Evidential Fusion for Gender Profiling

Jianbing Ma and Weiru Liu and Paul Miller

School of Electronics, Electrical Engineering and Computer Science,
Queen's University Belfast, Belfast BT7 1NN, UK
{jma03,w.liu}@qub.ac.uk, p.miller@ecit.qub.ac.uk

Abstract. Gender profiling is a fundamental task that helps CCTV systems to provide better service for intelligent surveillance. Since subjects being detected by CCTVs are not always cooperative, a few profiling algorithms are proposed to deal with situations when faces of subjects are not available, among which the most common approach is to analyze subjects' body shape information. In addition, there are some drawbacks for normal profiling algorithms considered in real applications. First, the profiling result is always uncertain. Second, for a time-lasting gender profiling algorithm, the result is not stable. The degree of certainty usually varies, sometimes even to the extent that a male is classified as a female, and vice versa. These facets are studied in a recent paper [16] using Dempster-Shafer theory. In particular, Denoeux's cautious rule is applied for fusion mass functions through time lines. However, this paper points out that if severe mis-classification is happened at the beginning of the time line, the result of applying Denoeux's rule could be disastrous. To remedy this weakness, in this paper, we propose two generalizations to the DS approach proposed in [16] that incorporates time-window and time-attenuation, respectively, in applying Denoeux's rule along with time lines, for which the DS approach is a special case. Experiments show that these two generalizations do provide better results than their predecessor when mis-classifications happen.

Keywords: Gender Profiling; Gender Recognition; Evidence Theory; Cautious Rule.

1 Introduction

Nowadays, CCTV systems are broadly deployed in the present world, e.g., Florida School Bus Surveillance project [1], the First Glasgow Bus Surveillance [21], Federal Intelligent Transportation System Program in the US [20], Airport Corridor Surveillance in the UK [19, 17, 18, 13], etc. However, despite the wide-range use of CCTVs, the impact on anti-social and criminal behaviour has been minimal. For example, assaults on bus and train passengers are still a major problem for transport operators. That is, surveillance systems are not capable of reacting events of interest instantly.

A key requirement for active CCTV systems is to automatically determine the threat posed by each individual to others in the scene. Most of the focus of the computer vision community has been on behaviour/action recognition. However, experienced security analysts profile individuals in the scene to determine their threat. Often they can identify individuals who look as though they may cause trouble before any anti-social behaviour has occurred. From criminology studies, the vast majority of offenders are young adolescent males. Therefore, key to automatic threat assessment is to be able to automatically profile people in the scene based on their gender and age. In this paper, we focus on the former.

Although it is a fundamental task for surveillance applications to determine the gender of people of interest, however, normal video algorithms for gender profiling (usually face profiling) have three drawbacks. First, the profiling result is always uncertain. Second, for a time-lasting gender profiling algorithm, the result is not stable. The degree of certainty usually varies, sometimes even to the extent that a male is classified as a female, and vice versa. Third, for a robust profiling result in cases were a person's face is not visible, other features, such as body shape, are required. These algorithms may provide different recognition results - at the very least, they will provide different degrees of certainties. To overcome these problems, in [16], an evidential (Dempster-Shafer's (DS) theory of evidence) approach is proposed that makes use of profiling results from multiple profiling algorithms using different human features (e.g., face, full body) over a period of time, in order to provide robust gender profiling of subjects in video. Experiments show that this approach provides better results than a probabilistic approach.

DS theory [2, 22, 8, 9] is a popular framework to deal with uncertain or incomplete information from multiple sources. This theory is capable of modelling incomplete information through ignorance. For combining difference pieces of information, DS theory distinguishes two cases, i.e., whether pieces of information are from distinct, or non-distinct, sources. Many combination rules are proposed for information from distinct sources, among which are the well-known Dempster's rule [22], Smets' rule [23], Yager's rule [24], and Dubois & Prade's hybrid rule [4], etc. In [3], two combination rules, i.e., the cautious rule and the bold disjunctive rule, for information from non-distinct sources are proposed. Therefore, gender profiling results from the same classifier, e.g. face-based, at different times are considered as from non-distinct sources while profiling results from different classifiers are naturally considered as from distinct sources.

In [16], for gender profiling results from the same classifier at different time points, Denoeux's cautious rule [3] is used to combine them. For profiling results from different classifiers (i.e., face profiling and full body profiling), Dempster's rule [2, 22] is introduced to combine them. And finally, the pignistic transformation is applied to get the probabilities of the subject being male or female.

However, if severe mis-classification happens at the beginning of the time line, the result of applying Denoeux's rule could be disastrous. For instance, if a subject is classified as a female with a certainty degree 0.98, and later on it is classified as a male with certainty degrees from 0.85 to 0.95, then by Denoeux's cautious rule, it will be always classified as a female. In order to remedy this weakness, in this paper, we propose two generalizations on applying Denoeux's rule through time lines, in which one uses time-window and the other uses time-attenuation, respectively. In the time-window generalization, Denoeux's rule is applied only for the most recent n frames where n is a pre-given threshold depending on the time length. In the time-attenuation generalization, the certainty degree is reduced gradually by time at a pre-defined attenuation factor. Experiments show that these two generalizations do provide better results when mis-classifications happen, but they have to pay the price of performing less accurate in other situations than the fusion method proposed in [16]. In summary, we can say these two generalizations are more robust than their predecessor.

The rest of the paper is organized as follows. Section 2 provides the preliminaries on Dempster-Shafer theory. Subsequently, Section 3 introduces the two generalizations of the DS approach. In Section 4, we discuss the difficulties in gender profiling in terms of scenarios. Section 5 provides experimental results which shows our generalizations perform better than its predecessor and a classic fusion approach as well as single profiling approaches. Finally, we conclude the paper in Section 6.

2 Dempster-Shafer Theory

For convenience, we recall some basic concepts of Dempster-Shafer's theory of evidence. Let Ω be a finite, non-empty set called the frame of discernment, denoted as, $\Omega = \{w_1, \cdots, w_n\}$.

Definition 1 *A basic belief assignment(bba) is a mapping* $m : 2^{\Omega} \to [0, 1]$ *such that* $\sum_{A \subseteq \Omega} m(A) = 1$.

If $m(\emptyset) = 0$, then m is called a mass function. If $m(A) > 0$, then A is called a focal element of m. Let $\mathscr{F}_m$ denote the set of focal elements of m. A mass function with only a focal element Ω is called a *vacuous* mass function.

From a bba m, *belief* function (Bel) and *plausibility* function (Pl) can be defined to represent the lower and upper bounds of the beliefs implied by m as follows.

$$Bel(A) = \sum_{B \subseteq A} m(B) \text{ and } Pl(A) = \sum_{C \cap A \neq \emptyset} m(C). \tag{1}$$

One advantage of DS theory is that it has the ability to accumulate and combine evidence from multiple sources by using *Dempster's rule of combination*. Let m_1 and m_2 be two mass functions from two distinct sources over Ω. Combining m_1 and m_2 gives a new mass function m as follows:

$$m(C) = (m_1 \oplus m_2)(C) = \frac{\sum_{A \cap B = C} m_1(A) m_2(B)}{1 - \sum_{A \cap B = \emptyset} m_1(A) m_2(B)} \tag{2}$$

In practice, sources may not be completely reliable, to reflect this, in [22], a *discount rate* was introduced by which the mass function may be discounted in order to reflect the reliability of a source. Let r $(0 \leq r \leq 1)$ be a discount rate, a discounted mass function using r is represented as:

$$m^r(A) = \begin{cases} (1-r)m(A) & A \subset \Omega \\ r + (1-r)m(\Omega) & A = \Omega \end{cases} \tag{3}$$

When $r = 0$ the source is absolutely reliable and when $r = 1$ the source is completely unreliable. After discounting, the source is treated as totally reliable.

Definition 2 *Let* m *be a bba on* Ω. *A pignistic transformation of* m *is a probability distribution* P_m *over* Ω *such that* $\forall w \in \Omega, P_m(w) = \sum_{w \in A} \frac{1}{|A|} \frac{m(A)}{1 - m(\emptyset)}$ *where* $|A|$ *is the cardinality of* A.

Let $\oplus$ be the conjunctive combination operator (or Smets' operator [23]) for any two bbas m, m' over Ω such that

$$(m \oplus m')(C) = \sum_{A\subseteq\Omega, B\subseteq\Omega, A\cap B=C} m(A)m'(B), \forall\, C \subseteq \Omega. \tag{4}$$

A simple bba m such that $m(A) = x, m(\Omega) = 1 - x$ for some $A \neq \Omega$ will be denoted as A^x. The vacuous bba can thus be noted as A^0 for any $A \subset \Omega$. Note that this notation, i.e., A^x, is a bit different from the one defined in [3] in which A^x in our paper should be denoted as A^{1-x} in [3].

Similarly, for two sets $A, B \subset \Omega$, $A \neq B$, let A^xB^y denote a bba m such that $m = A^x \oplus B^y$ where $\oplus$ is the conjunctive combination operator defined in Equation (4). For these kinds of bbas, we call them *bipolar* bbas. A simple bba A^x could be seen as a special bipolar bba A^xB^0 for any set $B \subseteq \Omega$, $B \neq A$.

It is easy to prove that any $m = A^xB^y$ is:

$$m(\emptyset) = xy, m(A) = x(1-y), m(B) = y(1-x), m(\Omega) = (1-x)(1-y) \tag{5}$$

In addition, when normalized, m in Equation 5 is changed to m' as follows.

$$m'(A) = \frac{x(1-y)}{1-xy}, m'(B) = \frac{y(1-x)}{1-xy}, m'(\Omega) = \frac{(1-x)(1-y)}{1-xy} \tag{6}$$

For two bipolar bbas $A^{x_1}B^{y_1}$ and $A^{x_2}B^{y_2}$, the cautious combination rule proposed in [3] is as follows.

Lemma 1 *(Denœux's Cautious Combination Rule) Let $A^{x_1}B^{y_1}$ and $A^{x_2}B^{y_2}$ be two bipolar bbas, then the combined bba by Denœux's cautious combination rule is also a bipolar bba A^xB^y such that: $x = max(x_1, x_2), y = max(y_1, y_2)$.*

Also, according to [3], for $m_1 = A^{x_1}B^{y_1}$ and $m_2 = A^{x_2}B^{y_2}$, the combined result by Equation (2) is[1]

$$m_{12} = A^{x_1+x_2-x_1x_2}B^{y_1+y_2-y_1y_2} \tag{7}$$

3 Two Generalizations

In this section, we discuss two generalizations for the Cautious rule, i.e., the time-window approach and the time-attenuation approach. Let $\oplus_C$ be the operator defined by the Cautious rule.

Definition 3 *(Time-Window Cautious Combination Rule) Let $A^{x_1}B^{y_1}, \cdots, A^{x_n}B^{y_n}$ be n successive bipolar bbas, then the combined bba by Time-Window cautious combination rule of window size t is $m_t = A^{x_{n-t+1}}B^{y_{n-t+1}} \oplus_C \cdots \oplus_C A^{x_n}B^{y_n}$.*

That is, a time-window cautious rule of window size t only combines the recent t bbas.

[1] In [3], the combined result is $m_{12} = A^{x_1x_2}B^{y_1y_2}$, but recall that we use a slightly different notation from [3].

Definition 4 *(Time-Attenuation Cautious Combination Rule) Let $A^{x_1}B^{y_1}, \cdots, A^{x_n}B^{y_n}$ be n successive bipolar bbas, then the combined bba by Time-Attenuation cautious combination rule of attenuation factor t, $0 < t < 1$, is $m_t = A^{x_1 t^{n-1}} B^{y_1 t^{n-1}} \oplus_C \cdots \oplus_C A^{x_n}B^{y_n}$.*

That is, in a time-attenuation cautious rule of attenuation factor t, the coefficient is reduced by t each time. Hence if a male is mis-classified as a female with a certainty degree 0.98, and hence is represented as $M^0F^{0.98}$, will be attenuated gradually that it will not affect the cautious combination result for long since $0.98t^n$ will grow smaller when $0 < t < 1$ and n increases.

4 Gender Recognition Scenario

In this section, we provide a detailed description of a gender profiling scenario, which lends itself naturally to a DS approach.

Figure 1 shows three images taken from a video sequence that has been passed through a video analytic algorithm for gender profiling. In this sequence, a female wearing an overcoat with a hood enters the scene with her back to the camera. She walks around the chair, turning, so that her face becomes visible, and then sits down.

Fig. 1. Three images taken from a video sequence

(a) (b) (c)

Fig. 1(a) shows that the subject is recognised by the full body shape profiling as a male. Note that her face is not visible. In Fig. 1(b), the subject is classified as female by the full body shape profiling algorithm. In Fig. 1(c), as she sits down, with her face visible, the face profiling algorithm classifies her as female, whilst the full body profiling classifies her as male. Note that the full body profiling algorithm is not as reliable as the face profiling algorithm. Conversely, full body profiling is always possible whilst the face information can be missing. That is why these two profiling algorithms should be considered together. In addition, as full body profiling is not as robust, discount operations should be performed on the algorithm output (cf. Equation (3)). The discount rate is dependent on the video samples and the training efficiency. For every video frame in which a body (face) is detected, gender recognition results are provided. The full body profiling algorithm and the face profiling algorithm, provided a person's face is detected, report their recognition results for every frame of the video, e.g., male with 95% certainty.

For a frame with only a body profiling result, for instance Fig. 1(a), the corresponding mass function m for body profiling will be M^x where M denotes that *the person is classified as a male* and x is the mass value of $m(\{M\})$. The corresponding mass function for face profiling is M^0F^0 where F denotes that *the person is classified as a female*, or the vacuous mass function. Alternatively, we can refer to this as the vacuous mass function.

Similarly, for a frame with both body profiling and face profiling, for instance Fig. 1(c), the corresponding mass function for body profiling will be M^x (or in a bipolar form M^xF^0) and the mass function for face profiling is F^y (or in a bipolar form M^0F^y) where x, y are the corresponding mass values. As time elapses, fusion of bipolar bbas by the cautious rule or its two generalizations are introduced, as shown by Lemma 1 and Definition 3 and Definition 4. And when it comes to present the final profiling result, we use Dempster's rule to combine the two fused bipolar mass functions from the two recognition algorithms, respectively. Namely, for the two bipolar bbas $m_1 = M^{x_1}F^{y_1}$ and $m_2 = M^{x_2}F^{y_2}$, it is easy to get that the combined result m_{12} by Dempster's rule is (normalized from the result of Equation 7):

$$m_{12}(\{M\}) = \frac{(x_1 + x_2 - x_1x_2)(1 - y_1)(1 - y_2)}{1 - (x_1 + x_2 - x_1x_2)(y_1 + y_2 - y_1y_2)},$$

$$m_{12}(\{F\}) = \frac{(1 - x_1)(1 - x_2)(y_1 + y_2 - y_1y_2)}{1 - (x_1 + x_2 - x_1x_2)(y_1 + y_2 - y_1y_2)},$$

$$m_{12}(\Omega) = \frac{(1 - x_1)(1 - x_2)(1 - y_1)(1 - y_2)}{1 - (x_1 + x_2 - x_1x_2)(y_1 + y_2 - y_1y_2)}.$$

Finally, we use the pignistic transformation (Def. 2) for the final probabilities. That is, $p(\{M\}) = m_{12}(\{M\}) + m_{12}(\Omega)/2$ and $p(\{F\}) = m_{12}(\{F\}) + m_{12}(\Omega)/2$. Obviously, we will say the subject is a male if $p(\{M\}) > p(\{F\})$, and a female if $p(\{M\}) < p(\{F\})$. In very rare cases that $p(\{M\}) = p(\{F\})$, we cannot know whether it is male or female.

The following example illustrates the computation steps.

Example 1 *Let us illustrate the approach by a simple scenario with four frames, and there is a mis-classification in the first frame. In the first frame, the corresponding both body profiling (m_b^1) and face profiling (m_f^1) results as $m_b^1 = M^{0.6}$ and $m_f^1 = F^{0.9}$ (mis-classification). In the second frame, there is only a body profiling (m_b^2) result which is $m_b^2 = M^{0.7}$. Frame three is associated with body profiling (m_b^3) and face profiling (m_f^3) results as $m_b^3 = F^{0.4}$ and $m_f^3 = M^{0.6}$, and frame four is associated with body profiling (m_b^4) and face profiling (m_f^4) results as $m_b^4 = M^{0.6}$ and $m_f^4 = M^{0.6}$.*

By Lemma 1, the fusion results by the cautious rule are $m_b = M^{0.7}F^{0.4}$ and $m_f = M^{0.6}F^{0.9}$.

By Definition 3 with window size 2, the fusion results by the time-window cautious rule are $m_b^W = M^{0.6}F^{0.4}$ and $m_f^W = M^{0.6}$.

By Definition 4 with attenuation factor 0.95, the fusion results by the time-attenuation cautious rule are $m_b^A = M^{0.6}F^{0.38}$ and $m_f^A = M^{0.6}F^{0.77}$.

Then by Equation 7, we get $m_{bf} = M^{0.88}F^{0.94}$*, which, when normalized, is equivalent to* $m_{bf}(\{M\}) = \frac{0.88(1-0.94)}{1-0.88*0.94} = 0.31$*,* $m_{bf}(\{F\}) = \frac{0.94(1-0.88)}{1-0.88*0.94} = 0.65$*,* $m_{bf}(\Omega) = \frac{(1-0.88)(1-0.94)}{1-0.88*0.94} = 0.04$*. And finally we get* $p(\{M\}) = 0.33$ *and* $p(\{F\}) = 0.67$ *which indicates that the subject is a female.*

Similarly, we have $m^W_{bf} = M^{0.84}F^{0.4}$*, and hence* $m^W_{bf}(\{M\}) = \frac{0.84(1-0.4)}{1-0.84*0.4} = 0.76$*,* $m^W_{bf}(\{F\}) = \frac{0.4(1-0.84)}{1-0.84*0.4} = 0.10$*,* $m^W_{bf}(\Omega) = \frac{(1-0.84)(1-0.4)}{1-0.84*0.4} = 0.14$ *and* $p^W(\{M\}) = 0.83$ *and* $p^W(\{F\}) = 0.17$*, which indicates that the subject is a male.*

Also, we have $m^A_{bf} = M^{0.88}F^{0.857}$*, and hence* $m^A_{bf}(\{M\}) = \frac{0.88(1-0.857)}{1-0.88*0.857} = 0.51$*,* $m^A_{bf}(\{F\}) = \frac{0.857(1-0.88)}{1-0.88*0.857} = 0.42$*,* $m^A_{bf}(\Omega) = \frac{(1-0.88)(1-0.857)}{1-0.88*0.857} = 0.07$ *and* $p^A(\{M\}) = 0.55$ *and* $p^A(\{F\}) = 0.45$ *which also supports that the subject is a male.*

5 Experimental Results

In this section we compare fusion results obtained by a classic approach, a Dempster-Shafer theory approach proposed in [16] and two of its generalization approaches. As there are no benchmark datasets for both body and face profiling, we simulate the output of both body and face classifiers on a sequence containing a male subject. For the body classifier, the probability of any frame being correctly classified as male/female is roughly 60-90%. For the face classifier, only 75% of the available frames are randomly allocated as containing a face. For each of these frames the probability of the frame being correctly classified as being male/female is 85-100%. In both cases the values for $m(\{M\})$ and $m(\{F\})$ are uniformly sampled from the ranges 0.6-0.9 and 0.85-1.0 for the body and face classifiers outputs respectively.

As mentioned before, for gender profiling results from the same classifier at different time points, we use the cautious rule to combine them. For profiling results from different classifiers (i.e., face profiling and full body profiling), we use Dempster's rule to combine them. And finally, we apply the pignistic transformation (Def. 2) to get the probabilities of the subject being male or female.

Classic fusion in the computer vision community [25] takes the degrees of certainty as probabilities, i.e., they consider the face profiling and the full body profiling output p^t_f and p^t_b indicating the probabilities of faces and full bodies being recognized as males at time t. Then it uses $p^t_{b,f} = c^t_f p^t_f + c^t_b p^t_b$ to calculate the final probability $p^t_{b,f}$ at time t, where c^t_f and c^t_b are the weights of the face and full body profiling at time t, proportional to the feasibility of the two algorithms in the last twenty frames. As full body profiling is always feasible, suppose face profiling can be applied n times in the last twenty frames, then we have:

$$c_b = \frac{20}{20+n}, c_f = \frac{n}{20+n}.$$

For this experiment, the performance of the DS and classic fusion schemes were characterised by the true positive rate:

$$T_{PR} = \frac{N_{PR}}{N}$$

where N_{PR} is the number of frames in which the gender has been correctly classified and N is the total number of frames in which the body/face is present. According to the training on the sample videos, the discount rate r for the full body profiling is set to 0.3. For comparison, we calculate the T_{PR} value for the body classifier alone, the face classifier, the DS fusion scheme and the classic fusion scheme.

Here, we first apply the approaches to 58 simulations each with 50 frames (so there are 2900 total frames), where a mis-classification happens at the beginning. The comparison results are presented as follows.

Methods	N	N_{PR}	T_{PR} (%)
Full Body	2900	1606	55.4
Face	2159	2002	92.7
Classic Method	2900	2078	71.7
DS Approach	2900	2380	82.1
Time-Attenuation (0.95)	2900	2194	75.7
Time-Attenuation (0.99)	2900	2431	83.8
Time-Window (5)	2900	2586	89.2

Table 1: Comparison of T_{PR} for body classification, face classification, classic fusion, DS fusion and its two adaptions - Mis-Classification Cases.

From Table 1, we can see that the two generalizations provide better results than the DS fusion scheme, except when the attenuation factor is 0.95. This may be because setting the attenuation factor to 0.95 reduces the certainty degrees too quickly.

An example simulation result comparing the classic, DS, Time-Attenuation (0.99) and Time-Window (5) approaches is shown in Fig. 2.

Now we apply the approaches to 20 simulations each with 150 frames (so there are 3000 total frames), where we do not assume mis-classification happened at the beginning. The comparison results are presented as follows.

Methods	N	N_{PR}	T_{PR} (%)
Full Body	3000	1792	59.7
Face	2229	2125	95.3
Classic Method	3000	2490	83.0
DS Approach	3000	2899	96.6
Time-Attenuation (0.95)	3000	2126	70.9
Time-Attenuation (0.99)	3000	2401	80.0
Time-Window (5)	3000	2395	79.8
Time-Window (20)	3000	2552	85.1

Table 2: Comparison of T_{PR} for body classification, face classification, classic fusion, DS fusion and its two adaptions - General Cases.

From Table 2, we can see that the two generalizations perform worse than the DS fusion scheme. This is not surprising since the former do not always hold the highest certainty

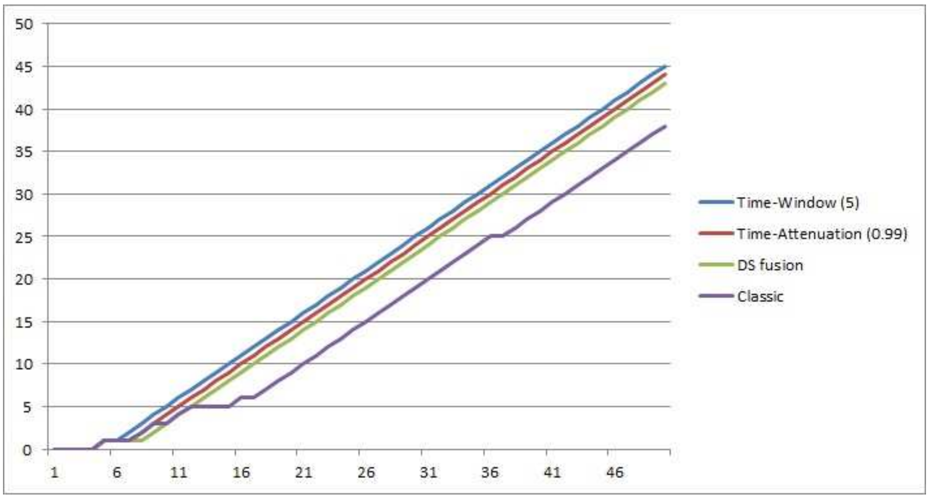

Fig. 2. An Example Simulation

degree as in the DS fusion scheme. Table 2 also shows that when the attenuation factor or the window size increases, the results improve. Actually, if the attenuation factor is one or the window size equals to the number of frames, then these two generalizations will provide the same results as the DS fusion one, or we can see the DS fusion scheme is a special case of these two generalizations.

6 Conclusion

In this paper, we have proposed two generalized fusion methods to combine gender profiling classifier results by modifying the application of the Cautious rule, i.e., the time-window fusion method and the time-attenuation fusion method. Experimental results show that these two generalizations provide more robust results than other approaches, especially to their predecessor DS fusion scheme.

From the experimental results, it suggests that the time-window fusion scheme performs slightly better than the time-attenuation fusion scheme. But we think this conclusion still depends on the choice of attenuation factor, window size and frame size.

For future work, we plan to apply the fusion schemes to profiling classifier results generated from real video sequences. Also, for the time-attenuation generalization, we are trying to use the well-known attenuation approach used in machine learning as:

$$x'_n = x'_{n-1}(1-\alpha) + x_n\alpha,$$

where α is an attenuation factor, and see whether this will be a better choice. In addition, we are also exploiting ideas from knowledge base merging [5, 11, 6, 7], statistical fusion [10, 12] and calculi on sequential observations [14, 15].

Acknowledgement This research work is sponsored by the EPSRC projects EP/G034303/1 and EP/H049606/1 (the CSIT project).

References

1. Bsia. Florida school bus surveillance. In *http://www.bsia.co.uk/LY8VIM18989_action;displaystudy_sectorid;LYCQYL79312_caseid;NFLEN064798*.
2. A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Statistics*, 28:325–339, 1967.
3. Thierry Denœux. Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artifical Intelligence*, 172(2-3):234–264, 2008.
4. D. Dubois, H. Prade, and R. Yager. Fuzzy set connectives as combinations of belief structures. *Information Sciences*, 66:245–275, 1992.
5. S. Konieczny and R. Pino-Pérez. On the logic of merging. In A. G. Cohn, L. Schubert, and S. C. Shapiro, editors, *KR'98, Principles of Knowledge Representation and Reasoning*, pages 488–498, San Francisco, California, 1998. Morgan Kaufmann.
6. J. Ma and W. Liu. Modeling belief change on epistemic states. In *Proc. of 22th Flairs*, pages 553–558. AAAI Press, 2009.
7. J. Ma and W. Liu. A framework for managing uncertain inputs: An axiomization of rewarding. *Inter. Journ. of Approx. Reasoning*, 52(7):917–934, 2011.
8. J. Ma, W. Liu, D. Dubois, and H. Prade. Revision rules in the theory of evidence. In *Procs. of ICTAI*, pages 295–302, 2010.
9. J. Ma, W. Liu, D. Dubois, and H. Prade. Bridging jeffrey's rule, agm revision and dempster conditioning in the theory of evidence. *International Journal on Artificial Intelligence Tools*, 20(4):691–720, 2011.
10. J. Ma, W. Liu, and A. Hunter. Incomplete statistical information fusion and its application to clinical trials data. In *Procs. of the First International Conference on Scalable Uncertainty Management (SUM07)*, pages 89–103, 2007.
11. J. Ma, W. Liu, and A. Hunter. The non-archimedean polynomials and merging of stratified knowledge bases. In *Procs. of ECSQARU*, pages 408–420, 2009.
12. J. Ma, W. Liu, A. Hunter, and W. Zhang. Performing meta-analysis with incomplete statistical information in clinical trials. *BMC Medical Research Methodology*, 8(1):56, 2008.
13. J. Ma, W. Liu, and P. Miller. Event modelling and reasoning with uncertain information for distributed sensor networks. In *Procs. of SUM*, pages 236–249, Toulouse, France, 2010. Springer.
14. J. Ma, W. Liu, and P. Miller. Belief change with noisy sensing in the situation calculus. In *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence (UAI'2011)*, pages 471–478, 2011.
15. J. Ma, W. Liu, and P. Miller. Handling sequential observations in intelligent surveillance. In *Proceedings of the 5th International Conference on Scalable Uncertainty Management (SUM'2011)*, pages 547–560, 2011.
16. J. Ma, W. Liu, and P. Miller. An evidential improvement for gender profiling. In T. Denoeux and M. Masson (Eds.), editors, *Belief Functions: Theory and Applications, Procs. of BELIEF'2012*, 2012.
17. J. Ma, W. Liu, P. Miller, and W. Yan. Event composition with imperfect information for bus surveillance. In *Procs. of AVSS*, pages 382–387. IEEE Press, 2009.
18. P. Miller, W. Liu, F. Fowler, H. Zhou, J. Shen, J. Ma, J. Zhang, W. Yan, K. McLaughlin, and S. Sezer. Intelligent sensor information system for public transport: To safely go .... In *Procs. of AVSS*, 2010.
19. ECIT Queen's University of Belfast. Airport corridor surveillance. *http://www.csit.qub.ac.uk/Research/ResearchGroups/IntelligentSurveillanceSystems*, 2010.
20. US Department of Transportation. Rita - its research program. *http://www.its.dot.gov/ITS_ROOT2010/its_program/ ITSfederal_program.htm*, 2010.

21. Gardiner Security. Glasgow transforms bus security with ip video surveillance. In *http://www.ipusergroup.com/doc-upload/Gardiner-Glasgowbuses.pdf*.
22. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
23. P. Smets. Non-standard logics for automated reasoning. In P. Smets, A. Mamdani, D. Dubois, and H. Prade, editors, *Belief functions*, pages 253–286, 1988.
24. R. Yager. On the dempster-shafer framework and new combination rules. *Information Sciences*, 41:93–138, 1987.
25. H. Zhou, P. Miller, J. Zhang, M. Collins, and H. Wang. Gender classification using facial and full body features. In *Technical Report, CSIT, Queen's University Belfast, UK*, 2011.