Two Decision Fusion Frameworks for Image Forensics

M. Fontani, A. Costanzo, M. Barni
University of Siena
Dept. of Information Engineering
Via Roma 56, 53100, Siena - ITALY

T. Bianchi, A. De Rosa, A. Piva
University of Florence
Dept. of Electronics and Telecommunications
Via S. Marta 3, 50139, Firenze - ITALY

***Abstract*—Image forensics research has mainly focused on the detection of artifacts introduced by a single processing tool, thus resulting in the development of a large number of specialized detection algorithms. In tamper detection applications, however, the kind of artifacts the forensic analyst should look for is not known beforehand, hence making it necessary that several tools developed for different scenarios are applied. The problem, then, is to devise a sound strategy to fuse the information provided by the different tools. In this paper we introduce two theoretical frameworks, based on Dempster-Shafer's Theory of Evidence and on Fuzzy Theory respectively, to perform the fusion of heterogeneous, incomplete or conflicting outputs of forensic algorithms. Both models are easily expandable to an arbitrary number of tools, do not require tools output to be probabilistic and take into account available information about tools reliability. To validate the proposed approaches, we carried out some experiments addressing a simple yet realistic scenario in which three forensic tools exploit different artifacts introduced by double JPEG compression to detect cut&paste tampering within a specified region of an image. The results we obtained are encouraging, especially when compared with the performance of a simple decision method based on the binary OR operator.**

SESSION: SIGNAL PROCESSING

I. INTRODUCTION

Nowadays the majority of images are created, stored and distributed in a digital format that is fairly easy to edit and tamper with. As a result, digital image forensics has become an important field of research to prove the authenticity and integrity of digital images. In the last years many techniques for detecting photographic tampering have been proposed [1], [2], [3], [4]. Each tool implementing a detection technique usually searches for a trace left by a specific processing, thus dealing with a single type of manipulation. However when an analyst is asked to judge the integrity of a given image, the kind of manipulation the image has undergone is not known beforehand. Therefore, if we are interested in finding whether an image has been tampered with, the application of a single detection method may not be enough and it is necessary to use more than one method. However, when using more than one tool, we are usually interested in obtaining a single global answer that allows us to decide whether the image is authentic or not. In other words, we need to fuse different outputs. Several critical problems may arise in this scenario: for example outputs can be heterogeneous, discording or unreliable. This problem can be addressed in different ways [5], the most important of them being fusion at "feature level" (that is, fusion of data provided by tools is performed before taking a final classification) and fusion at "measurement level" (each tool provides a score of detection, and these scores are fused). Most of the existing works are based on the former method; hybrid approaches have been developed in [6] and in [7]. In this work we propose two different frameworks to tackle the fusion problem at measurement level, based on Dempster-Shafer's Theory of Evidence and on Fuzzy Theory respectively. Both theories have been conceived to overcome some limitations of classic theory of probability when dealing with incomplete or unreliable information. For this reason we believe that decision fusion methods based on such theories may help to deal with the heterogeneous or conflicting outputs usually provided by different forensic algorithms.

Dempster-Shafer's (DS) theory of evidence [8] is a framework for reasoning under uncertainty that allows the representation of ignorance and of available information in a more flexible way with respect to Bayesian theory. Reasoning in the Bayesian framework often urges to apply insufficient reasoning to assign a-priori probabilities, thus introducing extraneous assumptions. Dempster-Shafer's theory [8], instead, abandons the classical probability frame and allows to reason without a-priori probabilities through a new formalism.

Fuzzy sets theory was conceived in 1965 by L. Zadeh as an extension of the classic set theory [9], [10]. From this initial concept a multi-value fuzzy logic has been derived as an extension of Boolean logic. Fuzzy logic aims to imitate the highly adaptive behavior of human reasoning to incomplete, unreliable or partially true information. Fuzzy logic has been used in many control applications in which robustness to noise and imprecise inputs is a critical requirement [11], [12].

The rest of this paper is organized as follows: in section II we provide a formalization of the decision fusion problem from an image forensics perspective; in sections III and IV we introduce two different frameworks to address the problem; in section V we test the two proposed models by fusing outputs coming from three tools based on the analysis of JPEG compression artifacts and we compare their performances with those of a decision method based on binary OR; finally, in section VI we outline some directions for future research.

II. Problem formalization

Let $\mathcal{T}$ be a set of K image forensic tools for detecting whether a certain region within an image I has been tampered with or not. Each tool $T_i \in \mathcal{T}$ analyzes a feature set in the specified region looking for tampering traces and generates an output that tells us whether the trace is present or not. At the end of this process we have K outputs. If we want to answer to the question "*has the selected region been tampered with?*" we need a method to merge these K outputs into a single value. Based on this value we can take a final decision on the authenticity of the region.

There are several approaches to perform this task. Amongst the simplest we can include majority decision and logical disjunction. In the first case a region is considered to be tampered with if more than half of the tools tells that a tampering has occurred; in the second case, a region is considered to be tampered with if at least one tool says that a tampering has occurred. As the number of adopted tools increases several problems may arise, thus making classic decision methods ineffective. Let us consider some examples. Two or more tools could be mutually exclusive: if one finds traces of tampering then the other(s) will not find anything. In this case a decision method based on majority may not work as intended. Moreover, tools are usually not perfect. Practical implementation of a forensic algorithm can be a delicate process: from tuning of parameters to choice of training dataset, many factors can affect the final performance. This may result in a tool that is prone to errors. Let us imagine a case where we have $K-1$ tools that work perfectly and one that is really bad. It may happen that most of times this tool will claim that the image region has been tampered with, thus inducing a simple logic disjunction operator to error. For these reasons we need to devise an alternative reliable fusion method. In the sequel we propose two different decision fusion approaches based on Dempster-Shafer's Theory of Evidence and fuzzy logic respectively.

A. Tool outputs.

In order to proceed with our decision fusion approaches, we need to assume that all the tools share the same output format, consisting of a pair of values (D, R), where:

- $D \in [0,1]$ is the degree of *detection*, that is a measure of the presence of the tampering trace within the analyzed image region. Values near 1 indicate a high presence of the tampering trace.
- $R \in [0,1]$ is the *reliability* of D, that is a measure of the confidence of tool on the detection value. Values near 1 indicate a high confidence.

D does not necessarily need to be a probability and generally changes from region to region. R can either be a constant value depending only on the overall performance of the tool or change depending on the characteristics of the analyzed region (i.e. size, color, visual content). Therefore, in order to define the reliability of each tool, we need some informations about its performance, drawn either from theory or from experimental analysis.

B. Definition of tampering tables.

Now that each tool provides a standard output, we need to describe the behavior we expect from them. Let us suppose that a region of image I has undergone a tampering. We ask ourselves: "*If everything goes smoothly, what kind of output are we expecting from the tools at our disposal?*". Depending on the nature of manipulation, a tool may or may not be able to detect a region as tampered. We indicate the capability of detecting the tampering with Y and the incapability with N. Therefore, if we have K tools, each manipulation (or absence of manipulation) is identified by one or more K-dimensional sequences of Y and N, each specifying the expected behavior of the tools in *ideal* conditions. Note that there may be some sequences that are not specified a priori: some sequences may be not ideally acceptable or correspond to an unknown type of tampering.

We refer with T_{true} and T_{false} to the tables whose columns correspond to the expected (*standard*) cases of detection and non-detection of tampering respectively; we refer with T_{doubt} to the table of unexpected (*non-standard*) cases belonging neither to T_{true} nor to T_{false}. Since the definition of these tables depends on tools and is based on knowledge of their performances, in the following we assume that they are always available.

III. Dempster-Shafer (DS) decision fusion framework

Dempster-Shafer's (DS) theory of evidence [8] is often cited in the field of decision fusion because it allows to combine evidences coming from different sources, interpreting them as "belief" on propositions, and provides a formalism for turning logical operations on propositions into operations among sets. Another key feature of DS Theory is its soundness in dealing with uncertainty.

When using classical probability theory for finding the probability of a certain event A, the additivity rule must be satisfied; so by saying that $Pr(A) = p_A$ one implicitly says that $Pr(\bar{A}) = 1 - p_A$, thus committing information about the probability of event A to its complementary $\bar{A}$. Another consequence of the rule of additivity involves the representation of ignorance: complete ignorance about a dichotomic event A is usually represented by setting $Pr(A) = Pr(\bar{A}) = 0.5$ (according to Laplace's principle of insufficient reasoning), but this probability distribution would also be used to model perfect knowledge about probability of each event being 0.5 (as for a fair coin tossing). These facts have an impact in reasoning: Bayesian inference framework typically requires to specify a-priori probabilities and, in scenarios where uncertainty is high, this results in an extensive use of the principle of insufficient reasoning, thus introducing extraneous assumptions. Dempster-Shafer's theory [8], instead, abandons the classical probability frame and allows to reason without a-priori probabilities through a new formalism.

A. Dempster-Shafer's formalism

Let the frame $\Theta_x = \{x_1, x_2, \ldots, x_n\}$ define a finite set of possible mutually exclusive and exhaustive values of a variable x. We are interested in quantifying the belief for propositions of the form "the true value of x is in H", where $H \subseteq \Theta_x$ (so the set of all possible propositions is the power set of Θ_x, 2^{Θ_x}). Contrary to the Bayesian case, in this framework belief for one proposition is only committed to any other logically *implied* proposition, i.e. belief for a given subset H of Θ is only committed to any subset B $\subset$ Θ containing H. Each proposition is mapped on a single subset and is assigned a basic belief *mass* through a Basic Belief Assignment (BBA).

Definition Let Θ be a frame. A function $m : 2^{\Theta} \to [0,1]$ is called a Basic Belief Assignment if the followings hold:

$$m(\emptyset) = 0 \qquad \sum_{A \subseteq \Theta} m(A) = 1 \tag{1}$$

where the summation is taken over every possible subset A of Θ. Each set S such that $m(S) > 0$ is called a *focal element* for m.

Intuitively, $m(A)$ is the atomic information for this framework: it is the part of belief that supports exactly A but, due to the lack of further information, does not support any strict subset of A. Thus if we want to obtain the total belief that a given BBA commits to A, we must add the mass of all proper subsets of A plus the mass of A itself, thus obtaining the *Belief* for A.

Definition A function $Bel : 2^{\Theta} \to [0,1]$ is a belief function over Θ if the following is satisfied:

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

$Bel(A)$ summarizes all our reasons to believe in A. Relationships and interpretations of $m(A)$, $Bel(A)$ and other function that derives from these are well explored in [13]. Here we just notice that $Bel(A) + Bel(\bar{A}) \leq 1 \; \forall A \subseteq \Theta$ and $1 - (Bel(A) + Bel(\bar{A}))$ is the lack of information, that is "the doubt", about the proposition A.

B. Combination Rule

We are interested in using DS framework to perform data fusion. Dempster defined a *combination rule* that allows to combine several belief functions defined over the same frame.

Definition Let Bel_1 and Bel_2 be belief functions over the same frame Θ with BBAs m_1 and m_2. Let us also assume that K, defined below, is positive. Then for all non-empty $A \subseteq \Theta$ the function m_{12} defined as:

$$m_{12}(A) \triangleq \frac{1}{1-K} \cdot \sum_{\substack{i,j: \\ A_i \cap B_j = A}} m_1(A_i) m_2(B_j) \tag{2}$$

where $K = \sum_{i,j: A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j)$, is a BBA function and is called the *orthogonal sum* of Bel_1 and Bel_2, denoted by $Bel_1 \oplus Bel_2$.

This rule has many properties [13], in this work we are mainly interested in its associativity and commutativity. Notice that K is a measure of the *conflict* between m_1 and m_2: the higher it is, the higher the conflict. Dempster's rule can give unintuitive results when the conflict is near to 1: Zadeh showed this drawback in its famous paradox [14]. However, Haenni proved that Dempster rule works well in combining "realistic" BBAs in which, for example, no oracles exist [15].

In the previous definition (2) it is assumed that the two BBAs, m_1 and m_2, are defined over the same frame. Whenever we need to combine BBAs that are defined on different domains, we have to redefine them on the same target frame: this can be done using *marginalization* and *vacuous extension*. Let us call domain D the set of variables on which evidence is defined, and let denote a BBA on domain D with m^D.

Definition Let m^{D_1} be a BBA function defined on a domain D_1, then its vacuous extension to $D_1 \cup D_2$, denoted with $m^{D_1 \uparrow (D_1 \cup D_2)}$, is defined as:

$$m^{D_1 \uparrow (D_1 \cup D_2)}(C) = \begin{cases} m^{D_1}(A) & \text{if } C = A \times \Theta_{D_2}, A \subseteq \Theta_{D_1} \\ 0 & \text{otherwise} \end{cases}$$

This allow to extend the frame of a BBA function without introducing extraneous assumptions (no new information is provided about variables that are not in D_1). The inverse operation of vacuous extension is marginalization.

Definition Let m^D be a BBA function defined on a domain D, its marginalization to the domain $D_0 \subseteq D$, denoted with $m^{D \downarrow D_0}$, is defined as

$$m^{D \downarrow D_0}(A) = \sum_{B \downarrow A} m^D(B)$$

where the index of the summation denotes all sets $B \subseteq \Theta_D$ such that the configurations in B reduce to those in $A \subseteq \Theta_{D_0}$ by the elimination of variables in D that are not also in D_0.

C. Model for the Multimedia Forensics scenario

In this section we define a model in the Dempster-Shafer framework that fits the problem formalized in sec. II. Firstly we just need to map tool outputs (that are easily thinkable as propositions, as we shall see) in a set of Basic Belief Assignments on some sets; then, we will deal with the fusion problem. The model is built in such a way that the introduction of new evidence, coming from new tools, is straightforward.

1) Formalization for one tool: For clarity of explanation, we start by formalizing the proposed model for just one tool, let us call it *ToolA*, which returns a value of detection $A \in [0,1]$ and has a reliability $R \in [0,1]$. Intuitively, this tool will provide evidence for the propositions: "image has undergone a tampering detectable using *ToolA*" and for the opposite "image has not undergone a tampering detectable using *ToolA*". We model this information introducing a variable T_a, with frame: Θ_{T_a} = {*ta*, *na*}, where *ta* stands for the first proposition and *na* stands for the second. The power set of Θ_{T_a} will contain also (*ta* $\cup$ *na*): it is the doubt that *ToolA* has about the presence of the trace, so it refers to the proposition "image has or has

not undergone a tampering detectable using *ToolA*".
From the detection value A provided by *ToolA* we specify the following BBA over the frame Θ_{T_a}:

$$m_A^{T_a}(X) = \begin{cases} A_T & \text{for } X = \quad \{(\mathit{ta})\} \\ A_N & \text{for } X = \quad \{(\mathit{na})\} \\ A_{TN} & \text{for } X = \quad \{(\mathit{ta}) \cup (\mathit{na})\} \end{cases} \tag{3}$$

The way A is mapped into A_T, A_N and A_{TN} is an *interpretation* of *ToolA* response: one possible, general, interpretation is to consider the response doubtful when detection value A is near 0.5, and increasingly sure when it approaches to the interval extremes. If more information is available about how the tool detection value should be interpreted, it can be implemented by choosing an appropriate mapping from A to A_T, A_N and A_{TN} (see fig. 1 for an example of three different mappings).

We have assumed *ToolA* coming with a value of reliability R, which can optionally depend on the specific image it is working on. This information can be formalized introducing a new variable R_a, with frame: $\Theta_{R_a} = \{ra, ua\}$ where *ra* is the event "*ToolA* is reliable" and *ua* is the event "*ToolA* in not reliable". We choose to summarize reliability information using a BBA that has only two focal elements:

$$m_A^{R_a}(X) = \begin{cases} A_R & \text{for } X = \quad \{(\mathit{ra})\} \\ 1 - A_R & \text{for } X = \quad \{(\mathit{ua})\} \end{cases}$$

This BBA does not assign a mass to the doubt: we are saying that knowing that a tool is not trustable and not knowing whether it is trustable turns out to be the same. Consequently, the most intuitive mapping from R to this BBA assignment is to choose $A_R = R$.

Being defined on different frames, $m_A^{T_a}$ and $m_A^{R_a}$ cannot be combined as they are. We need to extend them to a common domain: the simplest one is $T_a \times R_a$. We use vacuous extension to find $m_A^{R_a \uparrow (T_a \times R_a)}$ while, for extending $m_A^{T_a}$, we use a different approach, to give a specific interpretation of what tool reliability should mean: we assume that if the tool is unreliable, its detection should not be considered. This can be easily expressed by putting all elements representing propositions in which the tool is not reliable (i.e. all ($\cdot$, *ua*) elements) in every focal element of the combined BBA:

$$m_A^{T_a \times R_a}(X) = \begin{cases} A_T & \text{for } X = \quad \{(\mathit{ta}, \mathit{ra}) \cup (\mathit{ta}, \mathit{ua}) \cup (\mathit{na}, \mathit{ua})\} \\ A_N & \text{for } X = \quad \{(\mathit{na}, \mathit{ra}) \cup (\mathit{ta}, \mathit{ua}) \cup (\mathit{na}, \mathit{ua})\} \\ A_{TN} & \text{for } X = \quad \{(\mathit{ta}, \mathit{ra}) \cup (\mathit{na}, \mathit{ra}) \cup (\mathit{ta}, \mathit{ua}) \cup \\ & \qquad \cup (\mathit{na}, \mathit{ua})\} \end{cases}$$

Now, using (2) we can combine reliability and detection BBAs to yield m_A, which summarizes all our knowledge about *ToolA* by now:

$$m_A(X) = \begin{cases} A_R \cdot A_T & \text{for } X = \quad \{(\mathit{ta}, \mathit{ra})\} \\ A_R \cdot A_N & \text{for } X = \quad \{(\mathit{na}, \mathit{ra})\} \\ A_R \cdot A_{TN} & \text{for } X = \quad \{(\mathit{ta}, \mathit{ra}) \cup (\mathit{na}, \mathit{ra})\} \\ 1 - A_R & \text{for } X = \quad \{(\mathit{ta}, \mathit{ua}) \cup (\mathit{na}, \mathit{ua})\} \end{cases}$$

2) Introducing new tools: If another tool, say *ToolB*, respecting the assumptions in sec. II becomes available, we can use the same formalism defined in the previous section to introduce it into the model. We will get to m_B, a BBA that summarizes the knowledge for this new tool, defined over the frame $\Theta_{T_b} \times \Theta_{R_b}$.

Because we cannot combine m_A and m_B unless they are defined over the same frame, we choose the following strategy: first marginalize both the BBAs eliminating reliability variables (we are eventually interested only in the detection value); then redefine $m_A^{T_a}$ and $m_B^{T_b}$ on the new domain $T_a \times T_b$ using vacuous extension; finally use Dempster rule to combine these two BBAs, yielding m_{AB}:

$$m_{AB}(X) = \begin{cases} A_R \cdot A_T \cdot B_R \cdot B_T & \text{for } X = \quad \{ (\mathit{ta}, \mathit{tb})\} \\ A_R \cdot A_T \cdot B_R \cdot B_N & \text{for } X = \quad \{(\mathit{ta}, \mathit{nb}) \} \\ A_R \cdot A_T \cdot C_B & \text{for } X = \quad \{(\mathit{ta}, \mathit{tb}) \cup (\mathit{ta}, \mathit{nb}) \} \\ A_R \cdot A_N \cdot B_R \cdot B_T & \text{for } X = \quad \{(\mathit{na}, \mathit{tb}) \} \\ A_R \cdot A_N \cdot B_R \cdot B_N & \text{for } X = \quad \{(\mathit{na}, \mathit{nb}) \} \\ A_R \cdot A_N \cdot C_B & \text{for } X = \quad \{(\mathit{na}, \mathit{tb}) \cup (\mathit{na}, \mathit{nb}) \} \\ C_A \cdot B_R \cdot B_T & \text{for } X = \quad \{(\mathit{ta}, \mathit{tb}) \cup (\mathit{na}, \mathit{tb}) \} \\ C_A \cdot B_R \cdot B_N & \text{for } X = \quad \{ (\mathit{ta}, \mathit{nb}) \cup (\mathit{na}, \mathit{nb})\} \\ C_A \cdot C_B & \text{for } X = \quad \{(\mathit{ta}, \mathit{tb}) \cup (\mathit{na}, \mathit{tb}) \cup \\ & \qquad \cup (\mathit{ta}, \mathit{nb}) \cup (\mathit{na}, \mathit{nb}) \} \end{cases}$$

where $C_A = (1 - A_R(A_T + A_N))$ and $C_B = (1 - B_R(B_T + B_N))$, *tb* is the proposition "image has undergone a tampering detectable using *ToolB*" and *nb* is the proposition "image has not undergone a tampering detectable using *ToolB*".

This process can be iterated: if another new tool *ToolX* become available, the associativity of Dempster's rule allows to combine directly its BBA m_X with m_{AB}, so we will always need to extend the domain of only two BBAs: the one coming from the new tool and the last one we had for previous tools. This strategy makes this model easily expandable up to a arbitrary high number of tools.

3) Tool compatibility: By now we have considered tool responses independent from each other. This allowed us to avoid conflict between tools, keeping the model easily expandable. However, in sec. II-B we introduced tables T_{true}, T_{false} and T_{doubt}, with the last one ruling out some of the combination of tool responses. If we have three tools (*ToolA*, *ToolB*, *ToolC*), information coming from these tables can be easily written using a BBA, defined on domain $T_a \times T_b \times T_c$, that has only one focal set, which contains the union of all events that are considered possible (that is, events described in one between T_{true} and T_{false}) while all other events are assigned a null mass.

This BBA should be combined, as a last step, with m_{ABC} and, having some events being declared as impossible, some conflict will arise (represented by K in eq. 2), making the data fusion non trivial. Notice that, thanks to the commutative property of Dempster's rule of combination, we could introduce informations about tool relationships only in the last step, keeping the model easily expandable.

4) Model output: At the end of the decision fusion task we want to know whether a given region of an image has been tampered with or not. We can therefore define the final output of the proposed model just looking at the belief for two sets:

the first one, T, is the union of all events in which at least one algorithm revealed a tampering; the second one, N, is the event in which none of the tools detected a tampering (in the previous example we would have $N = (na, nb, nc)$). It can be useful to keep trace of the conflict between tools that raised during the fusion process, which is given by K in eq. 2. This conflict depends only on the compatibility tables defined in sec. II-B.

Finally, the output of the proposed model is made up of two belief values and a measure of the conflict detected during decision fusion; formally, the output is given by:

$$Bel(T); \quad Bel(N); \quad K$$

where K is defined in sec. III-B.

These outputs summarize the information provided by the available tools, without forcing a final decision: if a binary decision about image authenticity is required, an interpretation of these outputs has to be made, for example a comparison between $Bel(T)$ and $Bel(N)$ as will be shown later.

IV. Fuzzy decision fusion framework

In this section we first provide a brief overview of fuzzy logic principles and then describe a practical system implementing them in an image forensics scenario.

A. Fuzzy logic principles

Fuzzy logic relies on three simple concepts: fuzzy sets, fuzzy operators and if-then rules. In the following we briefly introduce each of these concepts.

Let $X = \{x\}$ be a space of objects. A fuzzy set A in X is a class of objects of X characterized by a membership function $\mu_A(x)$, that is a curve defining how each point $x \in X$ is mapped to a membership value (or grade of membership) in the interval $[0, 1]$. The value $\mu_A(x)$ represents the grade of membership of x in A[1] [9].

Let us now apply this concept to logic. Classic Boolean logic requires that a proposition is either true (1) or false (0). There are no other possible values to assign. Based on real world experience, Fuzzy logic affirms that a proposition is not always totally false or totally true but true or false to some grade in the interval $[0, 1]$. Doing so, it is possible to claim that a proposition is true, *more or less* true, *somewhat* true and so on. Since Boolean logic can be seen as a particular case of fuzzy logic where one can only assign values 0 and 1 to membership functions, the extension of logical operators is not too complicate. Given two fuzzy sets A and B, standard fuzzy logical operators can be redefined as follows:

$$\begin{aligned} \text{AND}(A,B) &= \min(A, B) \\ \text{OR}(A,B) &= \max(A, B) \\ \text{NOT}(A) &= 1 - A \end{aligned} \tag{4}$$

Let x and y be two fuzzy variables. Let A and B be fuzzy sets. A fuzzy if-then rule is commonly represented as follows:

$$\text{IF } x \text{ is } A \text{ THEN } y \text{ is } B \tag{5}$$

The first part of the rule (x is A) is called antecedent, the second part (y is B) consequent. An antecedent can also consists of an arbitrary number of expressions. The behavior of a system is usually described by means of a set of if-then rules because most of times one rule alone is not effective. In a nutshell, a fuzzy system receives input variables that are crisp numbers (e.g. a measure of temperature) and need to be turned into something fuzzy. This task is performed by means of fuzzy sets. Once input values are transformed into fuzzy entities, they are combined accordingly to if-then rules. Result is something fuzzy and usually needs to be turned again into something crisp.

B. Fuzzy decision fusion framework

The proposed framework is constructed as follows: a set of if-then rules is derived from tables T_{true} and T_{false}; sequences in T_{doubt} are mapped into standard cases resulting in another set of if-then rules; all rules are applied to pairs (D,R) provided by the forensic tools producing a number that needs to be compared with a threshold to obtain a final binary answer on image region authenticity.

1) Fuzzy sets: Intuitively, detection and reliability values can either be considered *low* or *high*, where with low and high we mean fuzzy sets characterized by a membership function. Similarly, the presence of tampering derived from fusion of (D, R) of all tools can have different degrees of intensity. In our implementation we have chosen five fuzzy sets for the presence of tampering: *very weak*, *weak*, *neither weak nor strong*, *strong* and *very strong*.

2) Standard if-then rules: By looking at columns of T_{true} and T_{false} we create the so called *standard* if-then rules. These tables describe the behavior we expect from the tools in presence of a certain tampering. Intuitively, we try to assign them a linguistic meaning. Generally there is not a tool that is either wholly capable or incapable of detecting a certain tampering, but rather a tool that is *more* or *less* capable or incapable. Let us focus on the capability of detection (Y): if a tool provides a high value of detection with a high reliability we consider it *more* capable of detecting. We consider the same tool less capable (but still able of correct discrimination) if it provides a high value of detection with a low reliability. Similarly for the incapability of detection (N). These concepts can be formalized as follows to produce the fuzzy output of a single forensic tool:

$$\begin{array}{lllll} \text{Y} & = & (\text{detection is } \textit{high} & \text{AND} & \text{reliability is } \textit{high}) \quad \text{OR} \\ & & (\text{detection is } \textit{high} & \text{AND} & \text{reliability is } \textit{low}) \\ \text{N} & = & (\text{detection is } \textit{low} & \text{AND} & \text{reliability is } \textit{high}) \quad \text{OR} \\ & & (\text{detection is } \textit{low} & \text{AND} & \text{reliability is } \textit{low}) \end{array} \tag{6}$$

In section II-B we saw that columns of T_{true} and T_{false} are K-dimensional arrays whose elements are either Y or

[1] For example, if X is the space of temperatures, a temperature value x participates to the fuzzy set $A = \textit{hot}$ with grade $\mu_A(x)$, to the fuzzy set $B = \textit{cold}$ with grade $\mu_B(x)$ and so on. Each of these sets is characterized by a specific membership function. A value of $\mu_A(x)$ near 1 indicates a high grade of membership of temperature x in *hot*.

N. Let $\boldsymbol{s}$ be one of these arrays. The antecedent of an if-then rule is built by substituting to each element of $\boldsymbol{s}$ the corresponding expression as shown in equation 6. The choice of the consequent of rule depends on whether $\boldsymbol{s}$ belongs to T_{false} or T_{true}:

$$IF \begin{cases} \boldsymbol{s} \in T_{true} & \text{THEN tampering is } \textit{very strong} \\ \boldsymbol{s} \in T_{false} & \text{THEN tampering is } \textit{very weak} \end{cases} \tag{7}$$

Depending on the crisp input values (D, R), the processes of fuzzification and logical combination assign to each rule a certain degree of support. Each degree is used to truncate the corresponding output fuzzy set.

3) Non-standard if-then rules: If-then rules for *non-standard* cases of T_{doubt} are built similarly to those for standard cases. Antecedents are generated again as described in equation 6. However, some further reasoning is required to define consequents. When a non-standard case occurs we do not have a support from theory or experiments. Therefore we need to map this case into something that we know, according to the reliability of the various tools. The more a tool is reliable, the more we are willing to trust it. Let $\boldsymbol{ns}$ be a non-standard sequence belonging to T_{doubt} and $\boldsymbol{s}$ a standard sequence belonging either to T_{true} or T_{false}. Let us create a binary sequence by assigning values 0 and 1 to N and Y respectively. We propose to evaluate the distance between $\boldsymbol{ns}$ and $\boldsymbol{s}$ by means of the following weighted Hamming distance:

$$d(\boldsymbol{ns}, \boldsymbol{s}) = \sum_{i=1}^{K} (1 - R_i) XOR\big[\boldsymbol{ns}(i), \boldsymbol{s}(i)\big] \tag{8}$$

where: K is the number of tools; $w_i = (1 - R_i)$ is a weight that depends on tool reliability; XOR is the bitwise exclusive-OR operator; $\boldsymbol{ns}(i)$ and $\boldsymbol{s}(i)$ are the i-th bits of $\boldsymbol{ns}$ and $\boldsymbol{s}$ respectively. With equation 8 we compute the distance of $\boldsymbol{ns}$ from all the M standard sequences and select the closest one as follows:

$$\boldsymbol{s}_{min} = \arg\min_{n} \big[d(\boldsymbol{ns}, \boldsymbol{s}_n)\big], n = 1, 2, .., M \tag{9}$$

Since this process is an approximation based on experimental parameters, it is not wise to lean too much towards presence or absence of tampering. Therefore we choose to mitigate the intensity of the consequent as follows:

$$IF \begin{cases} \boldsymbol{s}_{min} \in T_{true} & \text{THEN tampering is } \textit{strong} \\ \boldsymbol{s}_{min} \in T_{false} & \text{THEN tampering is } \textit{weak} \end{cases} \tag{10}$$

With this approach a problem may arise: if two or more tools are equally reliable, we may have more than just one $\boldsymbol{s}$ at distance d_{min}. Although this does not happen frequently in the experiments we conducted, we have chosen to proceed as follows:

- if all the sequences $\boldsymbol{s}$ at distance d_{min} belong to T_{true} or T_{false} we choose the first $\boldsymbol{s}$ of the set and we define consequent as described in equation 10.
- if $\boldsymbol{ns}$ is equally close to at least one $\boldsymbol{s}$ belonging to T_{true} and one to T_{false} we change the consequent as follows:

$$\text{THEN tampering is } \textit{neither weak nor strong} \tag{11}$$

Note that this is not a fuzzy task: mapping of non standard cases is performed before building the fuzzy inference system.

V. Experimental analysis

In this section we describe a practical implementation of the two proposed approaches. Our goal is to validate the ideas we formalized above on a realistic image forensic scenario. We first briefly describe the tools that we employed and the dataset of images used to evaluate fusion accuracy. We choose to compare experimental results of both approaches with two other techniques: the first one is the binary OR (maximum criterion) operator, applied to tool thresholded outputs; the second is the more complex classification obtained using a two-class Support Vector Machine, using the detection values obtained from algorithms as features (we use a RBF kernel with parameters C = 0.1 σ = 2.48).

A. Implemented forensic tools

The two proposed decision fusion models are evaluated by fusing outputs obtained from $K = 3$ algorithms for tampering detection. More specifically, we implemented the one from Luo et al. [1], the one from Lin et al. [2] and the one from Farid [3]. In the following we will refer to them as T_A, T_B and T_C. All of these tools can be used to check whether a certain region of the image has been substituted with one cropped from another image, before performing a last JPEG re-compression of the resulting image.

We conducted our experiments on a dataset of 1600 JPEG compressed images by checking integrity of a 256 × 256 region located in the center of each image. Among these 1600 images, 800 are kept unmodified and 800 are used to simulate 4 different classes of cut & paste tampering. Each class has been designed so that only a single tool (or a pair of tools) is able to detect the presence of the manipulation[2]. Depending on alignment or misalignment of 8×8 grids of first and latter JPEG compression and on their respective quality factors, a specific tool may or may not be able to detect a manipulation (see table I for a brief description of each tampering procedure). According to the principles underlying each tool and to a preliminary experimental analysis we carried out on them, compatibilities turn out to be as in table II.

According to the assumptions made in section II, each tool has to output a value of detection in [0,1], where values near 1 indicate a high confidence about the analyzed region being tampered. For T_A, this value is obtained using the approach in [16] to get a probabilistic output from the SVM (training has been performed on a separated dataset); for T_B, the detection is taken as the median (over the suspected region) of the probability map [2]; for T_C, the value of the KS statistics is directly used [3].

[2] Notice that these tampering are tailored to tools requirements, thus making unpredictable phenomena in image features very unlikely, while tampering conducted on real world may actually expose them.

TABLE I

Construction of image dataset. Starting from uncompressed images 4 tampering classes have been created by means of variation of number of compressions and grid alignment. A fifth class of non tampered images has been added.

Class	Tampering procedure
Class 1	Outer region is compressed once. Inner region is compressed twice with misaligned grids
Class 2	Outer region is compressed twice with aligned grids. Inner region is compressed twice with misaligned grids
Class 3	Outer region is compressed once. Inner region is compressed twice with aligned grids
Class 4	Outer region is compressed twice with aligned grids. Inner region is compressed once
Class 5	Image is compressed once with a random but fairly high $QF \in \{70, 75, 80, 90\}$

TABLE II

Expected interactions between the 3 tools. First 4 columns correspond to T_{true}, fifth column to T_{false}. Cases that are not present in the table belong to T_{doubt}.

Tool	Class 1	Class 2	Class 3	Class 4	Class 5
T_A	Y	Y	N	N	N
T_B	N	Y	N	Y	N
T_C	N	Y	Y	Y	N

B. *Dempster-Shafer system settings*

Based on knowledge about tools performance and on previous experiments, we noticed that T_A is more reliable when the second JPEG quality factor QF_2 is high. Reliability of T_B and T_C does not seem to be affected by QF_2. Therefore we have decided to use the following values for reliability: $R_A = 0.4 \cdot QF_2$ (where QF_2 is normalized in [0,1]), $R_B = 0.4$ and $R_C = 0.85$. The mapping of detection values into BBAs (eq. 3) are reported in figure 1: curves are chosen in order to have a low false alarm rate (P_F) for each algorithm considered separately, so the doubt is maximum in correspondence of the threshold that give a P_F of 3% on the dataset, estimated using the single tool ROC computed before. These curves are kept constant during the fusion process.

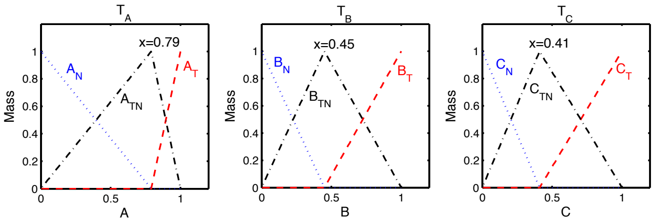

Fig. 1. Mapping from detection value to BBA for T_A (left), T_B (center) and T_C (right).

C. *Fuzzy inference system settings*

The fuzzy system features 6 inputs ($D_{A,B,C}$, $R_{A,B,C}$) and one output (tampering), that needs to be defuzzified to obtain the final decision. Similarly to the DS framework we have used $R_A = 0.4 \cdot QF_2$, $R_B = 0.4$ and $R_C = 0.85$. We used trapezoidal membership functions for all variables because they are simpler to implement and build automatically; however experiments have shown that use of smoother functions does not provide appreciable benefits. Figure 2 shows that each input can belong to two fuzzy sets: *low* and *high*. The point where the functions cross is where the maximum fuzziness is measured, since an input value is characterized by the same grade of membership for both classes. Values to the left of this point have a higher grade of membership in the fuzzy set *low*; values to the right of this point have a higher grade of membership in the fuzzy set *high*. Again, similarly to the configuration of DS framework, good values for such points are 0.79 for T_A, 0.45 for T_B, 0.41 for T_C and 0.5 for R_A, R_B and R_C.

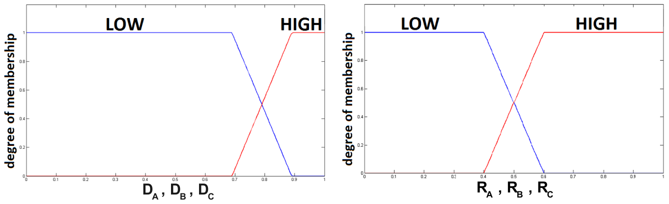

Fig. 2. Membership functions for input variables: detection (left) and reliability (right). Note that only D_A is shown here. D_B and D_C have the same shape but different points of maximum fuzziness.

Figure 3 shows membership functions for the output variable representing intensity of tampering. We have defined five possible fuzzy sets: from left to right *very weak*, *weak*, *neither weak nor strong*, *strong*, *very strong*.

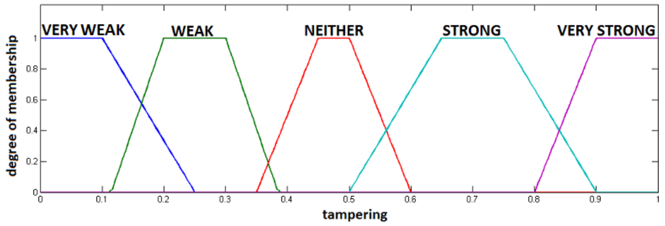

Fig. 3. Membership functions for output variable (tampering). From left to right: very weak, weak, neither weak nor strong, strong, very strong.

D. *Results and discussion*

The proposed methods have been tested on the dataset described in section V-A; the two-class (tampered vs. non-tampered) SVM has been trained using 800 out of the 1600 images. We decided to evaluate performance of all systems at low values of false alarm P_F since this is the most common operational condition used in practical applications. Table III reports the accuracy of detection (P_D) for low values of probability of false alarm (P_F) [3].

[3] According to [16], posterior probabilities obtained for the SVM are well fitted by a sigmoid, whose parameters are learned using training examples. When a new example must be classified, first its posterior is calculated, then the optimum threshold (which is proved to be 0.5) is applied; however, by sampling different thresholds, we can build a receiver operating curve. Values for P_F and P_D for the SVM are obtained in this way.

TABLE III
DETECTION ACCURACY OF THE TWO PROPOSED METHODS COMPARED TO LOGIC OR. WE CONSIDERED ONLY THOSE CASES WHERE P_F IS LOW.

	P$_F$=0%	**P$_F$=5%**	**P$_F$=10%**	**P$_F$=15%**	**P$_F$=20%**
DS	53.8%	76.5%	**84.3%**	**87.6%**	**91.3%**
Fuzzy	**59.7%**	**80.4%**	84.1%	85.4%	88.3%
SVM	50.5%	77.0%	84.0%	**87.6%**	90.7%
OR	55.5%	76.0%	81.2%	84.1%	86.2%

Results are promising, although not dramatically better than those obtained with a decision method based on logic OR and on SVMs. This can be explained by noting that the classes of tampering used in the experiments have been designed so that at least one tool is able to correctly detect the tampering (at least in principle). We did not introduce any unknown tampering that could alter the analyzed features. In addition, the number of tools we considered is quite limited. This is a case that is likely to be managed quite satisfactorily even by a simple OR operator, nevertheless, the proposed methods still perform better. In real world scenarios conflict is more likely to arise, thus making logical disjunction an hazardous approach. In this case we expect that benefits brought by our systems will be more significant. On the SVM side, it should be noted that when more tools become available the training of such a classifier gets increasingly complex, because it is difficult to create a suitable training set. On the contrary, parameterization of both DS and Fuzzy models can be performed by analyzing each tool separately, thus simplifying the setup.

The optimized versions of our systems perform very well also from a computational point of view. In particular, on our whole dataset of 1600 pairs (D, R), the DS framework completes the fusion in less than 0.1 seconds and the fuzzy framework in less than 2 seconds (about 1 second to build T_{true}, T_{false} and T_{doubt}, 0.2 seconds to build the inference system and 0.5 seconds to resolve if-then rules).

VI. CONCLUSIONS

In this work we focused on the problem of decision fusion from an image forensic point of view. When more than one forensic tool is employed, several problems may arise if we need to make a single decision from outputs that are heterogenous, discording or incomplete. To address the decision fusion problem we proposed two frameworks based on Dempster-Shafer's Theory of Evidence and on Fuzzy Theory respectively. Results are promising, however several aspects are not yet fully explored, including: implementation of a larger set of forensic tools; test of accuracy on a real-world dataset of tampered images; extension of frameworks to the most complex case where the suspicious tampered region is not known a priori. Our experiments show that both frameworks provide comparable results. Therefore, in this case of study, there are no reasons to prefer one rather than the other. However, experiments conducted on a wider set of tools and image processing techniques may highlight strengths and weaknesses of one approach with respect to the other

ACKNOWLEDGEMENTS

This work has been partially supported by the ICT-FET project REWIND - REVerse engineering of audio-VIsual coNtent Data funded by the European Commission.

REFERENCES

[1] W. Luo, Z. Qu, J. Huang, and G. Qiu, "A novel method for detecting cropped and recompressed image block," in *Proc. of ICASSP 2007*, vol. 2, 2007, pp. II–217 –II–220.

[2] Z. C. Lin, J. F. He, X. Tang, and C. K. Tang, "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," *Pattern Recognition*, vol. 42, no. 11, pp. 2492–2501, Nov. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2009.03.019.

[3] H. Farid, "Exposing digital forgeries from JPEG ghosts," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 1, pp. 154–160, 2009. [Online]. Available: http://dx.doi.org/10.1109/TIFS.2008.2012215

[4] E. Delp, N. Memon, and M. Wu, "Special issue on digital forensics," *IEEE Signal Processing Magazine*, vol. 26, no. 2, 2009.

[5] M. Kharrazi, H. T. Sencar, and N. Memon, "Improving steganalysis by fusion techniques: A case study with image steganography," in *T. Data Hiding and Multimedia Security*, 2006, pp. 123–137.

[6] S. Bayram, I. Avcibas, B. Sankur, and N. Memon, "Image manipulation detection," *J. Electronic Imaging*, vol. 15, no. 4, p. 041102, 2006. [Online]. Available: http://dx.doi.org/10.1117/1.2401138

[7] Y. F. Hsu and S. F. Chang, "Statistical fusion of multiple cues for image tampering detection," in *42nd Asilomar Conference on Signals, Systems and Computers*, 2008, pp. 1386–1390.

[8] G. Shafer, *A Mathematical Theory of Evidence*. Princeton: Princeton University Press, 1976.

[9] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[10] ——, "Outline of a new approach to the aanalysis of complex systems and decision," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-3, no. 1, pp. 28–44, 1973.

[11] T. Terano, K. Asai, and M. Sugeno, *Fuzzy Systems Theory and its Applications*. Academic Press Boston, 1992.

[12] E. H. Ruspini, P. Bonissone, and W. Pedrycz, *Handbook of Fuzzy Computation*. Institute of Physics Pub. Bristol and Philadelphia, 1998.

[13] A. Benavoli, L. Chisci, B. Ristic, A. Farina, and A. Graziano, *Reasoning under uncertainty: from Bayesian to Valuation Based Systems*. ISBN: 978-8886658430, 2007.

[14] L. A. Zadeh, "A mathematical theory of evidence (book review)," *AI Magazine*, vol. 5, no. 3, pp. 81–83, 1984.

[15] R. Haenni, "Shedding new light on Zadeh's criticism of Dempster's rule of combination," in *Information Fusion, 2005 8th International Conference on*, vol. 2, July 2005, p. 6 pp.

[16] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in large margin classifiers*. MIT Press, 1999, pp. 61–74.