In Proceedings of the International Conference on Internet of Things and Machine Learning (IML '17). Liverpool, United Kingdom, October 17-18, 2017. ACM Digital Library.
DOI:10.1145/3109761.3158409

A New Approach for Multi-Biometric Fusion Based on Subjective Logic

Kamer Vishi[*] and Audun Jøsang[*]

Abstract
Biometric verification systems have to address many practical requirements, such as performance, presentation attack detection (PAD), large population coverage, demographic diversity, and varied deployment environment. Traditional unimodal biometric systems do not fully meet the aforementioned requirements making them vulnerable and susceptible to different types of attacks. In response to that, modern biometric systems combine multiple biometric modalities at different fusion levels, such as sensor, feature, score and decision level. The fused score is decisive to classify an unknown user as a genuine or impostor. In this paper, we describe a new biometric fusion framework based on *Subjective Logic (SL)*; a type of probabilistic logic that explicitly takes uncertainty and trust into consideration. We principally evaluate our proposed fusion framework using two modalities, namely iris and fingerprint. Furethermore, the individual scores obtained from various comparators are combined at score level by applying four score fusion approaches (minimum score, maximum score, simple sum, and subjective logic) and three score normalization techniques (min-max, z-score, hyperbolic tangent). The experimental results show that the proposed score level fusion approach (subjective logic) gives the best authentication accuracy even when particular biometric classifiers give distinct comparison scores.

Keywords
Multimodal Biometrics – Belief Fusion – Subjective Logic – Quality Assessment – Fingerprint – Iris Recognition

[*] ***Affiliation:*** *SecurityLab, Department of Informatics, University of Oslo (UiO) - Norway*
*__E-mail__: kamerv@ifi.uio.no — josang@ifi.uio.no

1. Introduction

The most important practical consideration for a biometric system is its accuracy. The combination of multiple different samples or multiple biometric modalities is a natural approach to improve the quality of biometric systems, which in general is called biometric fusion.

Biometric fusion can be performed at various levels, such as sensor level, feature level, comparison[*] score level and decision level. Score level fusion is the most popular and suitable way of biometric fusion as all commercially available biometric sensors, and feature extractors do not provide access to the feature extraction algorithms (also known as *"black-box"*). Advantages of multimodal biometric systems over the unimodal systems have been discussed by Ross et al. [2].

This paper proposes a new framework for biometric data fusion based on the application of subjective logic (cumulative fusion), called *BDF-SL*. Prior to fusion in uncorrelated modalities, a score normalisation process is required as discussed by Jain et al. [3]; hence we have used three score normalisation techniques, such as Min-Max, Z-Score and Hyperbolic Tangent estimator method.

The rest of the paper is organized as follows: Section 2 discusses the, Section 3 describes multimodal biometrics and score level fusion, Section 4 describes the opinion representation in subjective logic as well as the relevant fusion models, Section 5 proposes how these fusion models can be applied in biometric fusion, in Section 6 our proposed multimodal biometric fusion technique at score level is described, in Section 7 and 8 experiments, data analysis and experimental results are given, and lastly Section 9 and 10 rounds off the paper with a discussion, indication of future work and conclusion.

[*]Note that the term *"matching"* as a synonym for *"comparison"* has been deprecated in the ISO SC37 Harmonized Biometric Vocabulary [1].

2. Related work

Different score fusion approaches have been proposed for fusing scores received from different biometric comparators.

Raghavendra et al. [4], proposed a multimodal biometric score level fusion scheme using Gaussian Mixture Model and Monte Carlo Method. The authors have fused face, speech and palm print modalities from three different databases and experimental results show that their method gives higher performance than some other fusion schemes, such as simple sum rule, weighted sum rule, and Likelihood Ratio (LR) methods.

Derawi et al. [5], fused gait and fingerprint scores for user authentication on mobile devices. The authors have used four normalisation techniques and four fusion approaches to combine comparison scores generated from fingerprint and gait recognition systems. Moreover, the authors show that fusion performance of these modalities gives better recognition performance than unimodal biometric system on mobile devices.

Vishi et al. [6], evaluated three normalization and four fusion techniques for multimodal biometric fusion using fingerprint and iris biometrics. The results shown in the paper indicate that the best performance of multimodal biometric system is achieved using hyperbolic tangent estimator (TanH) score normalization technique and simple sum fusion rule.

Peng et al. [7], used triangular norm technique to fuse four biometric characteristics, such as finger vein, fingerprint, finger shape and finger knuckle print from one finger of participants. Authors show that triangular norm fusion gives lower error rates and better performance than other fusion techniques.

Probabilistic and predicate logic based fusion techniques are not generally investigated in the multimodal biometric systems. Additionally, to the best of our knowledge subjective logic (SL) based fusion in biometrics is not investigated yet. It is worth mentioning that the Subjective Logic theory is invented by Prof. Dr. Audun Jøsang [8].

Conti et al. [9] and Lau et al. [10], proposed a multi-instance fusion that means fusing two comparison scores by using fuzzy logic rules from two different fingerprints. Authors show that score level fusion gives an increase on performance of 6.7% in comparison with unimodal biometric recognition systems.

Kisku et al. [11], have proposed a multimodal biometric system using face and ear biometrics through Gaussian Mixture Model (GMM) with belief fusion of the estimated scores characterized by Gabor responses and the proposed fusion method is accomplished by Dempster-Shafer (DS) decision theory. Results indicate that the technique provides an increase in accuracy and a significant improvement over the existing classical fusion rules.

3. Multibiometric Systems

Nowadays, biometric systems running in real world applications are primarily unimodal. These systems use a single biometric modality e.g., fingerprint with the purpose of identifying or verifying users. Unimodal biometrics are limited because none of the biometrics alone is considered robust enough to deal with hindrances caused by any external factor [12]. Some of the biggest concerns that unimodal systems deal with are as follows:

- **Noise**: made in the gained data due to differences in the biometric marker (e.g. surgically modification of the finger) [13]. Scars on the fingerprints and voice alteration due to variety of conditions are some of the examples of noise. The presence of noise on biometric data leads to false match or false non-match issues [12].

- **Intraclass variation**: that could happen when a user contact with the sensor or with transformations caused by a physiological factor that occurs with ageing [13].

- **Interclass similarities**: appears when a biometric database contains a large number of users, performing the need to increase the complexity to make differences between the users [13].

- **Non-universality**: The biometric system may not get precise biometric data from some of the users. *"The lack of universality is the primary reason for Failure-To-Enroll (FTE) situations"* [12].

- **Resistance to circumvention**: is the case when a user in the successful way masquerades as another person by falsifying biometric data taken from another person [13].

Some of the limitations of unimodal biometric are resolved by combining multiple biometric characteristics, multiple samples, and multiple algorithms to perform authentication. These systems are called multi-biometric systems and are expected to be more secure because of the integration of several independent pieces of evidence [14]. Systems that combine various biometric traits (modalities) are known as multimodal biometric systems, and their functionality is to overcome some shortcomings of unimodal biometrics mentioned above. For instance, the chance to get a valuable biometric increases with the number of involved biometric characteristics. They also prevent data spoofing because it is harder to spoof multiple biometric characteristics of real users [13].

3.1 Multiple Integration Strategies

There are six types of multi-biometric authentication systems based on the number of different modules or type of acquisition input. The definitions for these integration strategies are based on various sources [12, 15, 16] and are illustrated in Figure 1. These integration strategies or fusion scenarios are considered as follows:

1. **Multiple Sensors**: In these systems, multiple samples are produced for a single biometric trait to assuage noisy sensor data; therefore different sensors might be used to improve performance for the same biometric identifier such as fingerprinting (e.g., optical and capacitance sensors).

2. **Multiple Samples**: A single sensor might be used to capture the same biometric modality using the same instance (e.g., two attempts or templates of a person's right index finger vein).

3. **Single Biometric-Multiple Matchers**: In these systems, the same biometric trait is processed using different algorithms (approaches) for feature extraction and comparison module (e.g., minutiae vs non-minutiae based, like filter-based) matcher or algorithm.

4. **Multiple Biometric fusion**: usually referred as multimodal biometric system using different body traits of the same person (e.g., using finger vein and fingerprint

together, or gait and fingerprint, face, iris and fingerprint etc.)

5. **Multiple Instances**: These systems combine multiple instances from the same biometric trait (also knows as multi-unit or intramodal system in the literature) such as five fingers, ten fingers or both palms for vein recognition.

6. **"Soft Biometrics"**: These systems are rarely deployed, but worth noting here as *"soft biometrics"* are considered all those characteristics that have lack of the distinctiveness and permanence to identify an individual, on the other hand, they provide complementary information for primary biometric traits. Gender, age, eye colour, hair colour, height, weight, ethnicity etc., are some examples of "Soft biometrics".

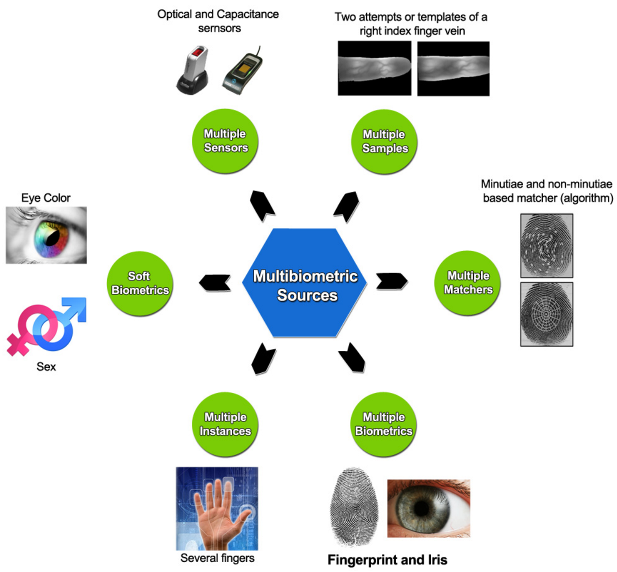

Figure 1. Types of multibiometric authentication systems

Businesses decide to apply a specific combination strategy based on their technical requirements, cost, and preferred biometric characteristic. These fusion strategies can be done successively (serial fusion) or simultaneously (parallel fusion). By applying multiple integration strategies, we improve biometric verification performance, achieve better decisions, and overcome many shortcomings of unimodal biometric systems.

Despite the advantages of multi-biometric systems, there are also some disadvantages, such as high cost, complexity and longer processing time.

3.2 Score Level Fusion

Score level (also known as opinion level) biometric fusion can be applied to both unimodal and multimodal biometric systems. Most commonly, scores from multiple samples of the same or of the different modalities are combined, in which case it is called score-level data fusion. However, it is also possible to combine other metrics such as ranks, in which case it is called rank-level data fusion.

In case of a unimodal system, multiple scores derived from multiple samples of the same modality are compared against the threshold in decision module. For example, multiple attempts or representations that are known to represent the same identity can be compared with the stored iris or fingerprint template of a given person. Each attempt will give a different score, so some form of a combination of each score can provide higher certainty about the likeness than each image separately.

In case of multimodal systems the scores are often derived from different scales, so scaling and normalisation are required.

The crucial aspect of score level data fusion is the weight that is to be given to the various scores, where specific methods can be weighted sum or average, weighted product, and decision trees. Intuitively, combining scores from the same modality is very different from combining scores from different modalities, this is discussed in more details below.

It can be argued that low-level biometric data fusion (e.g. signal or feature-level fusion) has the potential to offer higher accuracy than high-level biometric data fusion [17]. However, it is also important to consider the simplicity of implementation and performance issues. Hence, decision level fusion also can have many advantages, and score-level fusion might provide a good trade-off between potential accuracy, performance and ease of implementation [18].

A general observation of biometric analysis is that when the quality of the measure is questionable, then the derived score values and the decisions become uncertain. Therefore, we propose to perform biometric data fusion based on subjective logic which is a framework which explicitly can handle degrees of uncertainty with respect to biometric modalities.

4. Subjective Logic

In this section, we will first introduce the notation and formation of subjective opinions such as opinion representation used for biometric data fusion.

4.1 Opinion Representation

Let X be a random variable. A *subjective opinion* on X [19] is a tuple:

$$\omega_X = (\boldsymbol{b}_X, u_X, \boldsymbol{a}_X)\ , \tag{1}$$

where $\boldsymbol{b}_X : \mathbb{X} \to [0,1]$ is a *belief mass distribution* over X, $u_X \in [0,1]$ is an *uncertainty mass*, and $\boldsymbol{a}_X : \mathbb{X} \to [0,1]$ is a *base rate distribution* over X, satisfying the additivity constraints:[†]

$$u_X + \textstyle\sum_{x\in\mathbb{X}} \boldsymbol{b}_X(x) = 1 \qquad \text{and} \qquad \textstyle\sum_{x\in\mathbb{X}} \boldsymbol{a}_X(x) = 1\ . \tag{2}$$

[†]This definition is for a *multinomial subjective opinion*. In general, we can define *hyper opinions*, where $\boldsymbol{b}_X : \mathcal{R}(\mathbb{X}) = 2^{\mathbb{X}} \setminus \{\mathbb{X}, \emptyset\}$, and operate with them through their multinomial projections (see [8]).

The beliefs and the uncertainty mass are a result of a specific analysis of the random variable by applying expert knowledge, experiments, a personal judgment, etc. $\boldsymbol{b}_X(x)$ is the belief that X takes the value x expressed as a degree in $[0,1]$. It represents the amount of experimental or analytical evidence in favour of x. u_X is a scalar, representing the degree of uncertainty about the belief analysis. It represents lack of evidence that can be due to lack of knowledge or expertise, or insufficient experimental analysis. The base rate distribution $\boldsymbol{a}_X$ is the prior probability distribution of X that reflects domain knowledge relevant to the specific analysis, most usually relevant statistical information.

A multinomial opinion can be represented as a point inside a regular simplex. In particular, a trinomial opinion can be represented inside a tetrahedron, as shown in Figure 2. The beliefs are the distances to the sides, and the uncertainty mass is the distance to the base of the tetrahedron. The base rate distribution is represented as a point on the base.

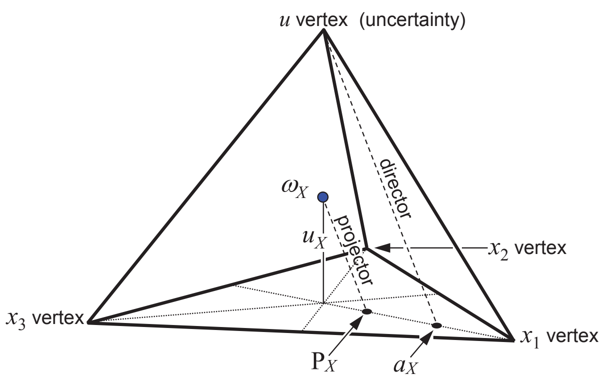

Figure 2. Visualisation of a trinomial opinion

A subjective opinion in which $u_X = 0$, i.e. an opinion without any uncertainty, is called a *dogmatic opinion*. Dogmatic opinions correspond to probability distributions. A dogmatic opinion for which $\boldsymbol{b}_X(x) = 1$, for some $x \in \mathbb{X}$, is called an *absolute opinion*. Absolute opinions correspond to observations. In contrast, an opinion for which $u_X = 1$ (and consequently $\boldsymbol{b}_X(x) = 0$, for every x) is called a *vacuous opinion*. For a given multinomial opinion ω_X we define its corresponding *projected probability distribution* $\mathbf{P}_X : \mathbb{X} \to [0,1]$ in the following way:

$$\mathbf{P}_X(x) = \boldsymbol{b}_X(x) + \boldsymbol{a}_X(x)\, u_X \,. \tag{3}$$

$\mathbf{P}_X(x)$ is an estimate for the probability of x which varies from the base rate value, in the case of complete ignorance ($u_X = 1$), to the belief in the case $u_X = 0$.

5. Determining Opinions from Biometric Sensors

In our approach, binomial opinions are used to represent biometric opinion scores. Each specific extracted feature from a biometric sample can be compared against the corresponding features stored in the database for a given person. Assume N features denoted by f_i where $i = 1 \dots N$. We assume that the degree of comparison of each feature f_i is given a rating $r(f_i)$ where $r(f_i) \in [0,1]$. This might require normalisation in case the underlying model for feature comparison provides ratings on a different scale. The more features analysed, the more ratings collected, and the smaller the uncertainty of the biometric score opinion becomes. The score opinion ω_X can be determined from the ratings $r(f_i)$ according to Eq.(4):

$$\omega_X \begin{cases} \boldsymbol{b} = \frac{\sum r(f_i)}{W+N} \\ \\ \boldsymbol{d} = \frac{\sum (1 - r(f_i))}{W+N} \\ \\ u = \frac{W}{W+N} \end{cases} \tag{4}$$

where $W = 2$ is the non-informative prior weight with default value dictated by the requirement of having a uniform PDF over binary frames in case of absence of evidence. The value would e.g. be $W = 3$ in case it were required to have a uniform probability density function (PDF) over a ternary frame. However, higher values for W make the probability distribution less sensitive to new evidence, so the value $W = 2$ is adopted [20].

Figure 3 illustrates the principle for deriving biometric opinion scores and using those scores for making decisions.

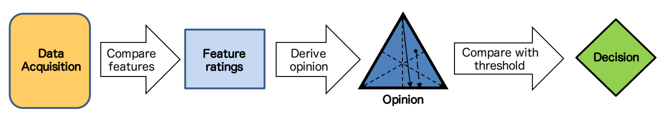

Figure 3. Deriving biometric opinion scores

An opinion derived according to Eq.(4) can thus represent a biometric score which can always be mapped to a probability value whenever required. The advantage of the opinion representation is that it directly takes into account the amount of evidence supporting the score.

6. The Proposed Biometric Data Fusion Based on Subjective Logic: *BDF-SL*

This section describes the proposed approach for multimodal biometric data fusion at opinion (score) level. Figure 4 illustrates a high-level design of our proposed fusion method, where subjective logic based fusion is combined with the statistical approach. In the proposed framework, the first step is to transform the comparison scores into subjective opinions using density estimation technique as described previously. In the next step, a subjective logic based operator known as *"cumulative fusion"* ($\oplus$ or $\diamond$) is used for comparison score fusion; and finally, a statistical likelihood ratio test is applied for classification. In this way, the properties of both statistical and subjective logic theories are combined for multi-biometric score fusion.

In upcoming sections, it is given an explanation of how subjective opinion fusion and cumulative fusion operator work, which is the main fusion component of our proposed method.

6.1 Subjective Opinion Fusion

Belief fusion involves belief arguments from multiple sources that must be combined in some way to produce a single belief argument. The purpose of opinion fusion is to create a new opinion that is more accurate or representative.

Different opinions can be fused in various ways, each having an impact on how the particular type of biometric fusion is modelled. It can be challenging to determine the correct or the most appropriate fusion operator for a given situation. This section describes cumulative fusion which is relevant for multi-biometric fusion cases.

6.2 Cumulative Fusion

Cumulative opinion fusion is when it is assumed that it is possible to collect an increasing amount of independent evidence by including more and more arguments, and that certainty about the correct state increases with the amount of evidence accumulated. A typical case depicting this type of fusion is when one makes statistical observations about possible outcomes, i.e. the more observations, the stronger the analyst's belief about the likelihood of each outcome.

For example, a mobile network operator could observe the location of a subscriber over time, which will produce increasing certainty about the most frequent locations of that subscriber. However, the result would not necessarily be suitable for indicating the exact location of the subscriber at a particular time.

The cumulative fusion rule is equivalent to *a posteriori* updating of Dirichlet distributions. Its derivation is based on the bijective mapping between the belief and evidence notations described in [21].

The symbol '$\diamond$' denotes the cumulative fusion of two observers A and B into a single imaginary observer $A \diamond B$.

Let $\mathbb{C} = \{C_1, C_2, \dots C_N\}$ be a frame of N sources with the respective opinions $\omega_X^{C_1}, \omega_X^{C_2}, \dots \omega_X^{C_N}$ over the same variable X. Let C denote a specific source $C \in \mathbb{C}$. The cumulative merger of all the sources in the source frame $\mathbb{C}$ is denoted $\diamond(\mathbb{C})$. The opinion $\omega_X^{\diamond(\mathbb{C})} \equiv \left(\boldsymbol{b}_X^{\diamond(\mathbb{C})},\ u_X^{\diamond(\mathbb{C})},\ \boldsymbol{a}_X^{\diamond(\mathbb{C})}\right)$ is the cumulative fused opinion expressed as:

Case I: $\exists\, u_X^C \neq 0$:

$$\begin{cases} \boldsymbol{b}_X^{\diamond(\mathbb{C})}(x) &= \dfrac{\sum\limits_{C\in\mathbb{C}} \left(\boldsymbol{b}_X^C(x) \prod\limits_{C_j \neq C} u_X^{C_j} \right)}{\sum\limits_{C\in\mathbb{C}} \left(\prod\limits_{C_j \neq C} u_X^{C_j} \right) - (N-1) \prod\limits_{C\in\mathbb{C}} u_X^C}, \\ u_X^{\diamond(\mathbb{C})} &= \dfrac{\prod\limits_{C\in\mathbb{C}} u_X^C}{\sum\limits_{C\in\mathbb{C}} \left(\prod\limits_{C_j \neq C} u_X^{C_j} \right) - (N-1) \prod\limits_{C\in\mathbb{C}} u_X^C}, \end{cases} \tag{5}$$

Case II: $u_X^C = 0, \quad \forall C \in \mathbb{C}$:

$$\begin{cases} \boldsymbol{b}_X^{\diamond(\mathbb{C})}(x) &= \sum\limits_{C\in\mathbb{C}} \gamma_X^C\, \boldsymbol{b}_X^C(x), \\ u_X^{\diamond(\mathbb{C})} &= 0 \end{cases} \tag{6}$$

$$\tag{7}$$

$$\text{where} \quad \gamma_X^C = \lim_{u_X^{\mathbb{C}} \to 0} \frac{u_X^C}{\sum\limits_{C_j \in \mathbb{C}} u_X^{C_j}}. \tag{8}$$

The cumulative fused opinion $\omega_X^{\diamond(\mathbb{C})}$ results from fusing the respective opinions ω_X^C of the sources $C \in \mathbb{C}$. The symbol '$\oplus$' denotes the cumulative belief fusion operator, hence we define

$$\omega_X^{\diamond(\mathbb{C})} \equiv \bigoplus_{C\in\mathbb{C}} (\omega_X^C) \tag{9}$$

$$\equiv \omega_X^{C_1} \oplus \omega_X^{C_2} \oplus \dots \omega_X^{C_N}. \tag{10}$$

It can be verified that the cumulative fusion operator is commutative, associative and non-idempotent. In Case II of Eq.(6), the associativity depends on preserving the relative weights of intermediate results with the additional weight parameter γ. In this case, the cumulative fusion operator is equivalent to the weighted average of probabilities.

The argument base rate distributions are normally equal. When that is not the case the fused base rate distribution over X is specified to be the evidence-weighted average base rate.

In case of N dogmatic arguments ω_X^C where $C \in \mathbb{C}$ it can be assumed that the limits in Eq.(6) are defined as $\gamma_X^C = 1/N$.

In our proposed fusion approach, we follow an information theoretic approach to analyse the significance of the correlation between biometric comparators in the design of a fusion scheme. Let $X \in \mathbb{X}$ denote the decision made by a biometric system, wherein biometrics $\mathbb{X} = \{genuine, impostor\}$. The

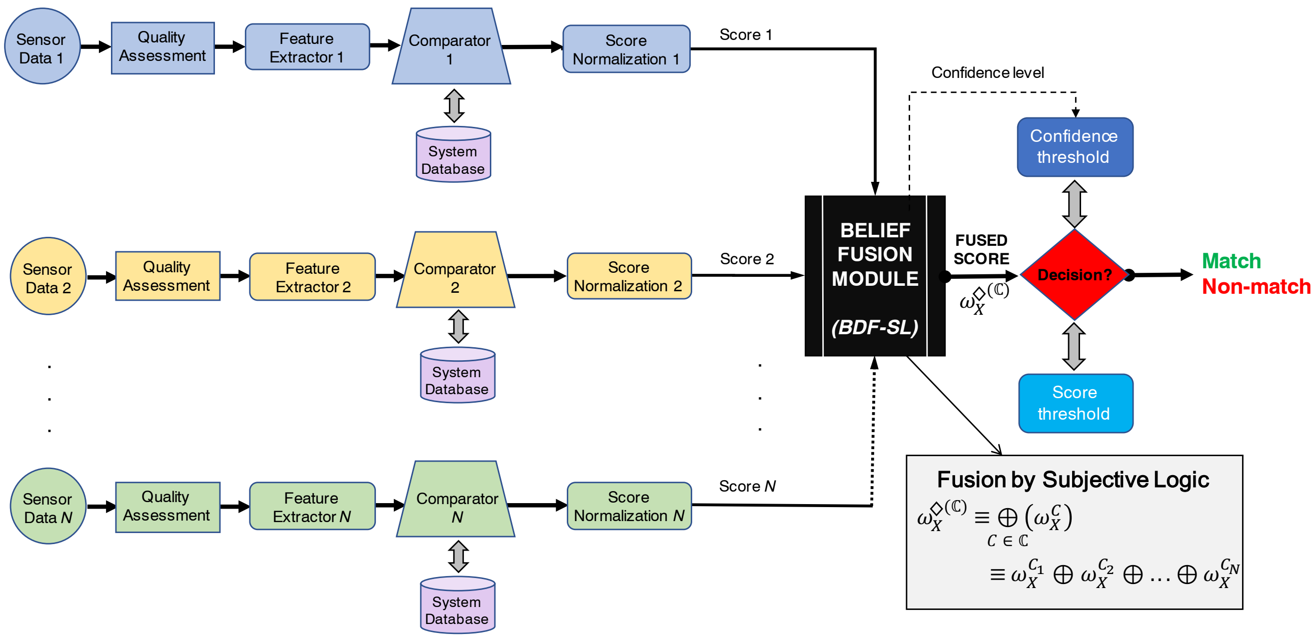

Figure 4. General scheme of the proposed biometric data fusion approach (BDF-SL)

entropy of X is a measure of the uncertainty of the random variable X, and in subjective logic it is defined as

$$H(X) = -\sum_{x \in \mathbb{X}} P_X(x) \log_2 (P_X(x)) \tag{11}$$

$P_X(genuine)$ and $P_X(impostor)$ are the prior probabilities of the genuine and impostor classes, respectively. The value of $H(X)$ is 1 bit when the prior probabilities of the two classes are equal. Let the biometric comparison scores be modeled as a multivariate random variable $C = [C_1, C_2, ..., C_n]$, where C_n corresponds to the comparison score of the n^{th} comparator. The mutual information between X and C is defined as

$$I(X;C) = H(C) - H(C|X) \tag{12}$$

where $H(C)$ is the entropy of C and $H(C|X)$ is conditional entropy of C given X. Mutual information can be viewed as reduction in the entropy (uncertainty) of X due to the knowledge of C.

7. Experiments and Data Analysis

There are several approaches to study multimodal biometric fusion. One approach is to use heterogeneous databases, explained by Ross et al. [22], i.e., combine biometric trait (e.g. fingerprint) from a database with the biometric characteristic (e.g. iris) from another database. From the experiment point of view, these combined biometric traits belong to the same person, and the resultant person is called as *chimeric user* or *virtual user* (please refer to section 7.3). Although this approach has been widely used in multimodal literature, it was questioned that whether this approach was reasonable during the 2003 Workshop on Multimodal User Authentication.

Poh et al. [23] studied this problem and showed that the performance measured with experiments carried out on chimeric users does not necessarily reflect the performance with real multimodal users. Obviously, the best way to study biometrics fusion is to use homologous multimodal biometric databases, which means that different biometric traits are indeed coming from the same real person. However, there are only a few multimodal biometric databases publicly available, and most of the existing multimodal databases are composed of two modalities.

The existing databases have several limitations, e.g., lack of import traits or lack of diversity of sensors or modalities. Therefore, for our experiments, we have chosen two fingerprint and two iris databases described below.

All fingerprint and iris recognition experiments are performed using MegaMatcher and Microsoft Visual Studio 2017 (C#). MegaMatcher is a commercial biometric feature extractor and comparator developed by Neurotechnology, as well as it is NIST MINEX-compliant fingerprint engine (VeriFinger), NIST IREX-compliant iris engine (VeriEye).

MegaMatcher ABIS 10.0 supports the following biometric standards: ANSI/ NIST-ITL-1, ISO/IEC 19794-2, ISO/IEC 19794-5, ISO/IEC 19794-6 biometric template standards and ICAO requirements [24].

7.1 Databases

Fingerprint and Iris experiments in this work are performed over four different data- bases (DB) collected from two separate institutions.

1. Fingerprint databases and an iris database are collected by Machine Learning and Applications (MLA) Group at Shandong University in China (SDUMLA-HMT) [25].

2. Another iris database is collected by Institute of Automation, Chinese Acade-my of Sciences (CASIA-Iris-Lamp) [26].

The fingerprint images on SDUMLA-HMT database are collected with five different sensors and is called a multi-sensor database. Furthermore, fingerprint images in SDUMLA-HMT database are acquired from six fingers of the same person such as *thumb finger, index finger and middle finger*, of both hands. It is worth mentioning that MLA Group has requested from participants eight impressions (attempts) for each of six fingers to five aforementioned sensors. Some of the fingerprint images are shown in Figure 5. Fingerprint database (DB) consists of *6(fingers) x 5(sensors) x 8(attempts) x 106(subjects) = 25,440 fingerprint images*.

We consider all images from each sensor as a database (DB), and we have checked the quality of images from five different sensors, for each of these databases. Then, we have selected the best quality and the worst quality databases among these which are DB2 from the FPR620 sensor and DB3 from FT-2BU sensor, see Figure 5.

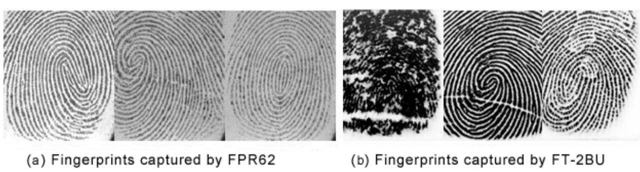

Figure 5. Fingerprint image samples from a) DB2 (best db) and b) DB3 (worst db)

It is to be noted that we have reduced the number of participants, fingers and impressions as following: 2 fingers, in particular, index fingers of both hands from 100 members out of 106, and from 8 samples we have used only first 5 impressions for finger in order to correlate fusion with *2 irides* and *5 iris attempts* (*1000 iris images*). After this modification we do have *2(fingers) x 100(subjects) x 5(attempts) = 1000 images* per database in total *2000 fingerprint images* from two databases (DB2 and DB3) out of 25,440 fingerprint images.

In our fingerprint experiment, we have assigned DB2 as FP-DB1 (for the best quality database) and DB3 as FP-DB2 (for the worst quality database).

Iris experiments in this work are conducted in two different databases (DB) collected from various institutions as shown in Figure 6.

1. First iris database (named as Iris-DB1) is collected by Institute of Automation, Chinese Academy of Sciences (CASIA-Iris-Lamp).

2. Second iris database (named as Iris-DB2) collected by Machine Learning and Applications (MLA) Group at Shandong University in China (SDUMLA-HMT Iris).

In our iris experiments, we have reduced the number of participants (images) as well to have the same number of samples to perform fusion with fingerprint databases.

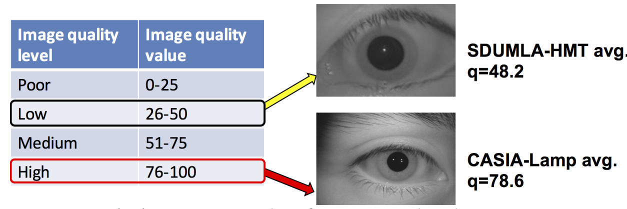

Figure 6. Iris image samples from two databases: SDUMLA-HMT Iris and CASIA-Iris-Lamp

- From SDUMLA-HMT iris database we have chosen only 100 subjects out of 106 in total 1000 iris images: *2(irises) x 5(irisimages) x 100(subjects) = 1000 iris images*

- From CASIA-Iris-Lamp we have used also only 100 subjects out of 411 and it is to be noted that we have reduced the number of attempts from 20 to 5, in order to comply with 1000 images in total: *2(eyes) x 5(iris images) x 100(subjects) = 1000 iris images*

7.2 Fusion Experiments

Based on the quality assessment results, datasets with the best quality of images are named with suffix 1, FP-DB1 for fingerprint and Iris-DB1, respectively. Futhermore, datasets with the worst quality of images are named with the suffix 2, FP-DB2 for fingerprint and Iris-DB2, respectively.

In this paper, we have formulated four fusion strategies to evaluate our proposed fusion framework, as follows:

- **Strategy #1: FP-DB1 fused with Iris-DB1**
- **Strategy #2: FP-DB1 fused with Iris-DB2**
- **Strategy #3: FP-DB2 fused with Iris-DB1**
- **Strategy #4: FP-DB2 fused with Iris-DB2**

7.3 Real vs. Virtual Users

For fusion scenarios 1 and 3 we have used heterogeneous databases for fingerprint and iris, thus we have created so called *"virtual users"*, while for fusion scenarios 2 and 4 we have used modalities (fingerprint and iris) from homogeneous databases or *"real users"* [27]. The methodology of real and virtual users is illustrated in figure 7.

A *"real user"* denoted as a subject (R) has provided both required modalities to the database, iris and fingerprint, this case is for SDUMLA-HMT database (same subjects for both iris and fingerprint). While a participant V, donated with only one biometric modality, either iris data (VEye) or fingerprint data (VFinger), therefore these traits are combined from different users to create another user [27]. In this case, we have combined iris modality from CASIA-Iris-Lamp and fingerprint modality from SDUMLA-HMT fingerprint database, and thus we have created so called a *"virtual user"*.

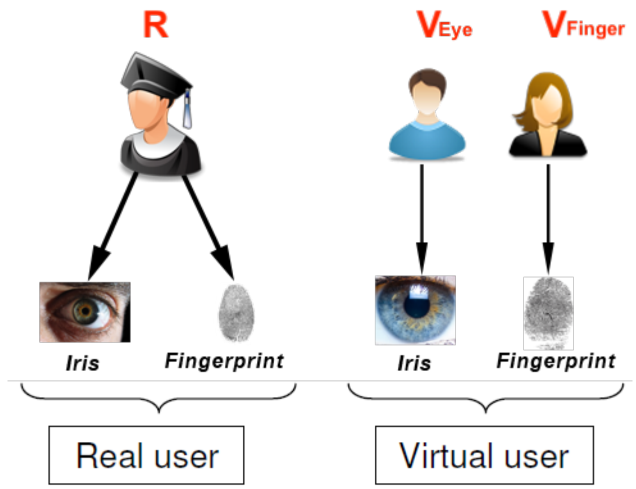

Figure 7. Methodology of real and virtual users

8. Experimental Results

We have retrieved both low and high Equal Error Rates (EER) from unimodal biometric systems (fingerprint and iris recognition). Unimodal biometric systems are often affected by several practical problems like noisy sensor data, unacceptable error rates, spoof attacks etc. Multimodal biometric systems overcome some of the aforementioned problems.

As shown in table 1, Hyperbolic Tangent score normalization technique gives better performance than two other normalization techniques, MinMax and Z-Score. Additionally, our proposed fusion method (BDF-SL) provides better results compared with other classical fusion approaches. Specifically, subjective logic fusion approach gives an average of 99.98% recognition accuracy, making it the most stable fusion framework.

Table 1. Some of comparison results for normalization and fusion techniques

EER (in %)			
	Normalization technique		
Fusion technique	MinMax	Z-Score	HypTan
Minimum score (MinS)	3.81836	7.32851	1.01881
Maximum score (MaxS)	0.11591	0.11206	0.07837
Simple sum (SS)	0.08949	0.15935	0.01763
Our Approach (BDF-SL)	0.08281	0.10955	**0.00011**

8.1 Biometric Fusion Results

In this section, we provide a comparison of performances for unimodal and mutimodal biometrics. In total we have generated 48 DET graphs for all our fusion strategies mentioned previously. However, here we show only one DET graph per fusion strategy (normalisation and fusion approach).

8.1.1 [Strategy #1:] *FP_DB1 fused with Iris_DB1*

This case shows a comparison of unimodal and multimodal databases of good quality images of iris *Iris_DB1* and fingerprint *FP_DB1*. The results show that the fingerprints dataset attains an EER of 0.86%, and iris dataset an EER 0.71%. The best fusion performance is achieved using Hyperbolic Tangent (TanH) and Subjective Logic fusion (proposed approach) with an EER of *0.00010%*. The results are visualized in form of DET curves in Figure 8.

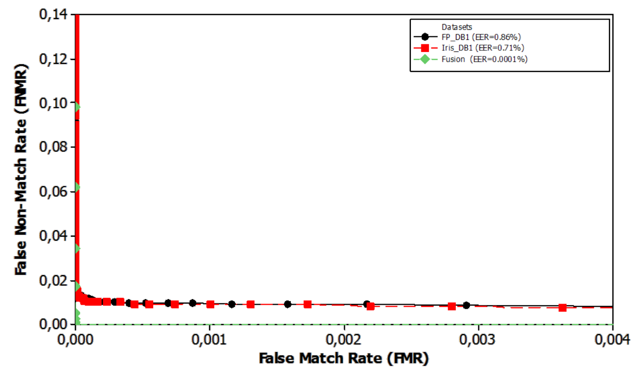

Figure 8. Fusion performance of good quality datasets *(FP_DB1 and Iris_DB1)*

8.1.2 [Strategy #2:] *FP_DB1 fused with Iris_DB2*

This case shows a comparison of unimodal and multimodal databases between good quality images of fingerprint *FP_DB1* and bad quality images of iris *Iris_DB2*. The results show that the fingerprints dataset attains an EER of 0.86%, and iris dataset an EER 7.35%. The best fusion performance is achieved using Hyperbolic Tangent (TanH) and Subjective Logic fusion (proposed approach) with an EER of *0.0320%*. The results are visualized in form of DET curves in Figure 9.

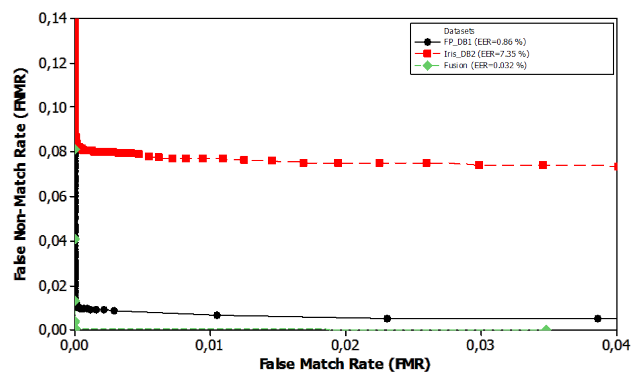

Figure 9. Fusion performance of good quality datasets *(FP_DB1)* and bad quality dataset *(Iris_DB1)*

8.1.3 [Strategy #3:] *FP_DB2 fused with Iris_DB1*

This case shows a comparison of unimodal and multimodal databases between bad quality images of fingerprint *FP_DB2* and good quality images of iris *Iris_DB1*. The results show that the fingerprints dataset attains an EER of 1.01%, and iris dataset an EER 0.71%. The best fusion performance is achieved using Hyperbolic Tangent (TanH) and Subjective Logic fusion (proposed approach) with an EER of *0.00015%*. The results are visualized in form of DET curves in Figure 10.

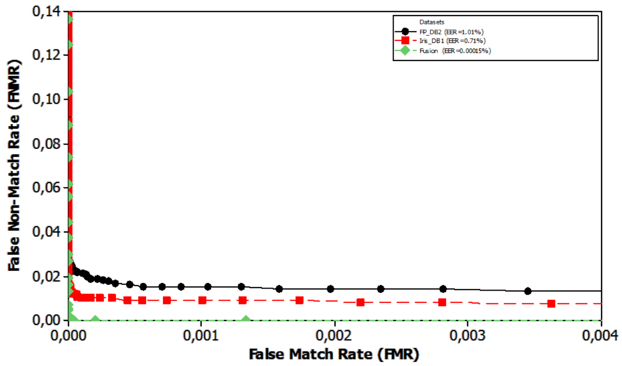

Figure 10. Fusion performance of bad quality datasets *(FP_DB2)* and good quality dataset *(Iris_DB1)*

8.1.4 [Strategy #4:] *FP_DB2 fused with Iris_DB2*

This case shows a comparison of unimodal and multimodal databases between bad quality images of both datasets, fingerprint *FP_DB2* and iris *Iris_DB2*. The results show that the fingerprints dataset attains an EER of 1.01%, and iris dataset an EER 7.35%. The best fusion performance is achieved using Hyperbolic Tangent (TanH) and Subjective Logic fusion (proposed approach) with an EER of *0.0038%*. The results are visualized in form of DET curves in Figure 11.

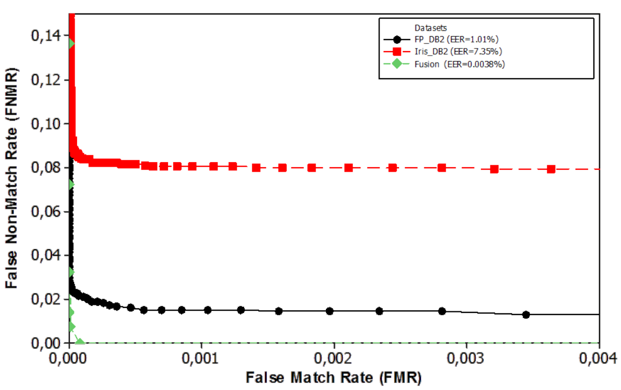

Figure 11. Fusion performance of bad quality datasets *(FP_DB2 and Iris_DB1)*

To summarize, the results indicate that fusion of fingerprint and iris improves the performance of biometric systems. We can define improvement as the percentage difference between the minimum EER of fingerprint or iris (we choose the lowest value between the two) minus the EER of their fusion and (divided by) the minimum EER of fingerprint or iris (we choose the lowest value between the two).

Table 2 summarizes the performances shown in this section with their improvements.

9. Discussion and Future Work

Experimental results show that in most cases subjective logic fusion performs better than classical fusion methods. However, we found out that there are few cases when subjective logic is not able to yield correct results, and sometimes decisions are not accurate because cumulative fusion generates non-specific belief opinions which reduce the performance.

Nevertheless, our future work will include experiments and tests with different modalities such as the face, voice, finger vein, fingerprint and iris and with a variety of databases, to assess the generalisation of our biometric data fusion based on subjective logic (BDF-SL).

Additionally, another research topic to be considered in our future work is an investigation of our proposed fusion approach using fingerprint and iris biometrics at feature extraction level or template level to secure biometric templates and enhance privacy of the users by hiding the meaning of extracted features (points) from iris and fingerprint in the stored template.

Finger	Iris	Fusion	Improvement
0.86 %	0.71 %	**0.0001** %	**99.98 %** (Our approach)
0.86 %	7.35 %	3.81 %	48.10 %
1.01 %	0.71 %	0.1872 %	80.50 %
1.01 %	7.35 %	0.0038 %	78.62 %

Table 2. Multimodal fusion improvements of fingerprint and iris recognition (values in percentage indicate EER)

10. Conclusion

An aspect of biometric measurements beeing often ignored is that the scores are derived based on varying degrees of the evidence which naturally leads to uncertainty about the score values. The advantage of subjective logic is that levels of confidence can be taken explicitly into account. It is therefore natural to apply subjective logic when deriving biometric scores and also for biometric data fusion.

From fingerprint and iris recognition experiments we have achieved both high and low performance based on the quality of datasets (images). The results show that the fingerprint dataset collected by the optical sensor (FPR620) attains an EER of 0.86%, while dataset collected by the capacitive sensor (FT-2BU) attains an EER of 1.01%. Moreover, iris dataset (SDUMLA-HMT) collected by the MLA group gave an EER of 7.35%, whereas CASIA-Iris-Lamp iris dataset collected by the OKI sensor gave an EER of 0.71%.

Fusion results show that the combination of fingerprint and iris biometrics using hyperbolic tangent score normalisation technique and subjective logic fusion approach gives an EER of 0.00011%. Specifically, we have achieved an improvement of 99.98% verification accuracy compared to other classical fusion methods. Furthermore, we proofed that quality of images (aqcuisition unit) has an important role in biometric systems; and recognition performance of the multimodal biometric system is better than unimodal biometrics.

Acknowledgments

This research is supported by the Research Council of Norway under Grant No.:IKTPLUSS 248030/O70 - SWAN Project. This research is also part of the SecurityLab of University of Oslo.

References

[1] ISO Standards. *ISO/IEC JTC 1/SC 37 Biometrics: SC 37 Standing Document 11 (SD 11), Part 1 Harmonization Document*.

[2] Anil K. Jain and Arun Ross. Multibiometric Systems. *Communications of the ACM*, 47(1):34, jan 2004.

[3] Anil Jain, Karthik Nandakumar, and Arun Ross. Score Normalization in Multimodal Biometric Systems. *Pattern Recognition*, 38(12):2270–2285, dec 2005.

[4] R Raghavendra, A Rao, and G Hemantha Kumar. A Novel Approach for Multimodal Biometric Score Fusion Using Gaussian Mixture Model and Monte Carlo Method. *International Conference on Advances in Recent Technologies in Communication and Computing*, 2009.

[5] Mohammad Derawi, Davrondzhon Gafurov, and Rasmus Larsen. Fusion of Gait and Fingerprint For User Authentication on Mobile Devices. *(IWSCN), 2010 2nd*, 2010.

[6] K. Vishi and S.Y. Yayilgan. Multimodal Biometric Authentication Using Fingerprint and Iris Recognition in Identity Management. In *Proceedings of the 9th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, pages 334–341, 2013.

[7] Jialiang Peng, Ahmed A. Abd El-Latif, Qiong Li, and Xiamu Niu. Multimodal Biometric Authentication Based on Score Level Fusion of Finger Biometrics. *Optik - International Journal for Light and Electron Optics*, 125(23):6891–6897, dec 2014.

[8] Audun Jøsang. *Subjective Logic: A Formalism For Reasoning Under Uncertainty*. Springer, Heidelberg, 2016.

[9] Vincenzo Conti, Giovanni Milici, Patrizia Ribino, Filippo Sorbello, and Salvatore Vitabile. *Fuzzy Fusion in Multimodal Biometric Systems*, pages 108–115. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[10] CW Lau, Bin Ma, HML Meng, YS Moon, and Yeung Yam. Fuzzy Logic Decision Fusion in a Multimodal Biometric System. *Interspeech*, 2004.

[11] Dakshina Ranjan Kisku, P Gupta, H Mehrotra, and J Sing. Multimodal Belief Fusion for Face and Ear Biometrics. *Intelligent Information Management*, 01(December):166–171, 2009.

[12] Arun A Ross, Karthik Nandakumar, and Anil K Jain. *Handbook of Multibiometrics*. Number ISBN-13: 978-0-387-22296-7. Springer-Verlag, 1st edition, 2006.

[13] B. Bhanu and V. Govindaraju. *Multibiometrics for Human Identification*. Cambridge University Press, 2011.

[14] L. I. Kuncheva, C. J. Whitaker, C. A. Shipp, and R. P. W. Duin. Is independence good for combining classifiers? In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, volume 2, pages 168–171, 2000.

[15] Davide Maltoni, Dario Maio, Anil K A.K. Jain, and Salil Prabhakar. *Handbook of Fingerprint Recognition*. Number ISBN: 978-1-84882-253-5. Springer-Verlag, 2nd edition, 2009.

[16] Anil K Jain, R. Bolle, S. Pankanti, Arun A Ross, and Karthik Nandakumar. *Introduction to Biometrics*. Springer, 2011.

[17] A.A. Ross, K. Nandakumar, and A.K. Jain. *Handbook of Multibiometrics*. Springer Science and Business Media, 2006.

[18] N. Morizet and J. Gilles. A New Adaptive Combination Approach to Score Level Fusion for Face and Iris Biometrics Combining Wavelets and Statistical Moments. In G. Bebis, editor, *Proceedings of the 4th International Symposium on Visual Computing (Advances in Visual Computing)*, volume 5359 of *LNCS*, pages 661–671, 2008.

[19] A. Jøsang. Conditional Reasoning with Subjective Logic. *Journal of Multiple-Valued Logic and Soft Computing*, 15(1):5–38, 2008.

[20] P. Walley. Inferences from Multinomial Data: Learning about a Bag of Marbles. *Journal of the Royal Statistical Society*, 58(1):3–57, 1996.

[21] Audun Jøsang and Robin Hankin. Interpretation and Fusion of Hyper-Opinions in Subjective Logic. In *Proceedings of the 15th International Conference on Information Fusion (FUSION 2012)*. IEEE, Los Alamitos, Singapore, July 2012.

[22] Arun Ross and Anil Jain. Information Fusion in Biometrics. *Pattern Recognition Letters*, 24:2115–2125, 2003.

[23] Norman Poh and Samy Bengio. Can Chimeric Persons Be Used in Multimodal Biometric Authentication Experiments? pages 87–100. Springer, Berlin, Heidelberg, 2006.

[24] Neurotechnology. MegaMatcher Automated Biometric Identification System for national-scale projects, 2017.

[25] Yilong Yin, Lili Liu, and Xiwei Sun. SDUMLA-HMT: A Multimodal Biometric Database. *Biometric Recognition*, pages 260–268, 2011.

[26] Chinese Academy of Sciences. CASIA-Lamp Image Database V4.0. Technical report, 2009.

[27] Franziska Wolf, Tobias Scheidat, and Claus Vielhauer. Study of Applicability of Virtual Users in Evaluating Multimodal Biometrics. pages 554–561. 2006.