generalized it for more than two sources. This generalization is known as PCR6. The same authors [21] also integrated a discounting mechanism in PCR6 resulting in the so-called discounted PCR (DPCR). They further proposed to combine the DPCR and the mix rule into the mix DPCR (MDPCR).

Delmotte *et al.* [3] investigated the integration of reliability coefficients $\{R_i\}_{i=1}^{M}$ in two combination rules. The first rule, referred to as Delmotte's averaging rule (DAR), averages input bbas using $\{R_i\}_{i=1}^{M}$ as weights. Note that averaging was also proposed by Murphy [25] and Jøsang [31]. The second rule, referred to as Delmotte's mix rule (DMR), is a mix rule with weight functions depending on $\{R_i\}_{i=1}^{M}$ (see [4] for a detailed definition of these functions).

Another family of rules was recently introduced by Denoeux [6]. Only the two most significant ones are examined in this paper: the cautious rule $\textcircled{\wedge}$ and the bold rule $\textcircled{\vee}$. They are based, respectively, on a conjunctive combination of conjunctive weight functions w and a disjunctive combination of disjunctive weight functions v[3]. Note that $\textcircled{\wedge}$ can only be applied to non-dogmatic bbas and, similarly, $\textcircled{\vee}$ can only be applied to non-normalized bbas. This problem can be solved by allocating a minimal residual belief ζ to $\emptyset$ and Ω, respectively. Other rules are listed in [32, 39, 34]. It is not intended in this article to review all existing rules, but only the most popular or relevant ones for our study.

3.2. Properties of combination rules

The way rules manipulate sources of information is described by several properties. These properties and their relevancies are briefly reviewed in this subsection. Formal definitions are not included but can be found in [39, 34, 32].

[2]The terms conjunctive and disjunctive and the underlying assumptions are presented in Section 3.2.2

[3]The function v is the disjunctive counterpart of function w, see [6] for a detailed definition of function v.

Table 1: Several combination rule equations for two sources of information. $A, B, C \subseteq \Omega$. Z is a normalization factor. A_v is a bba such that $A_v(\emptyset) = v$ and $A_v(A) = 1 - v$. $\Xi(A,B) = \frac{m_1(A)^2 m_2(B)}{m_1(A)+m_2(B)} + \frac{m_2(A)^2 m_1(B)}{m_2(A)+m_1(B)}$. ϵ is a discounting coefficient. R_i is the reliability coefficient of function m_i. $\wedge$ is the minimum operator. (*): closed-world assumption.

$\oplus$ (*)	$m_{\oplus}(A) = \frac{1}{1-\kappa} m_{\textcircled{\cap}}(A), \kappa = m_{\textcircled{\cap}}(\emptyset)$
DPR (*)	$m_{dpr}(A) = m_{\textcircled{\cap}}(A) + \sum_{B\cap C=\emptyset, B\cup C=A} m_1(B) m_2(C)$
$\textcircled{\cap}$	$m_{\textcircled{\cap}}(A) = \sum_{B\cap C=A} m_1(B) m_2(C)$
RCR (*)	$m_{rcr}(A) = \frac{\kappa}{1-\kappa+\kappa^2} m_{\textcircled{\cup}}(A) + \frac{1-\kappa}{1-\kappa+\kappa^2} m_{\textcircled{\cap}}(A)$
$\textcircled{\cup}$	$m_{\textcircled{\cup}}(A) = \sum_{B\cup C=A} m_1(B) m_2(C)$
MMR	$m_{mmr}(A) = \sum_{B\cup C=A} \left(1 - \frac{\lvert B\cap C\rvert}{\lvert B\cup C\rvert}\right) m_1(B) m_2(C) + \sum_{B\cap C=A} \frac{\lvert B\cap C\rvert}{\lvert B\cup C\rvert} m_1(B) m_2(C)$
$\underline{\textcircled{\cup}}$	$m_{\underline{\textcircled{\cup}}}(A) = \sum_{B\Delta C=A} m_1(B) m_2(C)$
PCR6	$m_{pcr6}(A) = m_{\textcircled{\cap}}(A) + \sum_{B\cap A=\emptyset} \Xi(A,B)$
$\textcircled{\wedge}$	$m_{\textcircled{\wedge}}(A) = \textcircled{\cap}_{A\subset\Omega} A^{w_1(A)\wedge w_2(A)}$
DPCR	$m_{dpcr}(A) = m_{\textcircled{\cap}}(A) + \sum_{B\cap A=\emptyset} \epsilon\Xi(A,B) + \sum_{B\cup C=A, B\cap C=\emptyset} (1-\epsilon) m_1(B) m_2(C)$
$\textcircled{\vee}$	$m_{\textcircled{\vee}}(A) = \textcircled{\cup}_{A\subset\Omega} A_{v_1(A)\wedge v_2(A)}$
MDPCR	$m_{mdpcr}(A) = \sum_{B\cup C=A, B\cap C\neq\emptyset} \left(1 - \frac{\lvert B\cap C\rvert}{\lvert B\cup C\rvert}\right) m_1(B) m_2(C) + \sum_{B\cap A=\emptyset} \epsilon\Xi(A,B)$ $+ \sum_{B\cap C=A, B\cap C\neq\emptyset} \frac{\lvert B\cap C\rvert}{\lvert B\cup C\rvert} m_1(B) m_2(C) + \sum_{B\cup C=A, B\cap C=\emptyset} (1-\epsilon) m_1(B) m_2(C)$
ZR (*)	$m_z(A) = \frac{1}{Z} \sum_{B\cap C=A} \frac{\lvert A\rvert}{\lvert B\rvert \lvert C\rvert} m_1(B) m_2(C)$
DAR (*)	$m_{dar}(A) = R_1 m_1(A) + R_2 m_2(A)$
YR (*)	$m_y(A) = m_{\textcircled{\cap}}(A), m_y(\Omega) = \kappa + m_{\textcircled{\cap}}(\Omega)$
DMR	$m_{dmr}(A) = \frac{1}{Z}\left[R_1 R_2 m_{\textcircled{\cup}}(A) + [R_1 R_2 - R_1 - R_2](1 - R_1 R_2) m_{\textcircled{\cap}}(A)\right]$

3.2.1. Algebraic properties

In these paragraphs, some of the most usual algebraic properties are listed:

- **Commutativity and associativity**: combined with commutativity, associativity allows a source-order-independent combination. The same goal can be reached by substituting quasi-associativity introduced by Yager [43] for associativity. If pieces of evidence are all available at the time of the combination (batch mode), then an $n-ary$ version of the rule suffices. If pieces of evidence are collected sequentially, then an updating scheme is often needed to avoid high computational cost.

- **Idempotence**: using an idempotent rule, no elementary piece of evidence is counted twice which is helpful in the case of non-distinct evidences. In practice, most pieces of evidences coming from sources relying on the same observation are not distinct but overlapping.

- **Existence of a neutral or absorbing element**: Some rules are designed so that the vacuous bba Ω^0 has no impact on the fusion result, *i.e.* the vacuous bba is a neutral element of the rule.

Table 2 lists the combination rules examined and their algebraic properties.

3.2.2. Conjunctive vs. disjunctive behaviours

The nature of a combination is closely related to informational ordering. Using a conjunctive rule, the result of the combination is more committed than each aggregated bba. Thus, conjunctive combinations

Table 2: Several combination rules and their properties - part 1. ×: the rule has the corresponding property. COM=commutativity, ASSO=associativity, Q-ASSO=quasi-associativity and IDEM=idempotence

	Algebraic properties					Algebraic properties			
	COM	ASSO	Q-ASSO	IDEM		COM	ASSO	Q-ASSO	IDEM
$\oplus$	×	×			DPR	×		×	
$\textcircled{\cap}$	×	×			RCR	×		×	
$\textcircled{\cup}$	×	×			MMR	×			
$\underline{\textcircled{\cup}}$	×	×			PCR6	×		×	
$\textcircled{\wedge}$	×	×		×	DPCR	×		×	
$\textcircled{\vee}$	×	×		×	MDPCR	×			
ZR	×				DAR	×		×	×
YR	×		×		DMR	×		×	

Table 3: Several combination rules and their properties - part 2. ×: the rule has the corresponding property. CONJ=conjunctive, κ-CONJ=purely conjunctive if $\kappa = 0$ and partially conjunctive and disjunctive if $\kappa > 0$, CONJ/DISJ=partially conjunctive and disjunctive, DISJ=disjunctive and OTHER=other kind

	Combination nature related properties						Combination nature related properties				
	CONJ	κ-CONJ	CONJ/DISJ	DISJ	OTHER		CONJ	κ-CONJ	CONJ/DISJ	DISJ	OTHER
$\oplus$	×					DPR		×			
$\textcircled{\cap}$	×					RCR		×			
$\textcircled{\cup}$				×		MMR			×		
$\underline{\textcircled{\cup}}$					×	PCR6		×			
$\textcircled{\wedge}$	×					DPCR		×			
$\textcircled{\vee}$				×		MDPCR			×		
ZR	×					DAR					×
YR					×	DMR			×		

and the underlying commitments are appropriate when sources tell the truth, *i.e.* are reliable.
Conversely, a disjunctive rule produces a bba that is less committed than the ones from which it originated. However, disjunctive combinations are appropriate when some of the sources tell the truth but we do not know which ones. Under such circumstances, it is too risky to commit oneself to one of the pieces of evidence. Further comments on conjunction and disjunction can be found in [8, 38, 39].
There is a third category of combination nature, if it is thought that only one source is true, but it is not known which one. This view corresponds to the exclusive disjunctive rule. Generally speaking, α-junction rules, whose behaviours are in between conjunctive and disjunctive, are not easy to interpret (see [38, 28]). The exclusive disjunctive rule assumption does not suit VT context and therefore is no longer discussed in this article.
Table 3 shows which class of behaviour a rule belongs to. A majority of rules are partially conjunctive and disjunctive. Some of them use $\textcircled{\cup}$ only as part of the conflict redistribution (denoted κ-CONJ in table 3). The others use $\textcircled{\cup}$ in a broader way (denoted CONJ/DISJ in table 3).

3.2.3. Subsets related properties

In this section, we introduce two properties arising from VT application. Using a conjunctive rule, the mass allocated to $\emptyset$ increases and that of Ω decreases. Similarly, using a disjunctive rule the mass allocated to Ω increases and that of $\emptyset$ decreases. We denote respectively these phenomena by: $\emptyset \succ \{\textcircled{\cap}, \textcircled{\wedge}\}$, $\Omega \succ \{\textcircled{\cup}, \textcircled{\vee}\}$, $\emptyset \dashv \{\textcircled{\cup}, \textcircled{\vee}\}$ and $\Omega \dashv \{\textcircled{\cap}, \textcircled{\wedge}\}$. These phenomena are a consequence of the nature of the combination and cannot be avoided but, in practice, they may yield bbas assigning excessive masses to $\emptyset$ or Ω. In the end, the result of the fusion may be unexploitable (see a remark of Smets in the very last paragraph of [39]).
A generalization of these phenomena can be formalized for any subset and any rule but, in this article, our practical needs are limited to monitoring rule behaviours toward $\emptyset$ and Ω.

3.3. Multiple-rule combination schemes

Clearly, the more properties needed, the more difficult it is to design an appropriate rule as in any constraint satisfaction problem. If a property is needed in a situation A whereas in a situation B an antagonistic property is needed, then a conditioning step can be used to identify the present situation and then select a rule. Even after conditioning, and therefore reducing the number of constraints, some incompatibilities may remain. It is then necessary to define priorities among these properties, *i.e.*, a hierarchy.
As explained in the introduction, more complex fusion schemes can be designed in order to overcome single rule limitations. A discounting conditioned by contextual data can be applied to bbas before combination. This process can tune bbas, alleviating inconvenient evidences that would prevent bbas from being efficiently processed by a single rule. The conditions under which a bba must be discounted are context-dependent. Generally speaking, discounting should not be opposed to other methods because it is a complementary tool that can help to solve simply some unreliability related problems.
Some rules such as DAR or DMR enclose a conditioning step because the output bba depends on external data. Kallel and Le Hégarat-Mascle [15] worked on partial distinctness. The proposed rule, called the cautious-adaptative rule, varies from the conjunctive rule to the cautious rule depending on a parameter $Q \in [0, 1]$. The value of Q is obtained using *a priori* knowledge on evidence distinctness.
In [39], Smets proposes a conditional scheme. A large number of conditions are examined leading to some rule selection and discounting. This scheme is limited to conflict management issues. Furthermore, only one rule is selected and applied to all bbas. The assumption that one rule meets all requirements is not verified in our VT application even if it is carefully selected (see 4.2.2).
Ha-Duong [12] proposes a hierarchical approach based on two combination rules. This approach is designed for the fusion of groups of experts' opinions. Within each group of experts the cautious rule is used, and the bbas resulting from these combinations are then aggregated using the disjunctive rule. The process is not conditioned by input bbas and therefore the choice of rules is static. It does not fit VT application in which sources evolve and necessitate dynamic rule selections.
Quost *et al.* [30, 29] designed an approach both hierarchical and conditional. Bbas are clustered regarding bba dependency criteria. An appropriate rule is then applied within each bba cluster. Following a dendrogram hierarchy, other levels of fusion are then needed to aggregate newly generated bbas. The method is dedicated to non-distinct source combination issues. *A priori* information is needed for combination rule learning.
In this article, we intend to develop a hierarchical and conditional combination scheme (HCCS) allowing VT requirements to be met.

4. Evidential fusion scheme induced by visual tracking

In this section, the VT problem is formalized and an evidential particle filter (EPF) is proposed as a solution. Then, the data fusion constraints induced by VT are examined and translated into combination rule property requirements. A hierarchical and conditional combination scheme is introduced so as to deal with these requirements.

4.1. Visual tracking using EPFs

The VT problem can be expressed as follows: the position of a target object must be identified in each image of a video. In this article, a bounding box is used to represent an object position within an image. This representation has the advantage of being coded by only four time-dependent parameters: $(x_{1,t}, x_{2,t}, W_t, H_t)$ with $(x_{1,t}, x_{2,t})$ the box centre coordinates, W_t the box width and H_t the box height at time t. Finding out the actual values of these parameters for each t is equivalent to solving the VT problem.
There are many ways to estimate these parameters. Particle filters (PFs) [14, 26] have gained popularity among the computer vision community because of the compromise they offer in terms of both precision and computation time. They are notably preferred to Kalman filters which are restricted to linear models and Gaussian noises. PFs estimate a state vector X_t whose value is the answer to the problem at stake, hence $X_t = (x_{1,t}, x_{2,t}, W_t, H_t)$. In each image of the sequence, the filter samples several sub-images, *i.e.*, several

Figure 1: A source turning from normal to weak. Red squares are locations to which the source assigns low weights. Green ones are heavily weighted.

values of X_t also called particles and denoted $X_t^{(i)}$. It is then necessary to evaluate to what degree each sub-image is likely to actually contain the target object. This evaluation is a crucial step of the filter and is performed using an observation model. The efficiency of the filter relies on the relevance of the model chosen.

Data fusion is frequently used to design more robust models. Some authors initially proposed Bayesian fusion solutions since PFs rely on probability theory [2, 27]. The efficiency of Bayesian fusion depends on the precision of density models and information delivered by the sources. Some sources in VT applications happen to be highly imprecise whenever a situation causes a source to be out of its observation capacity bounds.

As in many other applications, DST is an alternative to the Bayesian approach. Partial knowledge of the problem can be modelled and a large number of fusion tools are available as shown in the previous section. Thus, several authors have designed evidential particle filters (EPFs). Initially, extensions of the Kalman filter to the DST were proposed in [20, 22, 40]. In [10], a DST step produces features that can be used as in a classical PF. In [2], a PF is used for tracking along with DST for target classification into several object categories. In [41], a particle filter using a DSmT fusion step is presented for multiple target tracking. Each particle observation is compared to each target model. Two cues (location and colour) are combined using a worked out combination rule, yielding a new bba whose belief or plausibility function updates particle weights. In [17], three features are used and combined using the conjunctive rule. The pignistic transform is applied to the combination result so as to update particle weights. In [24], a similar approach is extended to multi-target multi-sensor tracking leading to further model developments.

In this article, we are not interested in how an EPF should be designed, but in determining a bba combination scheme adapted to visual tracking EPFs. We also limit the problem to monocular single object tracking. The EPF employed in our experiments is further presented in appendix A.

4.2. Information source disruptions in a VT context

To understand the relevance of adapting combination techniques to VT, we need to investigate what events disrupt VT algorithms. Occlusion is one of these events. PFs have shown better robustness to occlusions than other VT algorithms thanks to the fact that particles spread out during the course of an occlusion, giving PFs a better chance to detect the object after the occlusion. An occlusion bans access to visual information; image-based sources become ignorant, which is a case of major imprecision. Two other events generate imprecision: illumination changes and particular movements to which feature extraction methods may be sensitive. Figure 1 shows an example of an illumination change causing disruption to a colour-texture feature based source.

Clutter also causes severe disruptions. As opposed to other events, clutter induces unreliability of sources because they may identify two distinct locations for the target object whereas the tracking problem has only one solution. One of the proposed locations is thus wrong and the source delivers wrong information. Figure 2 illustrates a clutter situation causing a shape feature based source to turn unreliable. If the objects are perfectly identical, then one cannot expect to distinguish them. In this extreme case, other tracking techniques, such as multi-target tracking or trajectory analysis, should be applied. In this article, it is

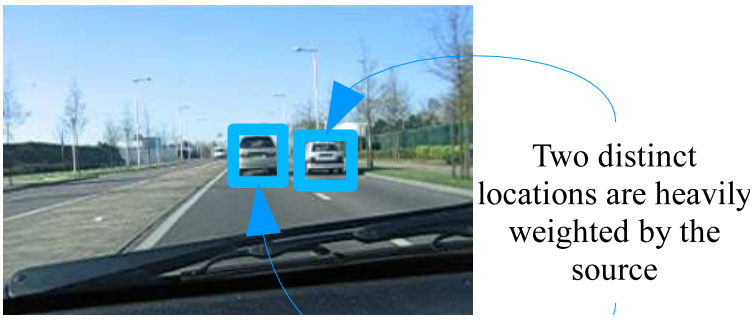

Figure 2: A Source turning unreliable

assumed that the objects can be distinguished using at least one source. However, this capacity may not be constant over time.

In conclusion to the above remarks, VT makes it necessary to design a combination technique taking high imprecision and unreliability into account. In the rest of this document, highly imprecise sources are referred to as **weak** sources. **Normal** sources comprise all sources that are neither weak nor unreliable. Thus, a normal source may contain some imprecision or unreliability which is considered as negligible. We now present our HCCS that is composed of two main steps: bba analysis and hierarchical fusion.

4.2.1. BBA analysis

This step consists in detecting weak and unreliable sources. It relies on the possibility to collect contextual information and therefore it is application dependent. We present here a simple method based on particle filtering.

Only two tests are needed to identify sources as weak and unreliable. The first one is the "weakness test", which separates weak sources from others. The weakness of the bba provided by S_j is determined by thresholding the ignorance $m_j(\Omega)$. To take a safer decision, the contextual information brought by a particle filter can be used. EPFs evaluate $m_j(\Omega)$ for each particle. Let us denote $m\left[S_j, X_t^{(i)}\right]$ the bba of source S_j at the location coded by particle $X_t^{(i)}$. The condition for detection as weak is thus:

$$1 - \min_{i=1..N} m\left[S_j, X_t^{(i)}\right](\Omega) < t_{weak}, \tag{6}$$

with N the number of particles and t_{weak} a threshold defined *a priori*. S_j is labelled as weak if there is no particle yielding $m[S_j](\Omega) < 1 - t_{weak}$. To set t_{weak}, one must identify up to what value labelling a rather imprecise source as weak is risk-free. Under this condition, highly imprecise bbas do not impact inappropriately on the fusion result.

The second test to perform is the "unreliability test" that separates non-weak sources into unreliable sources and normal sources. In this article, unreliability is detected if S_j identifies disjoint parts of the image as containing the tracked object. One of them actually contains the object, whereas the others contain something else that is similar to the object. PFs assign heavy weights to particles located near image regions resembling the target. Thus, it is possible to use these weights as part of a dispersion measure, *disp*. The unreliability condition for source S_j is then:

$$disp > t_{disp}, \text{ with } disp = \sum_{i=1}^{N} \lambda_t^{(i,j)} \left(\left(x_{1,t}^{(i)} - \bar{x}_{1,t} \right)^2 + \left(x_{2,t}^{(i)} - \bar{x}_{2,t} \right)^2 \right), \tag{7}$$

t_{disp} a threshold, and $\lambda_t^{(i,j)}$ the weight generated by the source S_j at time t for particle $X_t^{(i)}$. The weights are computed as if a one-source PF was used. $\bar{X}_t = \left(\bar{x}_{1,t}, \bar{x}_{2,t}, \bar{W}_t, \bar{H}_t\right)$ is the weighted mean of particles:

$\bar{X}_t = \sum_{i=1}^{N} \lambda_t^{(i,j)} X_t^{(i)}$. The mean values of height and width are not used because the size dispersion is not relevant for clutter detection. The value of the threshold is easily obtained because the dispersion value varies significantly when an outlier occurs. In our experiments, the threshold value was fixed dynamically: $t_{disp} = a \min(H_{t-1}, W_{t-1})$, with $a \in \mathbb{R}$.
In the experiments, parameters of the bba analysis step are set by the user with an error/correction procedure. Typical values of t_{weak} and a are shown in the experiments, see Table 8.

4.2.2. First fusion level

Since the appropriate combination behaviour depends on labels, each possible pair of source types must be examined:

- **case 1:** a weak source S_1 combined with any type of source S_2: the result should be a bba close the one provided by S_2. A weak source assigns a large mass to Ω; therefore the chosen rule should lower the ignorance. In some sense, this is a generalization of the neutral impact of the vacuous bba property.
- **case 2:** a normal source S_1 combined with another normal source S_2: both bbas represent reliable pieces of evidence. A conjunctive combination is typically useful in this case so as to extract as much common reliable information as possible. Normal sources are not always fully reliable, so a partially conjunctive rule also fits.
- **case 3:** an unreliable source S_1 combined with a normal or unreliable source S_2: at least one bba contains some erroneous evidence. A fully disjunctive combination is typically useful so as not to discard any possibly relevant evidence. Given that $\Omega \succ \textcircled{\cup}$, it would be unwise to aggregate weak sources using this rule. Consequently, if a source is both weak and unreliable, then the weak aspect prevails over the unreliable aspect.

These three interlocking cases cover all possible pairwise source combinations provided that the rule is commutative, which is implicitly assumed. The fact that the cases are interlocked implies a hierarchy.
Cases 1 and 2 match pretty well, since $\Omega \dashv \{\textcircled{\cap}, \textcircled{\wedge}\}$. However, a strong incompatibility is raised by case 3, since a source cannot be fully disjunctive and partially conjunctive at the same time, hence the need for a fusion scheme instead of a single rule. Weaker requirements can be added: quasi-associativity (to limit the computational cost of the combination) and the ability to process non-distinct sources (the process should be as generic as possible). The order with which rules are applied and bbas combined is obviously important. As a consequence, HCCS is not associative but if the chosen rules are quasi-associative, so is HCCS.
The combination of requirements from cases 1 and 2 indicates that bbas detected as normal or weak can be jointly fused using a conjunctive rule; $\textcircled{\cap}$ and $\textcircled{\wedge}$ are potential candidates. Case 3 requires a fully disjunctive rule, therefore $\textcircled{\cup}$ and $\textcircled{\vee}$ are potential candidates. Following the secondary requirement concerning non-distinct sources, bold and cautious rules should be preferred. However, because of the bba model used in our experiments, most of the bbas produced are normalized, which impairs the use of the bold rule. The experiments will show that the bold rule's performances are pretty poor *c.f.* Subsection 5.4. The proposed assignment is the following:

- weak and normal bbas, as well as bbas that are both unreliable and weak, are jointly aggregated with rule $\textcircled{\wedge}$;
- non-weak unreliable bbas are aggregated with rule $\textcircled{\cup}$.

Figure 3 summarizes the first fusion level.

4.2.3. Second fusion level

The first combination step yields two bbas: $m_{\textcircled{\cup}}$ and $m_{\textcircled{\wedge}}$. To fuse these two bbas, the first fusion step can be applied to them if they are identified as normal, weak or unreliable. Normality is a dominant character compared to weakness when using $\textcircled{\wedge}$, so $m_{\textcircled{\wedge}}$ is weak only if all the bbas combined with $\textcircled{\wedge}$ are weak. $m_{\textcircled{\wedge}}$

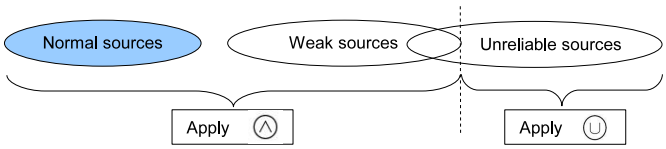

Figure 3: First fusion level of HCCS

is normal otherwise. As a combination of unreliable sources $m_{\textcircled{\cup}}$ is also unreliable. As a consequence of $\Omega \succ \textcircled{\cup}$, $m_{\textcircled{\cup}}$ can be imprecise too. This imprecision is considered to be artificial and is not taken into account.
Following requirements expressed in Subsection 4.2.2: if $m_{\textcircled{\wedge}}$ is weak, $m_{hccs} = m_{\textcircled{\wedge}}\textcircled{\wedge}m_{\textcircled{\cup}}$ (case 3) and if $m_{\textcircled{\wedge}}$ is normal, $m_{hccs} = m_{\textcircled{\wedge}}\textcircled{\cup}m_{\textcircled{\cup}}$ (case 1).
Note that, compared to related works, HCCS is not more computationally costly than a single rule in a batch mode, since the same number of aggregations is needed. In a sequential mode, a incoming bba is integrated to $m_{\textcircled{\cup}}$ or $m_{\textcircled{\wedge}}$ by associativity, and the final fusion result is obtained by repeating the second fusion substep, which is just one combination of two bbas.

5. Experiments

In this section, HCCS is evaluated in terms of VT performances. We use the EPF proposed in [17]. Sources are limited to visual information extractors. These extractors work on the same experiment (the image sequence), and consequently sources are not necessarily distinct. So as to respond differently to disrupting events, some of the following sources are used in the experiments:

- colour-texture source $\mathbf{S_{c-t}}$. The colour-texture extraction method is based on cooccurrence matrices [16]. Pairs of colours of neighbour pixels are counted and stored in these matrices.
- colour source $\mathbf{S_c}$. As in [26], a colour density is used for this source. A colour density is a 3D colour histogram in which colour occurrences are weighted by the distance from the centre pixel of the image region.
- shape source $\mathbf{S_s}$. The shape feature is a symmetry card [1]. Each image region column is considered as a potential axis of symmetry and equally distant pixels are compared using a colour distance.
- motion source $\mathbf{S_m}$. The motion is treated by detecting movement significantly different from the global motion of the scene [18]. A histogram of movement is built from pixel movement intensities.

By cleverly combining these extractors' assets, tracking can be maintained even if all kinds of previously cited events occur. When HCCS performances are compared with another fusion technique, it is important to stress that the same sources and the same EPF are used. We present below some implementation details before showing experimental results.

5.1. Implementation

This subsection describes the frame of discernment on which bbas are defined and how visual information is processed to construct bbas. These bbas feed the fusion technique which itself feeds the EPF.

5.1.1. Frame of discernment

The frame of discernment is defined as $\Omega = \{\omega_1, \omega_2, \omega_3\}$, with

- ω_1: *the sub-image contains the targeted object.*
- ω_2: *the sub-image contains a piece of the scene background.*
- ω_3: *the sub-image contains any other object independent from background.*

Ω can be reasonably thought to cover all possibilities and to be composed of exclusive hypotheses. In this frame of discernment, our goal is now to estimate each source bba and then to combine them using HCCS.

Table 4: Focal elements of implemented sources and their semantics.

source	focal element	semantics
S_{c-t}	$\{\omega_1\}$	the sub-image contains the object
	$\{\omega_2\}$	the sub-image contains a piece of the scene background
S_c	$\{\omega_1\}$	the sub-image contains the object
S_s	$\{\omega_1\}$	the sub-image contains the object
S_m	$\{\omega_1, \omega_3\}$	the sub-image contains an object independent from the background

5.1.2. BBA construction

For any source S_j and particle $X_t^{(i)}$, features are extracted from the analysed sub-image and compared to reference features known *a priori*. The set of reference features constitute the object model (or models if other elements of the scene are analysed). Model learning is rudimentarily performed on the first image of the processed sequence, since at $t = 0$ the object location is known. By matching a model with the observations drawn in the sub-image a distance d_A [4] can be computed, with $A \subsetneq \Omega$ a hypothesis in accordance with the model semantics. Indeed, model/observation matching is not possible for any subset A, it depends on the information source and its interpretation regarding the VT problem. Focal element selection is summarized in Table 4.

The Bhattacharyya distance is an efficient metric for histograms [26] and it can be directly applied to all outputs of feature extraction techniques presented above. 2D or 3D feature arrays are processed as 1D arrays so as to obtain d_A. A Gaussian model is used to define simple bbas:

$$m_A[S_j] \quad = \quad A^{1-\exp\left(-\frac{d_A^2}{2\sigma_{A,j}^2}\right)}, \tag{8}$$

with $\sigma_{A,j}$ the standard deviation of the Gaussian function. These parameters are supposed to be fixed *a priori*; in our experiments they were set using an error/correction procedure.
The global bba for the colour-texture source is obtained by fusing the two colour-texture sbbas: $m[S_{c-t}] = m_{\{\omega_1\}}[S_{c-t}] \textcircled{\wedge} m_{\{\omega_2\}}[S_{c-t}]$. Other source bbas are obtained straightforwardly since they are composed of only one sbba. For practical reasons, it was not possible to design more sbbas; for example, the shape of the background is generally impossible to model. The movement source is much more imprecise than others; therefore, the choice of focal elements for this source reflects its imprecision. Note that this bba construction model can be regarded as a 1-ppv version of Denoeux's model [4]. Further information on more refined bba models can be found in [5, 7].

5.2. Results and discussions

VT results were ground-truth-validated using a tracking rate measure $r \in [0, 1]$, more widely known as the dice coefficient. This measure evaluates a tracking algorithm performance by comparing its estimation of vector X_t to a ground-truth estimation of the same vector: $r = \frac{2 \times S(A \cap B)}{S(A)+S(B)}$. S is the surface area of a set in pixels, A is the estimated object bounding box, and B is the ground-truth object bounding box. The closer r is to 1, the more precise the algorithm. The smaller the variations of r, the more robust the algorithm.
In the following paragraphs, tracking efficiency is tested on videos containing the disrupting events mentioned in Section 4.2. HCCS is first tested on three different sequences corresponding to different data fusion challenges. In each case, HCCS is compared to $\textcircled{\cap}$ and $\textcircled{\cup}$ using the measure r. For the sake of clarity, comparisons to other rules are gathered in Subsection 5.4, where more general conclusions on the experiments are also proposed.

5.2.1. Impact of imprecise sources

This experiment focuses on weak source combination therefore the unreliability test is disabled. The experiment is carried out on a sequence named "tennis ball". Its characteristics are summarized in Table 5.

[4] The dependency on indexes j, i and t is omitted to simplify notations.

Table 5: Tennis ball sequence characteristics.

sources used	conflict	disrupting event
S_{c-t}	$m_{\{\omega_2\}}[S_{c-t}]$ is not used, no conflict is generated	a brutal illumination change occurs, turning the colour-texture source from "normal" to "weak"
S_s	no conflict	
S_m	no conflict	the source varies from "normal" to "weak" as the target moves or not.

This experiment is a typical case for which one-source approaches relying on colour information fail[5]. With the help of other types of sources, the tracking can be maintained provided that the fusion technique is not sensitive to imprecision. In Figure 4, tracking results obtained thanks to HCCS are compared with those of a PF relying only on S_{c-t}.

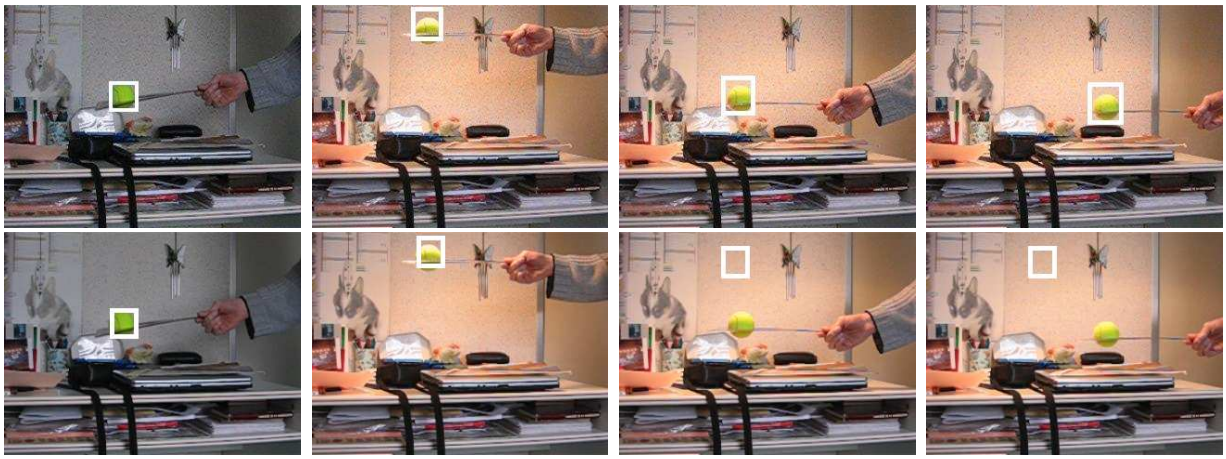

Figure 4: "tennis ball" sequence. Top: successful tracking in presence of an illumination change with HCCS. Bottom: tracking failure with a PF relying on S_{c-t}.

These bba analysis results are consistent with the scene description. An illumination change occurs around image 52. Since no source is identified as unreliable in this example, HCCS is equivalent to the cautious rule.
In Figure 5, the tracking rate evolution is represented for $\textcircled{\cup}$, $\textcircled{\cap}$ and HCCS. Unlike the disjunctive rule, HCCS is clearly insensitive to imprecise sources. Using the conjunctive behaviour of the cautious rule, HCCS exploits informative pieces of evidence. As the HCCS tracking rate shows, the tracking is maintained throughout the 149-frame sequence. There is no background model defined on that sequence, so no conflict is generated. Consequently, most of the other rules give very similar results compared to $\textcircled{\cap}$, see Subsection 5.4.

5.2.2. Impact of unreliable sources

As this experiment focuses on unreliable source combination, the weakness test is disabled. The experiment was carried out on a sequence named "two cars". Its characteristics are summarized in Table 6. The target is a grey car. The presence of a white car generates a clutter situation and consequently unreliability. When the discriminative powers of the sources are damaged by an outlier, data fusion helps to accumulate relevant information from each source so as to make a safer decision. Successful tracking on this sequence using HCCS is presented in Figure 6. Figure 7 presents HCCS's bba analysis step results. After frame 30,

[5]Illumination invariant colour features cannot usually overcome massive and sudden illumination changes.

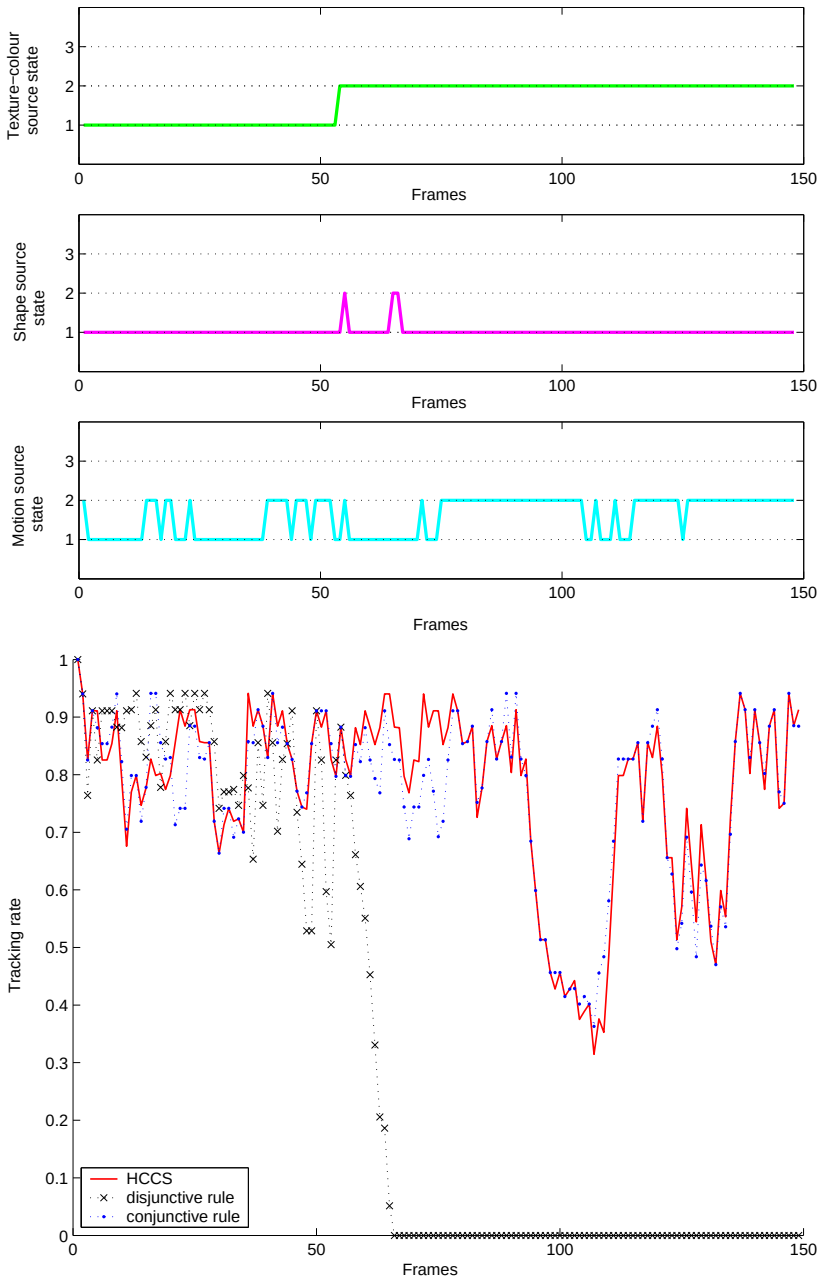

Figure 5: Tennis ball sequence. Top: bba analyses **1**: "normal", **2**: "weak", **3**: "unreliable". Bottom: tracking rates several combination techniques.

Table 6: Two cars sequence characteristics.

sources used	conflict	disrupting event
S_{c-t}	$m_{\{\omega_2\}}[S_{c-t}]$ is not used, no conflict is generated	a clutter situation occurs: two objects (a grey car and a white car) have similar colour-texture properties depending on the sun reflection on their bodies
S_s	no conflict	a clutter situation occurs: two objects (a grey car and a white car) have similar shape properties

Figure 6: "two cars" sequence: successful tracking in presence of unreliable sources.

Table 7: Dog and ball sequence characteristics.

sources used	conflict	disrupting event
S_{c-t}	$m_{\{\omega_2\}}[S_{c-t}]$ is used, conflict is generated	target occlusion: another object hides the target (the source becomes weak) noise: the colour-texture model efficiency is damaged, (the source becomes unreliable)
S_s	no conflict	target rotational motion: the symmetry features are useless
S_m	no conflict	target occlusion both the dog and the ball are moving causing unreliability
S_c	no conflict	target occlusion noise: the colour model efficiency is damaged

the white car is shadowed therefore the frequency of detection as unreliable increases for S_{c-t}. After frame 67, the white car leaves the scene and the clutter stops. S_s is less sensitive to the clutter situation which means that the shape model of the grey car is more relevant than the colour-texture one.
In Figure 7, the tracking rate evolution is represented for $\textcircled{\cap}$, $\textcircled{\cup}$ and HCCS. During the clutter, tracking algorithms using $\textcircled{\cap}$ or $\textcircled{\cup}$ "hesitate" between the two cars. This accounts for their rates dropping under 0.5. $\textcircled{\cap}$ does not produce satisfactory results because if a single source gives credit to ω_1 on two distinct regions of the image, a conjunctive fusion will most likely maintain this credit for each of these regions. Under such circumstances, HCCS selects the disjunctive rule and outperforms the conjunctive rule. Nonetheless, the disjunctive rule may favour one of the two image regions even if the sources are very imprecise. Fortunately, on this experiment, sources are not imprecise and unreliable at the same time. To ensure a safer fusion process, the weakness test must be used as well.

5.3. Dealing with multiple failures

This experiment contains unreliable, weak as well as conflicting sources. It covers a wide range of data fusion issues, thereby helping to validate HCCS in a broader context. The experiment is carried out on a sequence named "dog and ball". Its characteristics are summarized in Table 7.

The target is a ball. A dog playing with the ball causes an occlusion and blinds all sources. Figure 8 presents HCCS's results on this video. HCCS's bba analysis step results are given in Figure 9. The bba analysis results are consistent with the explanations given in Table 7. During the occlusion, S_{c-t} remains "normal" because it also gives credit to $\{\omega_2\}$ on image regions corresponding to the background (the lawn). S_m is also active during the occlusion because the dog keeps moving.
The tracking rates are presented in Figure 9. The tracking rate of $\textcircled{\cup}$ is damaged by S_s in the same way as S_{c-t} causes damage in the "tennis ball" sequence. Compared to $\textcircled{\cap}$, HCCS takes better advantage of unreliable sources in the same way as in the "two cars" sequence. Although S_{c-t} produces some conflict mass, rules relying on conflict reallocation are not adapted to the present experiment, see Table 9. During the major occlusion period, tracking rates are forced to 0 because, if the object cannot be seen, the meaning

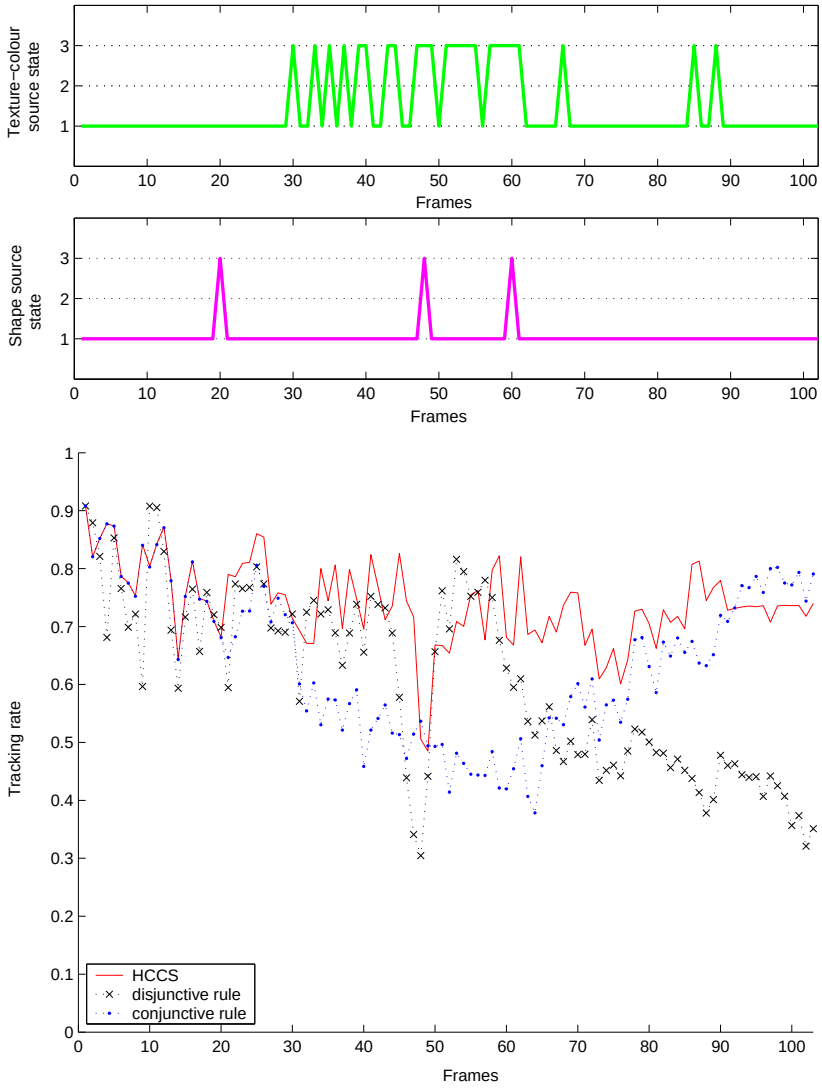

Figure 7: Two cars sequence. Top: bba analyses **1**: "normal", **2**: "weak", **3**: "unreliable". Bottom: tracking rates several combination techniques.

Figure 8: "dog and ball" sequence: successful tracking in the presence of an occlusion.

of a positive rate is questionable. As can be seen on Figure 9, the EPF recovers from the occlusion a lot faster thanks to HCCS.

5.4. Comparison between HCCS and single combination rules

In this sub-section, rules' and HCCS's performances are compared through the three sequences presented in the previous paragraphs using several measures: two quantitative measures μ and σ corresponding to the average tracking rate and its standard deviation, and a qualitative measure MTF/ST. MTF means Major Tracking Failure and ST means Satisfactory Tracking. The MFT/ST measure is evaluated by an expert.

5.4.1. Protocol

To compare HCCS with other approaches on an equal footing, the same amount of information must be available for all of them. Indeed, HCCS can outperform combination rules thanks to the contextual information that allows bba analysis. Consequently, analysis results were integrated to classical rules too. Each rule is thus tested in three different situations:

- the rule is applied without the help of bba analysis results.
- the rule is applied on weak sources and on discounted unreliable sources. A relevant value for the discounting coefficient α must be determined. If α is too small, the impact of unreliable sources is nearly unchanged. If α is too high, the pieces of evidence brought by the source are suppressed. We chose $\alpha = 0.7$ which is a good compromise in our experiments.
- the rule is applied on discounted unreliable sources and weak sources are evicted. As explained in the previous sections, weak sources damage the fusion result of several rules. Unlike unreliable sources, they can be discarded as they carry little information.

In addition, some rules require a parameter tuning; test values are:

- for parameter ζ of $\textcircled{\vee}$: $\zeta \in \left\{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\right\}$. ζ gives a minimal mass to $\emptyset$. Given the implemented bba model, a source never produces a dogmatic bba, therefore ζ is never used for $\textcircled{\wedge}$.

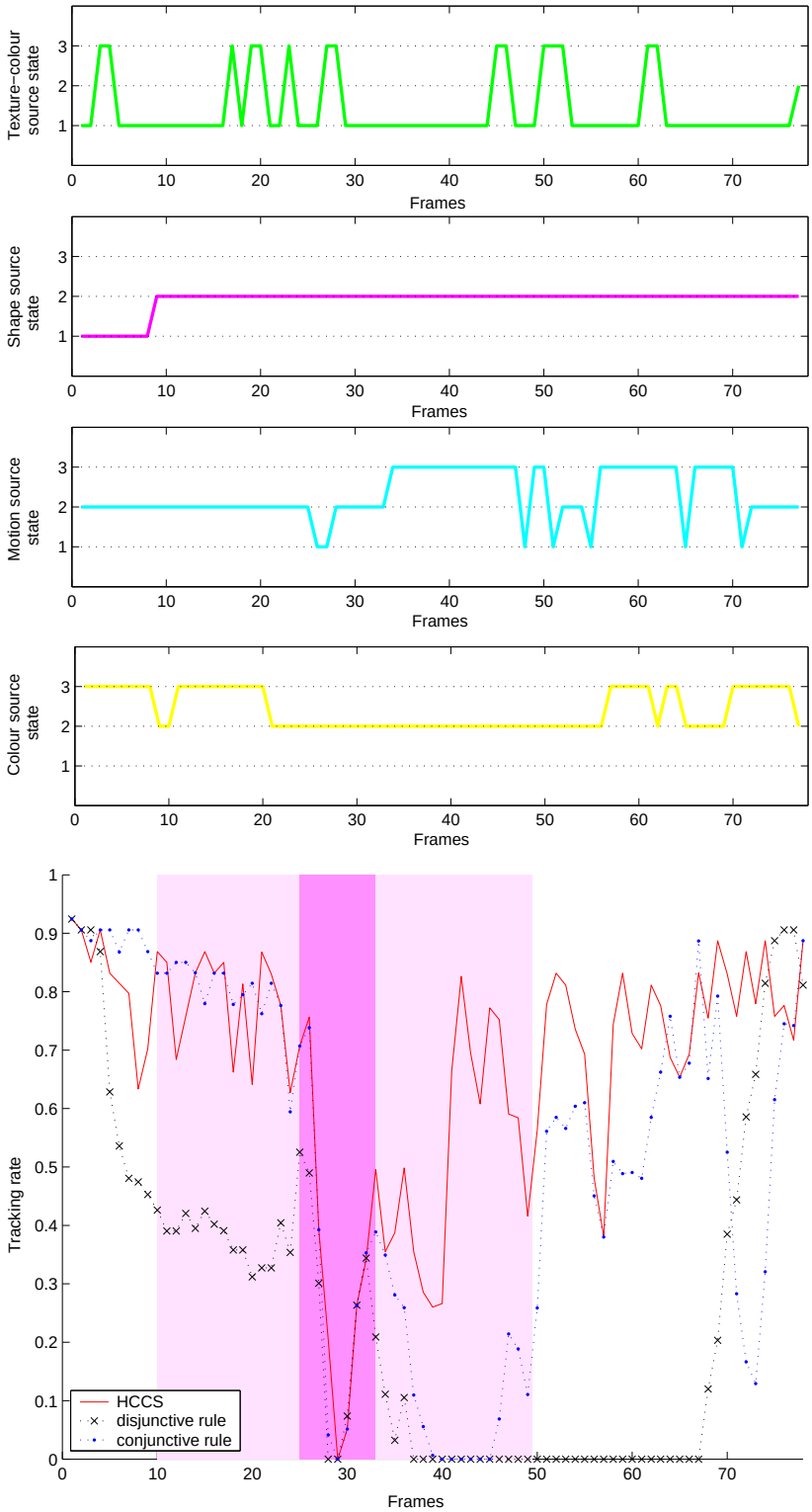

Figure 9: Dog and ball sequence. Top: bba analyses **1**: "normal", **2**: "weak", **3**: "unreliable". Bottom: tracking rates several combination techniques. Partial occlusion periods in pink, major occlusion period in magenta.

Table 8: Parameters of several rules giving highest value of μ. $\times$: test not relevant.

sequence name	Tennis ball	Two cars	Dog and ball	sequence name	Tennis ball	Two cars	Dog and ball
$\textcircled{\vee}$	$\zeta = 0.1$	$\zeta = 0.1$	$\zeta = 0.1$	DAR	$\times$	$R = 0.1$	$R = 0.9$
$\textcircled{\vee}+$	$\times$	$\zeta = 0.1$	$\zeta = 0.1$	DAR+	$\times$	$R = 0.1$	$R = 0.8$
$\textcircled{\vee}+^*$	$\zeta = 0.1$	$\times$	$\zeta = 0.1$	DAR+*	$\times$	$\times$	$R = 0.4$
DPCR	$\times$	$\times$	$\epsilon = 0.7$	DMR	$\times$	$R = 0.2$	$R = 0.4$
DPCR+	$\times$	$\times$	$\epsilon = 0.8$	DMR+	$\times$	$R = 0.1$	$R = 0.2$
DPCR+*	$\times$	$\times$	$\epsilon = 0.8$	DMR+*	$\times$	$\times$	$R = 0.9$
MDPCR	$\times$	$\times$	$\epsilon = 0.6$	HCCS	$t_{weak} = 0.0001$	$t_{weak} = \times$	$t_{weak} = 0.05$
MDPCR+	$\times$	$\times$	$\epsilon = 0.8$		$a = \times$	$a = 7$	$a = 2$
MDPCR+*	$\times$	$\times$	$\epsilon = 0.4$				

- for parameter ϵ of DPCR and MDPCR: $\epsilon \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. ϵ redirects the conflict mass on different terms of the rule's equation.
- for reliability coefficients R_j of DAR and DMR: $R_j \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$ if source S_j is unreliable and $R_j = 1$ otherwise.

Among all these parameter values, the ones giving the highest values of μ were retained and are presented in Table 8. Note that there is no unreliability test in the first experiment and no weakness test in the second one, so some cells of the table are marked with the symbol $\times$.

5.4.2. Results and discussion

Table 9 summarizes the performances of HCCS and the various rules investigated in the VT application.

Several remarks can be made on examination of Tables 8 and 9:

- The nature of the combination is the only combination rule property clearly producing different results depending on the situation. As HCCS analyses these situations for each bba and selects the adequate rule nature, it is therefore the only combination technique achieving a satisfying tracking on the three test sequences.
- Conflict redistribution rules are not adapted to the type of unreliability arising in the experiments. As demonstrated by several authors, conflict management can be of major importance in many applications [11]. If so, $\textcircled{\wedge}$ can be replaced with a conflict redistribution rule in HCCS.
- Weak source eviction improves the performances of $\textcircled{\cup}$, proving that $\Omega \succ \textcircled{\cup}$ must be taken into account inside the fusion process.
- In the second experiment, which focuses on unreliability, conditional discounting slightly improves the performances of most rules, but is not sufficient to prevent a partial loss of tracking.
- the bba models used in the experiments are inadequate for the bold rule, whose best performance is always met for $\zeta = 0.1$.

Note that, if parameters a and t_{weak} are not set correctly, the bba analysis results are erroneous and the VT performances of HCCS are then no better than those of single rules. This proves that the contextual information brought by the bba analysis is meaningful and adequately exploited by HCCS.
Furthermore, it is important to stress that these results depend on the bba models and the tracking problem definition. Better bba models [4, 5, 7] or other problem representations [22] should significantly improve the VT performances.

6. Conclusion

In this article, we have presented a novel evidential fusion scheme adapted to visual tracking challenges. Visual tracking induces specific data fusion issues regarding notably highly imprecise or unreliable sources.

Table 9: Global performances for all implemented rules and HCCS. The most remarkable results for each rule are in bold font. +: conditional discounting was used. *: weak sources evicted. ×: test not relevant. ST=satisfactory tracking, MTF=major tracking failure.

sequence name	Tennis ball			Two cars			Dog and ball		
sequence characteristics	Imprecision (illumination change, periodic motion)			Unreliability (clutter)			Imprecision and unreliability (occlusion, inaccurate object models)		
sources	S_{c-t}, S_s and S_m			S_{c-t} and S_s			S_{c-t}, S_c, S_s and S_m		
	μ	σ	MTF/ST	μ	σ	MTF/ST	μ	σ	MTF/ST
$\oplus$	**0.7673**	0.1471	ST	0.6392	0.1327	MTF	**0.5217**	0.3124	MTF
$\oplus$+	×	×	×	**0.6409**	0.1318	MTF	0.5164	0.3120	MTF
$\oplus$+*	0.6773	0.2649	ST	×	×	×	0.3568	0.3358	MTF
$\cap$	**0.7673**	0.1471	ST	0.6392	0.1327	MTF	**0.5217**	0.3124	MTF
$\cap$+	×	×	×	**0.6409**	0.1318	MTF	0.5152	0.3122	MTF
$\cap$+*	0.6773	0.2649	ST	×	×	×	0.3568	0.3358	MTF
$\cup$	0.3376	0.4060	MTF	0.6052	0.1554	MTF	0.2697	0.2991	MTF
$\cup$+	×	×	×	**0.6052**	0.1554	MTF	0.2918	0.3043	MTF
$\cup$+*	**0.6501**	0.3248	ST	×	×	×	**0.4848**	0.3296	MTF
$\wedge$	**0.7756**	0.1546	ST	0.6390	0.1348	MTF	**0.5087**	0.3168	MTF
$\wedge$+	×	×	×	**0.6398**	0.1327	MTF	0.5053	0.3163	MTF
$\wedge$+*	0.6773	0.3006	ST	×	×	×	0.3568	0.3358	MTF
$\vee$	0.7570	0.1438	ST	0.5212	0.1722	MTF	0.4677	0.3177	MTF
$\vee$+	×	×	×	**0.5864**	0.1463	MTF	**0.4942**	0.3154	MTF
$\vee$+*	**0.7669**	0.1473	ST	×	×	×	0.4101	0.3376	MTF
ZR	**0.7685**	0.1506	ST	0.6422	0.1300	MTF	**0.5356**	0.3164	MTF
ZR+	×	×	×	**0.6545**	0.1195	MTF	0.5306	0.3138	MTF
ZR+*	0.5726	0.3645	ST	×	×	×	0.4188	0.3417	MTF
YR	**0.7673**	0.1471	ST	0.6392	0.1327	MTF	**0.5209**	0.3136	MTF
YR+	×	×	×	**0.6409**	0.1318	MTF	0.5164	0.3120	MTF
YR+*	0.6773	0.2649	ST	×	×	×	0.3568	0.3358	MTF
DPR	**0.7673**	0.1471	ST	0.6392	0.1327	MTF	**0.5216**	0.3131	MTF
DPR+	×	×	×	**0.6409**	0.1318	MTF	0.5135	0.3127	MTF
DPR+*	0.6773	0.2649	ST	×	×	×	0.3564	0.3354	MTF
RCR	**0.7673**	0.1471	ST	0.6392	0.1327	MTF	**0.5205**	0.3134	MTF
RCR+	×	×	×	**0.6409**	0.1318	MTF	0.5130	0.3128	MTF
RCR+*	0.6773	0.2649	ST	×	×	×	0.4059	0.3353	MTF
MMR	**0.7779**	0.1483	ST	0.6424	0.1303	MTF	**0.5153**	0.3106	MTF
MMR+	×	×	×	**0.6529**	0.1225	MTF	0.5126	0.3109	MTF
MMR+*	0.7771	0.1491	ST	×	×	×	0.4035	0.3341	MTF
PCR6	**0.7673**	0.1471	ST	0.6392	0.1327	MTF	**0.5204**	0.3138	MTF
PCR6+	×	×	×	**0.6409**	0.1318	MTF	0.5164	0.3120	MTF
PCR6+*	0.6773	0.2649	ST	×	×	×	0.3566	0.3357	MTF
DPCR	**0.7673**	0.1471	ST	0.6392	0.1327	MTF	**0.5215**	0.3132	MTF
DPCR+	×	×	×	**0.6409**	0.1318	MTF	0.5150	0.3138	MTF
DPCR+*	0.6773	0.2649	ST	×	×	×	0.3568	0.3358	MTF
MDPCR	**0.7779**	0.1483	ST	0.6424	0.1303	MTF	**0.5160**	0.3108	MTF
MDPCR+	×	×	×	**0.6529**	0.1225	MTF	0.5129	0.3109	MTF
MDPCR+*	**0.7771**	0.1491	ST	×	×	×	4064	0.3326	MTF
DAR	0.7075	0.2347	ST	0.6742	0.1023	MTF	**0.5203**	0.3133	MTF
DAR+	×	×	×	**0.7116**	0.0767	ST	0.5140	3138	MTF
DAR+*	**0.7758**⋆	0.1567	ST	×	×	×	0.4223	0.3371	MTF
DMR	0.3376	0.4060	MTF	**0.6286**	0.1417	MTF	**0.5216**	0.3125	MTF
DMR+	×	×	×	0.6235	0.1343	MTF	0.5168	0.3125	MTF
DMR+*	**0.7976**	0.1499	ST	×	×	×	0.3708	0.3422	MTF
HCCS	**0.7756**	0.1546	ST	**0.7388**	0.0713	ST	**0.6705**	0.2122	ST

The constraints imposed by these abnormal sources made it necessary to design a broader fusion technique named the hierarchical and conditional combination scheme (HCCS). We propose to analyse bbas so as to detect and identify unreliable or highly imprecise sources. The fusion problem can then be separated in sub-problems, thereby reducing the number of constraints. We justify that two groups of bba can be processed by the cautious and disjunctive rules respectively. The two output bbas are analysed as well and aggregated using one of these two rules depending on the analysis results.
The bba analysis step is performed using contextual information brought by an evidential particle filter. This filter is used as the tracking algorithm for our visual tracking application. The experiments show that HCCS produces satisfactory visual tracking performances in spite of the presence of unreliable or highly imprecise sources. As compared to single combination rules, HCCS responds adequately to all the tracking scenarios examined. HCCS is a novel tool in the sense that it adapts itself individually to each source whenever a new image arrives. In addition, HCCS computation cost is equivalent to that of a combination rule.
The flexibility and robustness of combination schemes open new perspectives for other information fusion applications. Indeed, HCCS can be extended to other types of bbas provided that some contextual information allows an *ad hoc* bba analysis step. In future works, it is also intended to use HCCS in a multiple object tracking context. The frame of discernment and bba construction presented in [22] can be used to achieve multiple object tracking. It comprises a data association process between previous tracks and newly obtained observations. HCCS can be used directly on each track for source combination. It may also be possible to select a set of relevant particles for each object so that the bba analysis step can be run more locally.

A. EPF used in our experiments

In this appendix, some details concerning the EPF used in our tests are presented. This EPF relies on a sampling-importance-resampling with a multinomial resampling. Its evidential part is the one proposed in [17]. The procedure is summarized in algorithm 1. $\lambda_t^{(i)}$ is the weight of particle $X_t^{(i)}$. Y_t is the random variable representating observations. $\sigma_1, ..., \sigma_4$ are the standard deviation of the sampling density. These parameters as well as the initial position of the object X_0 are kwown *a priori*. The sampling density is sub-optimal but allows a simplification of the particle weight update step.

Algorithm 1 EPF used in our experiments

EPF used in our experiments

for t=1 to end of sequence **do**
  for i=1 to N **do**
    Sample particles $X_t^{(i)}$ using $p\left(X_t^{(i)}|X_{t-1}^{(i)}\right) = \left(\mathcal{N}\left(0,\sigma_1\right),\mathcal{N}\left(0,\sigma_2\right),\mathcal{N}\left(0,\sigma_3\right),\mathcal{N}\left(0,\sigma_4\right)\right)$
    for j=1 to M **do**
      Evaluate $m\left[S_j\right]$
    end for
    Use a DST combination method to aggregate the sources and calculate the pignistic transform $BetP$
    Obtain the likelihood $p\left(Y_t|X_t^{(i)}\right) = BetP\left(\omega_1\right)$
    Update weights using $\lambda_t^{(i)} \propto \lambda_{t-1}^{(i)} p\left(Y_t|X_t^{(i)}\right)$
    Normalize weights $\tilde{\lambda}_t^{(i)} = \frac{\lambda_t^{(i)}}{\sum_{j=1}^{N}\lambda_t^{(i)}}$
  end for
  Estimate the filtering density $p\left(X_t|Y_{1:t}\right) = \sum_{i=1}^{N}\tilde{\lambda}_t^{(i)}\delta_{X_t^{(i)}}\left(X_t\right)$
  Re-sample N new particles $\tilde{X}_t^{(j)}$ among $X_t^{(i)}$ with probability $\lambda_t^{(i)}$ and assign them the weights $\lambda_t^{(j)} = \frac{1}{N}$
end for
End

Note that HCCS can be used with any EPF. The one described in this appendix was chosen for its simplicity, therefore the influence of a combination method on the VT performances is easier to interpret.

References

[1] A. Bensrhair, M. Bertozzi, A. Broggi, A. Fascioli, S. Mousset, and G. Toulminet. Stereo-vision based feature extraction for vehicle detection. In *IEEE Intelligent Vehicles Symposium (IV'02)*, pages 465–470, 2002.

[2] F. Caron, M. Davy, E. Duflos, and P. Vanheeghe. Particle filtering for multisensor data fusion with switching observation models: applications to land vehicle positioning. *IEEE Trans. on Signal Processing*, 55(6):2703–2719, june 2007.

[3] F. Delmotte, L. Dubois, A.-M. Desodt, and P. Borne. Using trust in uncertainty theories. *Information and Systems Engineering*, 1:303–314, 1995.

[4] T. Denoeux. A k-nearest neighbour classification rule based on Dempster-Shafer theory. *IEEE trans. on Systems Man and Cybernetics*, 25(5):804–813, 1995.

[5] T. Denoeux. A neural network classifier based on Dempster-Shafer theory. *IEEE trans. on Systems, Man and Cybernetics A*, 30:131–150, 2000.

[6] T. Denoeux. Conjunctive and disjunctive combination of belief functions induced by non-distinct bodies of evidence. *Artificial Intelligence*, 172(2-3):234–264, february 2008.

[7] T. Denoeux and P. Smets. Classification using belief functions: the relationship between the case-based and model-based approaches. *IEEE Trans. on Systems, Man and Cybernetics B*, 36(6):1395–1406, 2006.

[8] D. Dubois and H. Prade. A set-theoretic view of belief functions: logical operations and approximatons by fuzzy sets. *Int. Journal of General Systems*, 12(3):193–226, 1986.

[9] D. Dubois and H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Comput. Intell.*, 4:244–264, 1988.

[10] F. Faux and F. Luthon. Robust face tracking using color Dempster-Shafer fusion and particle filter. In *Int. Conf. on Information Fusion (FUSION'06)*, pages 1–7, 2006.

[11] M. C. Florea, A-L Jousselme, E. Boiss, and D. Grenier. Robust combination rules for evidence theory. *Information Fusion*, 10:183–197, 2009.

[12] M. Ha-Duong. Hierarchical fusion of expert opinion in the transferable belief model, application to climate sensitivity. *Int. Journal of Approximate Reasoning*, 49:555–574, 2008.

[13] T. Inagaki. Interdependence between safety-control policy and multiple-sensor schemes via Dempster-Shafer theory. *IEEE trans. on Reliability*, 40(2):182–188, 1991.

[14] M. Isard and A. Blake. Condensation-conditional density propagation for visual tracking. *Int. Journal of Computer Vision*, 29(1):5–28, 1998.

[15] A. Kallel and S. Le Hégarat-Mascle. Combination of partially non-distinct beliefs: the cautious-adaptative rule. *Int. Journal of Approximate Reasoning*, 50(7):1000–1021, 2009.

[16] J. Klein, C. Lecomte, and P. Miché. Fast color-texture discrimination: application to car-tracking. In *IEEE Int. Conf. on Intelligent Transportation Systems (ITSC'07)*, pages 541–546, 2007.

[17] J. Klein, C. Lecomte, and P. Miché. Preceding car tracking using belief functions and a particle filter. In *IEEE Int. Conf. on Pattern Recognition (ICPR'08)*, pages 864–871, Tampa (USA), december 2008.

[18] G. Lefaix, E. Marchand, and P. Bouthemy. Motion-based obstacle detection and tracking for car driving assistance. In *IEEE Int. Conf. on Pattern Recognition (ICPR'02)*, pages 74–77, 2002.

[19] E. Lefevre, O. Colot, and P. Vannoorenberghe. Belief function combination and conflict management. *Information Fusion*, 3:149–162, 2002.

[20] R. Mahler. Can the Bayesian and Dempster-Shafer approaches be reconciled? yes. In *Int. Conf. on Information Fusion (FUSION'05)*, pages 864–871, 2005.

[21] A. Martin and C. Osswald. Toward a combination rule to deal with partial conflict and specificity in belief functions theory. In *Int. Conf. on Information Fusion (FUSION'07)*, pages 1–8, Quebec (Canada), 9-12 July 2007.

[22] N. Megherbi, S. Ambellouis, O. Colot, and F. Cabestaing. Multimodal data association based on the use of belief functions for multiple target tracking. In *Int. Conf. on Information Fusion (FUSION'05)*, pages cd–rom, Philadelphia , PA (USA), july 2005.

[23] D. Mercier, B. Quost, and T. Denoeux. Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 9(2):246–258, april 2008.

[24] R. Muñoz-Salinas, R. Medina-Carnicer, F.J. Madrid-Cuevas, and A. Carmona-Poyato. Multi-camera people tracking using evidential filters. *Int. Journal of Approximate Reasoning*, 50:732–749, 2009.

[25] C.K. Murphy. Combining belief functions with evidence conflicts. *Decision Support Systems*, 29:1–9, 2000.

[26] P. Perez, C. Hue, J. Vermaak, and M. Gangnet. Color-based probabilistic tracking. In *IEEE European Conf. on Computer Vision (ECCV'02)*, pages 661–675, 2002.

[27] P. Perez, J. Vermaak, and A. Blake. Data fusion for visual tracking with particles. *Proceedings of the IEEE*, 92(3):495–513, 2004.

[28] F. Pichon and T. Denoeux. Interpretation and computation of α-junctions for combining belief functions. In *6th Int. Symposium on Imprecise Probability: Theories and Applications (ISIPTA '09)*, Durham, U.K., 2009.

[29] B. Quost, T. Denoeux, and M.-H. Masson. Adapting a combination rule to non-independent information. In *12th Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU'08)*, pages 448–455, 2008.

[30] B. Quost, M.-H. Masson, and T. Denoeux. Refined classifier combination using belief functions. In *Int. Conf. on Information Fusion (FUSION'08)*, pages 776–782, 2008.
[31] A. Jøsang, M. Daniel, and P. Vannoorenberghe. Strategies for combining conflicting dogmatic beliefs. In *Int. Conf. on Information Fusion (FUSION'03)*, pages 1133–1140, Cairns (Australia), 2003.
[32] K. Sentz and S. Ferson. Combination of evidence in Dempster-Shafer theory. Technical report, SANDIA tech. report, 2002.
[33] G. Shafer. *A Mathematical Theory of Evidence*. Princeton Univ.press, 1976.
[34] F. Smarandache. An in-depth look at information fusion rules and the unification of fusion theories. Univeristy of New Mexico, 2004.
[35] F. Smarandache and J. Dezert. Information fusion based on new proportional conflict redistribution rules. In *Int. Conf. on Information Fusion (FUSION'05)*, pages 1–8, Philadelphia (USA), 25-29 july 2005.
[36] F. Smarandache and J. Dezert. *Advances and Applications of DSmT for Information Fusion (Collected works), 2nd volume*. Am. Res. Press, 2006.
[37] P. Smets. The canonical decomposition of weighted belief. In *Int. Joint. Conf. on Artificial Intelligence*, pages 1896–1901, 1995.
[38] P. Smets. The alpha-junctions: Combination operators applicable to belief functions. In *First Int. Joint Conference on Qualitative and Quantitative Practical Reasoning*, volume 1244, pages 131–153. LNCS, 1997.
[39] P. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion*, 8:387–412, 2006.
[40] P. Smets and B. Ristic. Kalman filter and joint tracking and classification based on belief functions in the TBM framework. *Information Fusion*, 8:16–27, 2007.
[41] Y. Sun and L. Bentadet. A sequential Monte Carlo and DSmT based approach for conflict handling in case of multiple targets tracking. *Lecture Notes in Computer Science*, 4633:526–537, 2007.
[42] R. Yager. On the Dempster-Shafer framework and new combination rules. *Information Sciences*, 41:93–138, 1987.
[43] R. Yager. Quasi-associative operations in the combination of evidence. *Kybernetes*, 16:37–41, 1987.
[44] L. Zhang. *Advances in the Dempster-Shafer theory of evidence*, chapter Representation, Independence, and combination of evidence in the Dempster-Shafer theory, pages 51–69. John Wiley & Sons, 1994.