A DSmT Based System for Writer-Independent Handwritten Signature Verification

Nassim ABBAS[1], Youcef CHIBANI[1], Bilal HADJADJI[1], Zayen AZZOUZ OMAR[1]

nabbas@usthb.dz, ychibani@usthb.dz, bhadjadji@usthb.dz, azzouz-omar.zayen.1@ens.etsmtl.ca

&

Florentin SMARANDACHE[2]

smarand@unm.edu

[1]Communicating and Intelligent System Engineering Laboratory
Faculty of Electronic and Computer Science
University of Sciences and Technology "Houari Boumediene"
Algiers, Algeria

[2]Department of Mathematics
University of New Mexico
Gallup, NM 87301, U.S.A.

ICIF 2016

Outline

- **Introduction**
- **One Class Support Vector Machines Classifier**
- **Belief Function Theories**
- **Proposed Combination Scheme for Handwritten Signature Verification**
- **Conclusion and future works**

Statement of the problem

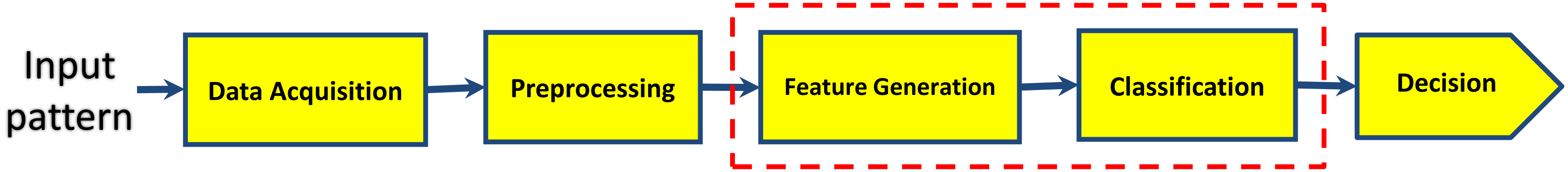


Figure 1. Structure of the recognition system

Solution

Parallel combination scheme

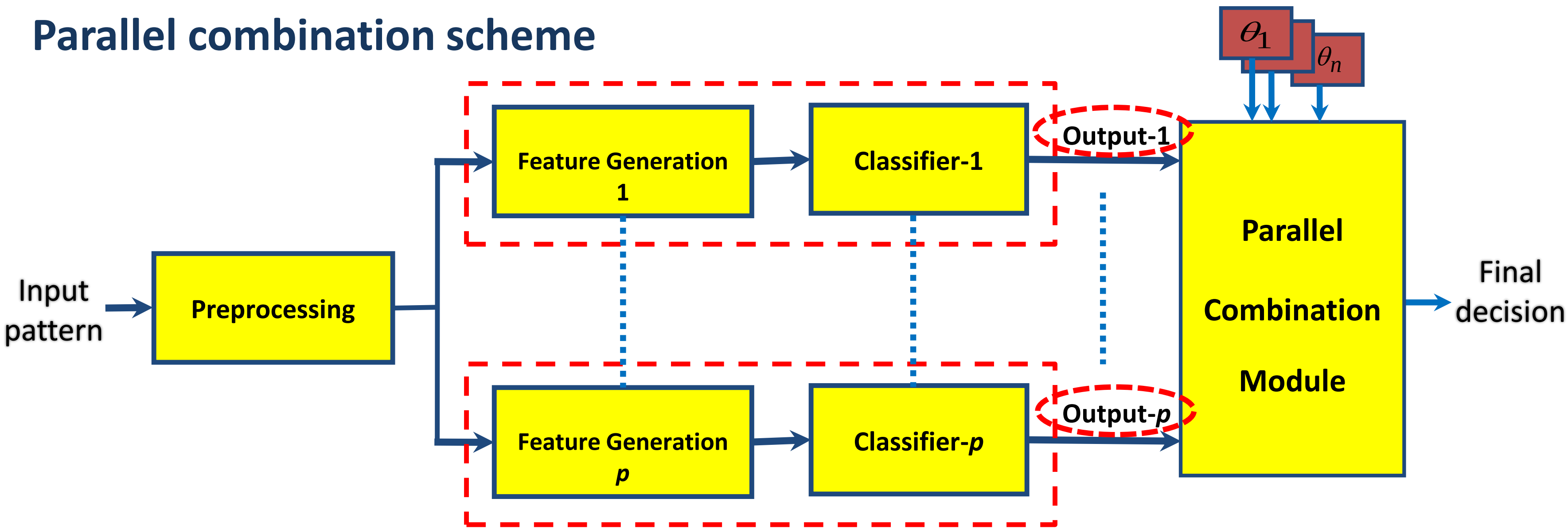


Figure 2. Parallel combination of classifiers

1. Introduction (2)

Combination levels

- Class level combination
- Rank level combination
- Measure level combination
  - ✓ Distance
  - ✓ Posterior probability
  - ✓ Confidence value
  - ✓ Match score
  - ✓ Belief function
  - ✓ Credibility
  - ✓ Possibility
  - ✓ Fuzzy measure
  - ✓ …

Belief functions take into considerations two notions:

- **Uncertainty:** *is an unrealistic measure induced by the outputs of classifier, which leads to interpret the response of the classifier as the result of a random phenomenon*
- **Imprecision:** *is measure representing the uncertainty linked to incomplete knowledge*

- **Three theories dealing with uncertainty and imprecise information have been introduced**
  - Probability theory (PT): *uncertainty*
  - Evidence theory (Dempster-Shafer Theory): *uncertainty + imprecision*
  - Plausible and paradoxical reasoning theory (Dezert-Smarandache Theory): *uncertainty + imprecision*

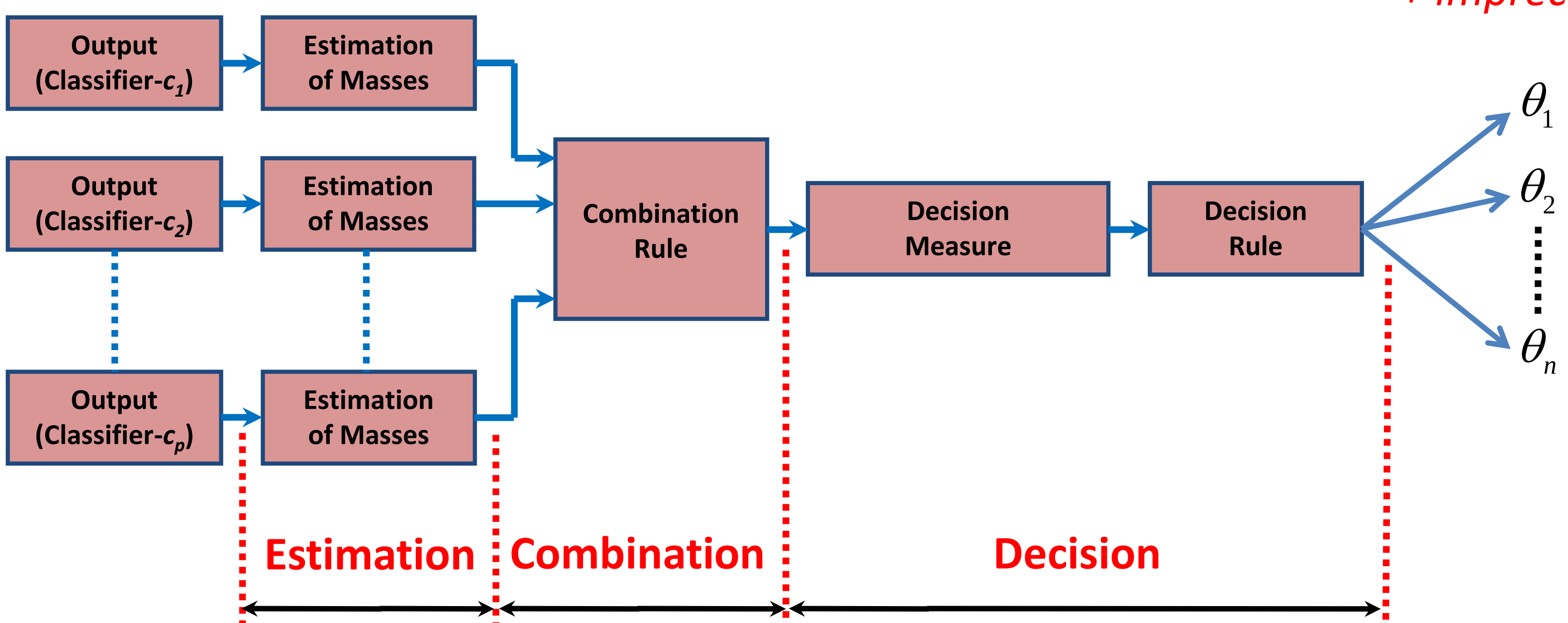
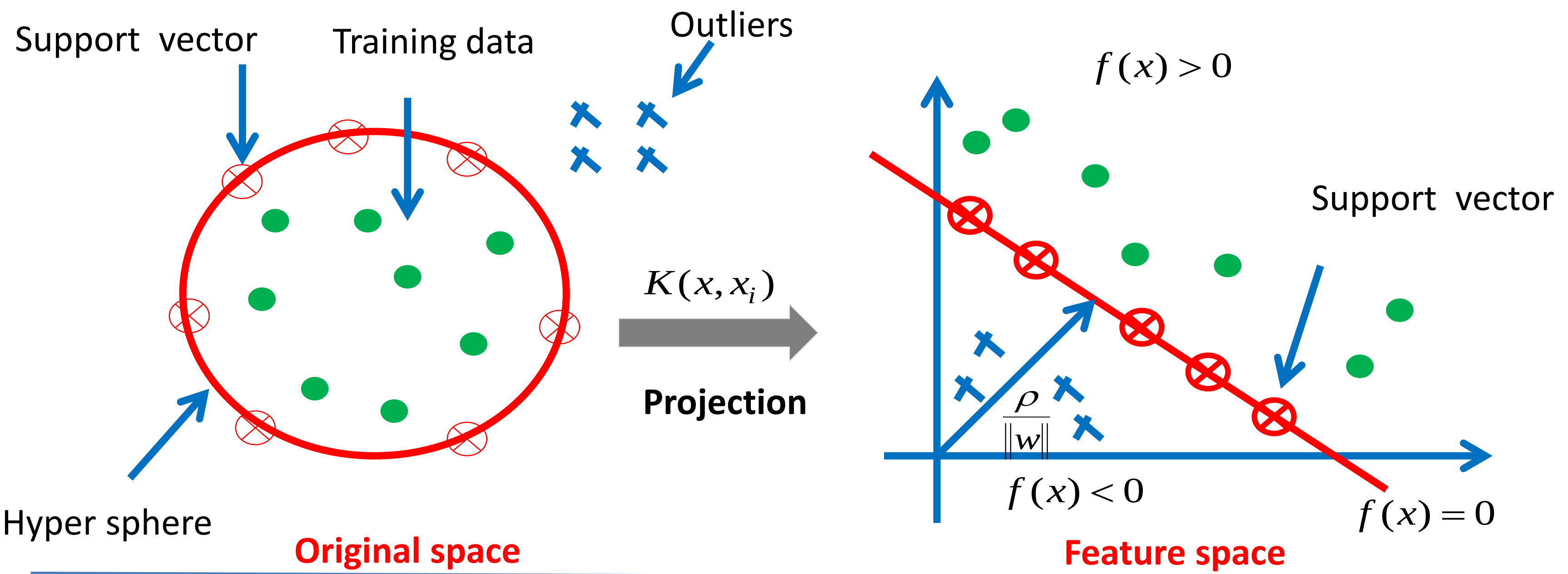


Figure 3. Belief Function Theories-Based Parallel Combination of Classifiers

One Class Support Vector Machines (OC-SVM)



Decision function:

$$f(x) = \sum_{j=1}^{Sv} \alpha_j K(x, x_j) - \rho$$

- Sv : Number of Support Vectors
- α_j : Lagrangian multipliers
- ρ : Distance of the hyperplane from the origin

Pattern ***x*** is accepted when $f(x) \geq 0$ Otherwise it is rejected

Mathematical Formalism

- **Discernment space:** is defined as a finite set of exhaustive and mutually exclusive hypotheses

$$G = \Theta = \{\theta_1, \theta_2 .., \theta_n\}$$

- **Basic probability assignment (bpa):**

$$\begin{cases} m \equiv P : \Theta \to [0,1] \\ \qquad \theta_i \mapsto m(\theta_i) \end{cases} \qquad \begin{cases} m(\varnothing) = 0 \\ \sum_{\theta_i \in \Theta} m(\theta_i) = 1,\ m(\theta_i) \geq 0 \end{cases}$$

- **Bayesian rule:**

$$P(\theta_i / x) = \frac{P(x/\theta_i).P(\theta_i)}{\sum_{i=1}^{n} P(x/\theta_i).P(\theta_i)}$$

Basic Sum combination rule

$$m_c(A) = m_{sum}(A) = \begin{cases} \frac{1}{p}\sum_{i=1}^{p} m_i(\theta_j) & \text{if } A = \theta_j, \\ 0 & \text{otherwise.} \end{cases}$$

A : Simple class belonging to discernment space Θ

p : Number of information sources

$m_i(.)$: bpa issued from the *i*-th source

- Advantage: Simple
- Limitation: No managing conflict between two sources

3. Belief Function Theories: *Dempster-Shafer Theory (1)*

Evidence Theory

- Dempster-Shafer theory (DST) allows to model both ignorance and imprecision, and to consider compound hypotheses such as the union of classes.
- It is generally recognized as a convenient and flexible alternative to the bayesian theory.

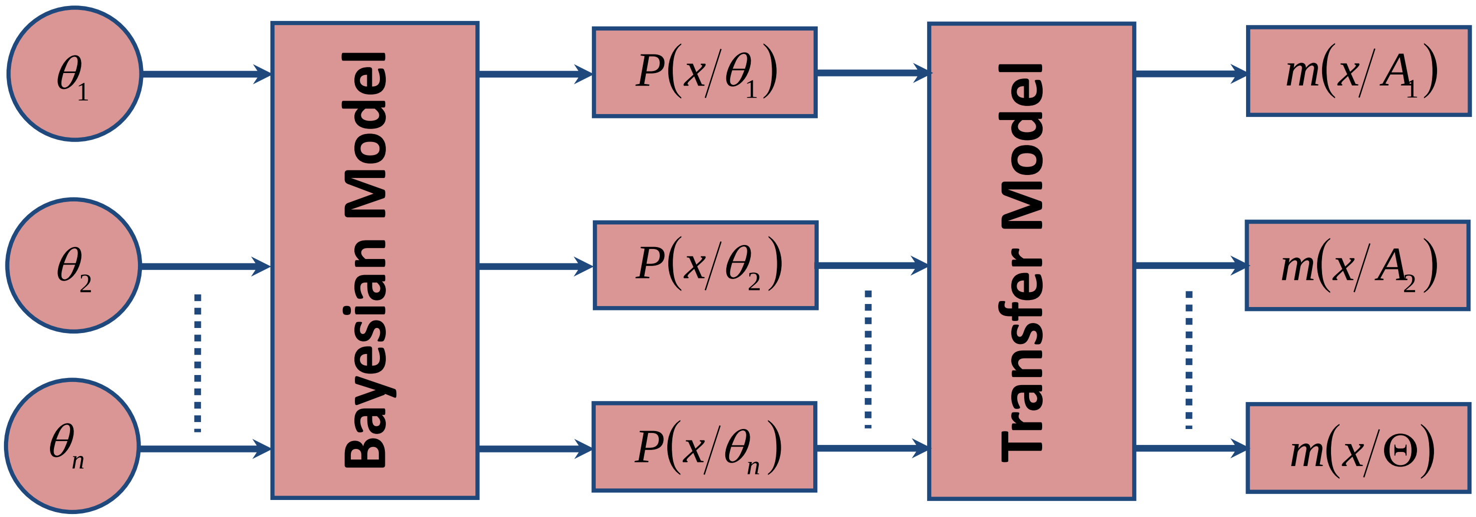
Mathematical Formalism

- **Discernment space:** $\Theta = \{\theta_1, \theta_2 .., \theta_n\}$
- **Power-set:** $G = 2^\Theta = \{\emptyset, \theta_1, \theta_2, \ldots, \theta_n, \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \ldots, \theta_1 \cup \theta_2 \cup \ldots \cup \theta_{n-1}, \Theta\}$
- **Basic belief assignment (bba):**

$$\begin{cases} m : 2^\Theta \to [0,1] \\ A_i \mapsto m(A_i) \end{cases} \qquad \begin{cases} m(\emptyset) = 0 \\ \sum_{A_i \in 2^\Theta} m(A_i) = 1,\ m(A_i) \geq 0 \end{cases}$$

Estimation of belief mass functions

- It's not directly explicit in term of modelling of the problem under consideration.
- It's specific to each application area according the nature of the data.
- Handwriting recognition.



Dissonant model of Appriou

- **Axiom (1):** Consistency with the Bayesian approach
- **Axiom (2):** Separability of the evaluation of the hypotheses
- **Axiom (3):** Consistency with the probabilistic association of sources

$$m_j(\theta_i) = \frac{\alpha_i . R_j . P(x/\theta_i)}{1 + R_j . P(x/\theta_i)},$$

$$m_j(\overline{\theta_i}) = \frac{\alpha_i}{1 + R_j . P(x/\theta_i)},$$

$$m_j(\Theta) = 1 - \alpha_i \ ,$$

$$R_j \in \left[0, \left\{ \max_{i \in [1, n]} [\sup(P(x/\theta_i))] \right\} \right]$$

$P(x/\theta_i)$: Conditional probability of an object x given the class θ_i .

R_j : Normalization factor

α_i : Coarsening factor.

Dempster's orthogonal sum rule

$$m_{\wedge}(A) = \sum_{\substack{B_1, B_2, \ldots, B_p \in 2^{\Theta} \\ B_1 \cap B_2 \cap \ldots \cap B_p = A}} \prod_{k=1}^{p} m_k(B_k),\ \forall A \in 2^{\Theta}$$

$$m_c(A) = m_{DS}(A) = \frac{m_{\wedge}(A)}{1 - K_c},\ \forall A \in 2^{\Theta} - \{\varnothing\}$$

$$K_c = m_{\wedge}(\varnothing) = \sum_{\substack{B_1, B_2, \ldots, B_p \in 2^{\Theta} \\ B_1 \cap B_2 \cap \ldots \cap B_p = \varnothing}} \prod_{k=1}^{p} m_k(B_k)$$

A : Focal element of the power-set 2^{Θ} .

$m_{\wedge}(A)$: Combined mass of Dempster's conjunctive rule.

K_c : Conflict measure between the different masses $m_k(.)$ issued from information sources S_k , respectively.

- Advantage: Taking into account the imprecise and uncertain information
- Limitation: No managing high conflict between two sources of information

3. Belief Function Theories: ***Dempster-Shafer Theory (5)***

Decision rules

- Combined mass function : uncertainty
- [Belief function, Plausibility function] : imprecision

Selecting the more realistic hypothesis.

Rules used for decision-making:

- Maximum of belief function.
- Maximum of plausibility function.
- Maximum of Pignistic Probability.
- Minimization of mass function with an acceptance threshold.

3. Belief Function Theories: *Dempster-Shafer Theory (6)*

Limitations of DST

- **Foundation of the DST** ⟹ **Does not take into account the paradoxical information**

- **Significant conflict measure** ⟹ **DST based combination is not possible**

▶ Solution: Dezert-Smarandache Theory (DSmT)

3. Belief Function Theories: ***Dezert-Smarandache Theory (1)***

Plausible and Paradoxical Reasoning Theory

- It has been originally developed since 2003 by Jean Dezert and Florentin Smarandache.
- It has the advantage of being able to represent explicitly the uncertainty from imprecise knowledge.
- It was elaborated for dealing with paradoxical sources of information (i.e. classes, descriptors, classifiers, sensors,...etc).
- It is based on a particular framework where the finite discrete frame of discernment is exhaustive but not necessarily exclusive.

3. Belief Function Theories: ***Dezert-Smarandache Theory (2)***

Mathematical Formalism

- **Discernment space:** $\Theta = \{\theta_1, \theta_2 .., \theta_n\}$

- **Hyperpower-set:** $G = D^{\Theta}$
  1. $\emptyset, \theta_1, \ldots, \theta_n \in D^{\Theta}$.
  2. If $A, B \in D^{\Theta}$, then $A \cap B \in D^{\Theta}$ and $A \cup B \in D^{\Theta}$.
  3. No other elements belong to D^{Θ}, except those obtained by using rules 1 or 2.

- **Generalized belief assignment (gbba):**

$$\begin{cases} m : D^{\Theta} \to [0,1] \\ \quad A_i \mapsto m(A_i) \end{cases} \qquad \begin{cases} m(\emptyset) = 0 \\ \sum_{A_i \in D^{\Theta}} m(A_i) = 1,\ m(A_i) \geq 0 \end{cases}$$

- **Estimation techniques of masses in DST framework** $\Longrightarrow$ **Valid in DSmT framework**

3. Belief Function Theories: ***Dezert-Smarandache Theory (3)***

Combination rules

- Classical DSm combination rule (DSmC)
- DSm Hybrid combination rule (DSmH)
- Proportional Conflict Redistribution rules (PCR1, …, PCR5, PCR6)
- …

Decision rules

- Minimum of mass function with an acceptance threshold

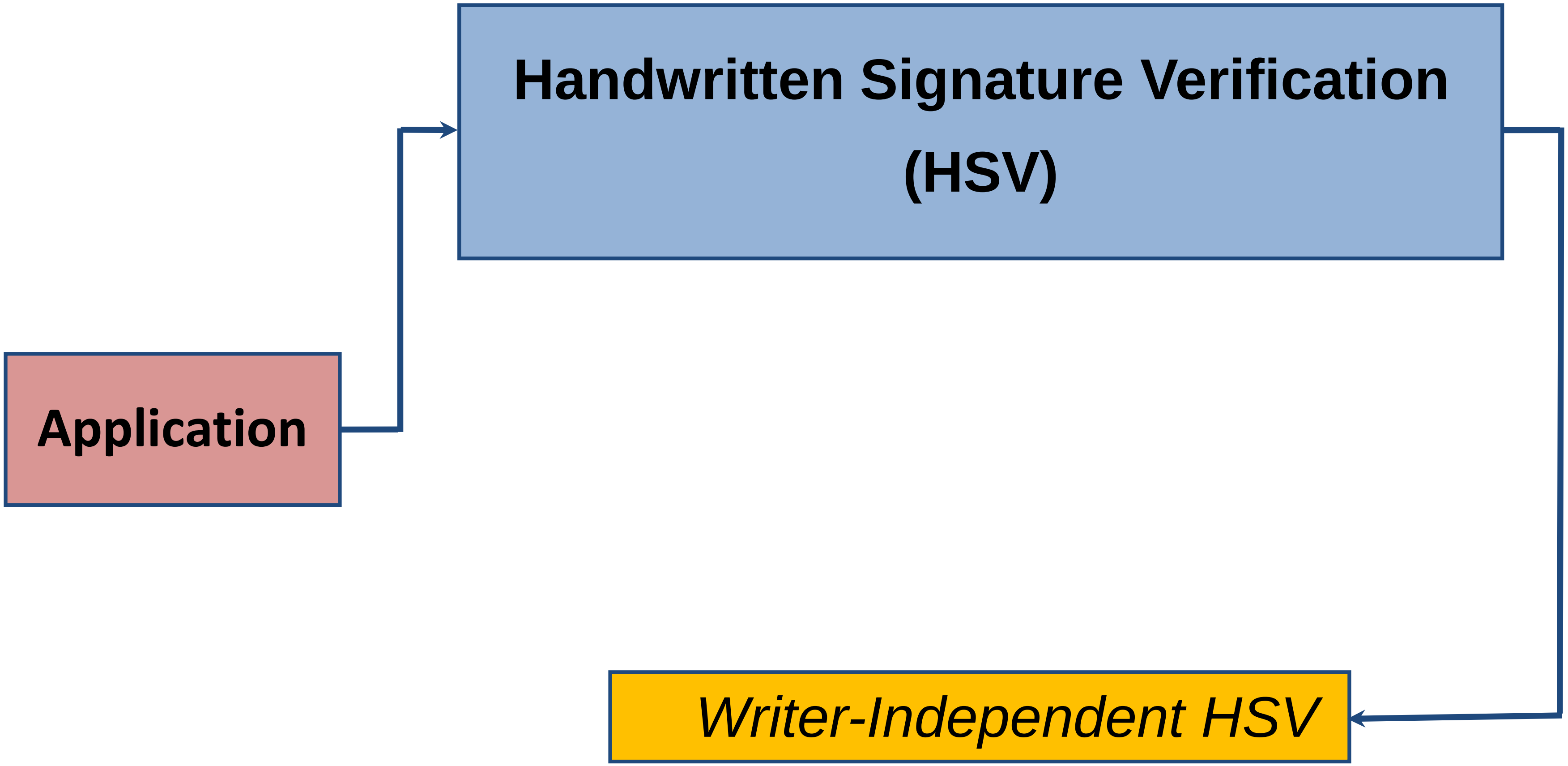
$$\text{Decision} = \begin{cases} \text{Accepted} & \text{if } \min\{m_{test}(\theta_1), m_{test}(\theta_2)\} \geq t_{opt} \\ \text{Rejected} & \text{otherwise} \end{cases}$$

Note:

t_{opt} : Denotes the optimal value of the acceptance decision threshold

$m_{test}(.)$: Defines the combined mass of the simple class θ_i

Parallel Combination of Classifiers for Handwritten Signature Verification



4. Proposed Combination Scheme for Handwritten Signature Verification (1)

Motivation

Physiological

Fingerprint

Face

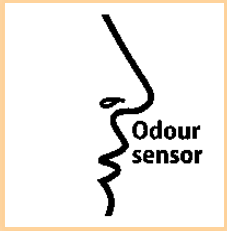
Iris

Hand geometry

Biological

DNA

Smell



Behavioral

Writing

Keyboarding

Gait

Signature

4. Proposed Combination Scheme for Handwritten Signature Verification (2)

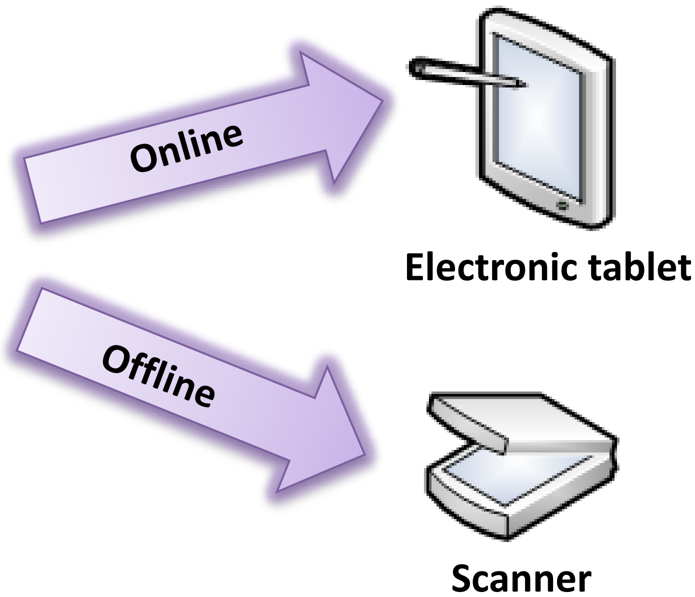
Motivation

- Sign a document to identify himself is a natural gesture.
- Handwritten signature is the biometric modality the most accepted by many peoples.
- It is used in many countries as legal or administrative element.
- Design of a signature verification system is cheaper and more simple comparatively to other biometric systems (for instance iris or face).

4. Proposed Combination Scheme for Handwritten Signature Verification (3)

Motivation

Acquisition of the handwritten signature



Dynamic features (Velocity, Pressure, ….)

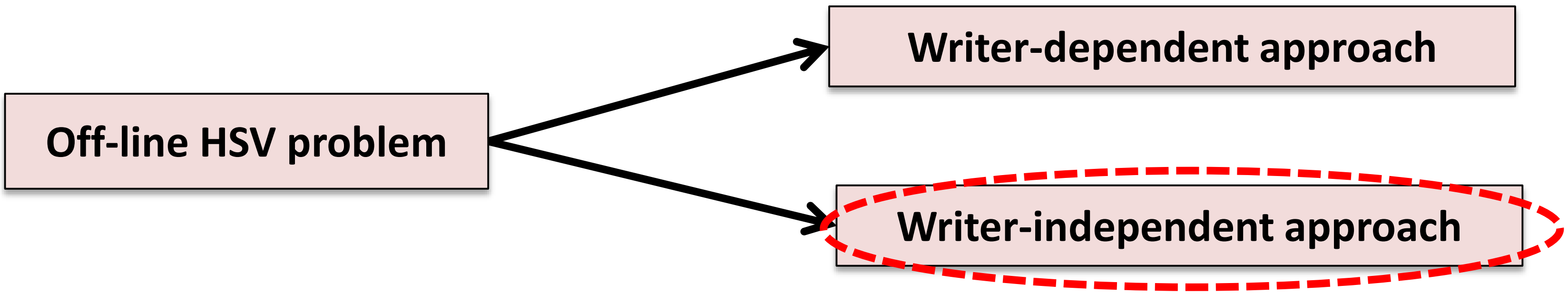
Static features (Image)

Difficulties of the offline handwritten signature verification:

- High variability intra-writer
- Easy to imitate
- Quality of the signature (Paper, Pen, Scanner)

4. Proposed Combination Scheme for Handwritten Signature Verification: ***Writer-Independent Handwritten Signature Verification***

Why use a writer-independent HSV approach ?



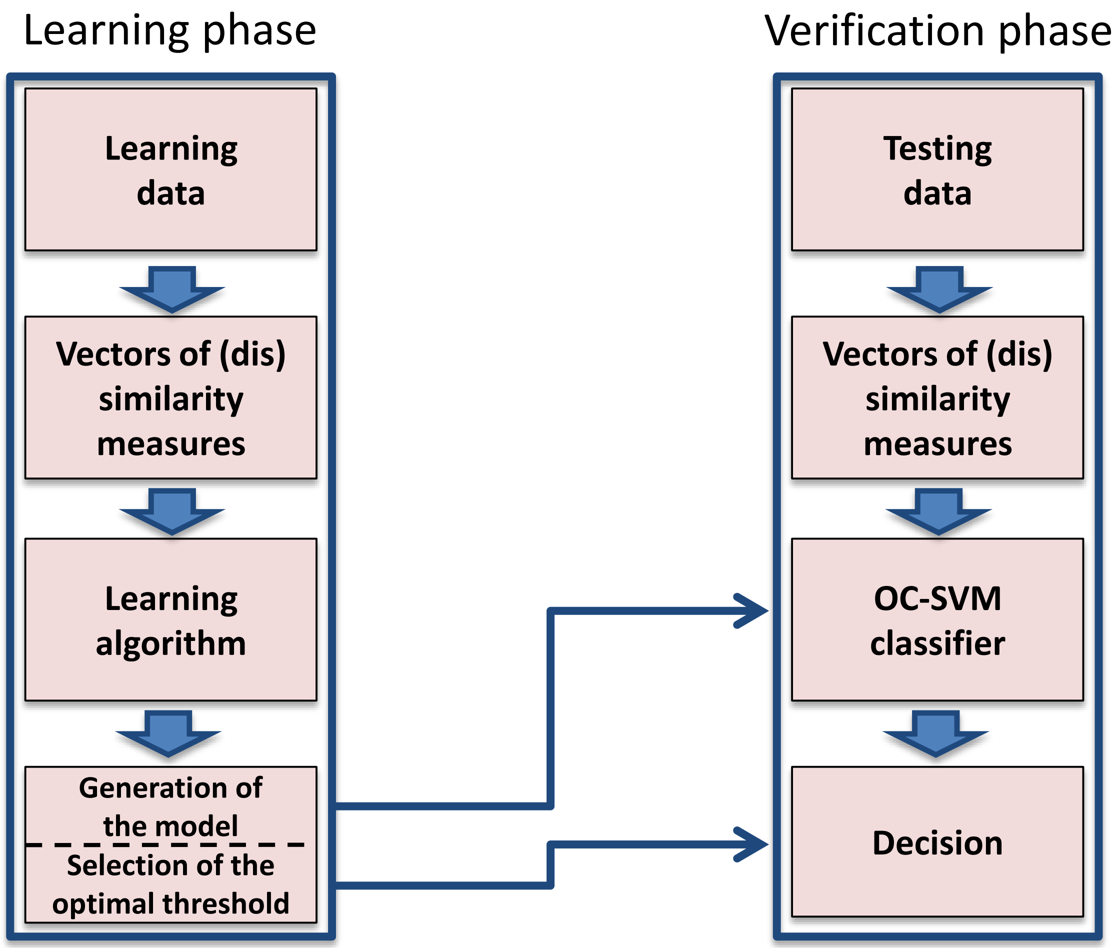
Off-line HSV writer-dependent approach:

- Advantage: Providing a high performance verification
- Limitation: Need of learning the model each time when a new writer should be included in the system

- Solution: (1) Off-line HSV writer-independent approach, (2) Using only genuine signatures, (3) through combination scheme of two individual verification systems

4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Verification scheme using an OC-SVM classifier



4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Classification based on DSmT

Space of discernment: $\Theta = \{\theta_1 \equiv \theta_{descriptor\,1}, \theta_2 \equiv \theta_{descriptor\,2}\}$

Mass function: $m_i(A_j) \in [0,1]$ Such that: $i = \{1,2\}$ (Source 1 or 2)

Source 1 $(\mathrm{OC-SVM}\,1)$

Source 2 $(\mathrm{OC-SVM}\,2)$

Combination
(Conflict management)

Combined sources

Combination space:

- PT: $G = \Theta = \{\theta_1, \theta_2\}$
- DST: $G = 2^{\Theta} = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2\}$
- DSmT: $G = D^{\Theta} = \{\emptyset, \theta_1, \theta_2, \theta_1 \cup \theta_2, \theta_1 \cap \theta_2\}$

Combined masses:

$$m_c(A) = m_1(X) \oplus m_2(Y) \qquad (A, X, Y) \in G \times G \times G$$

$$\sum_{j=1}^{Card\,(G)} m_c(A_j) = 1$$

4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Decision making in both DST and DSmT frameworks

Select outputs of both classifiers (validation phase)

$$t_{mass_1} = \min\{m_c(\theta_1), m_c(\theta_2)\}$$

Combinaison of masses (validation phase)

Combinaison of masses (learning and validation)

$$t_{mass_2} = \min\{\min(m_{learn}(\theta_1)), \min(m_{learn}(\theta_2))\}$$

Compute the optimal decision threshold

$$t_{mass_{opt}} = \frac{t_{mass_1} + t_{mass_2}}{2}$$

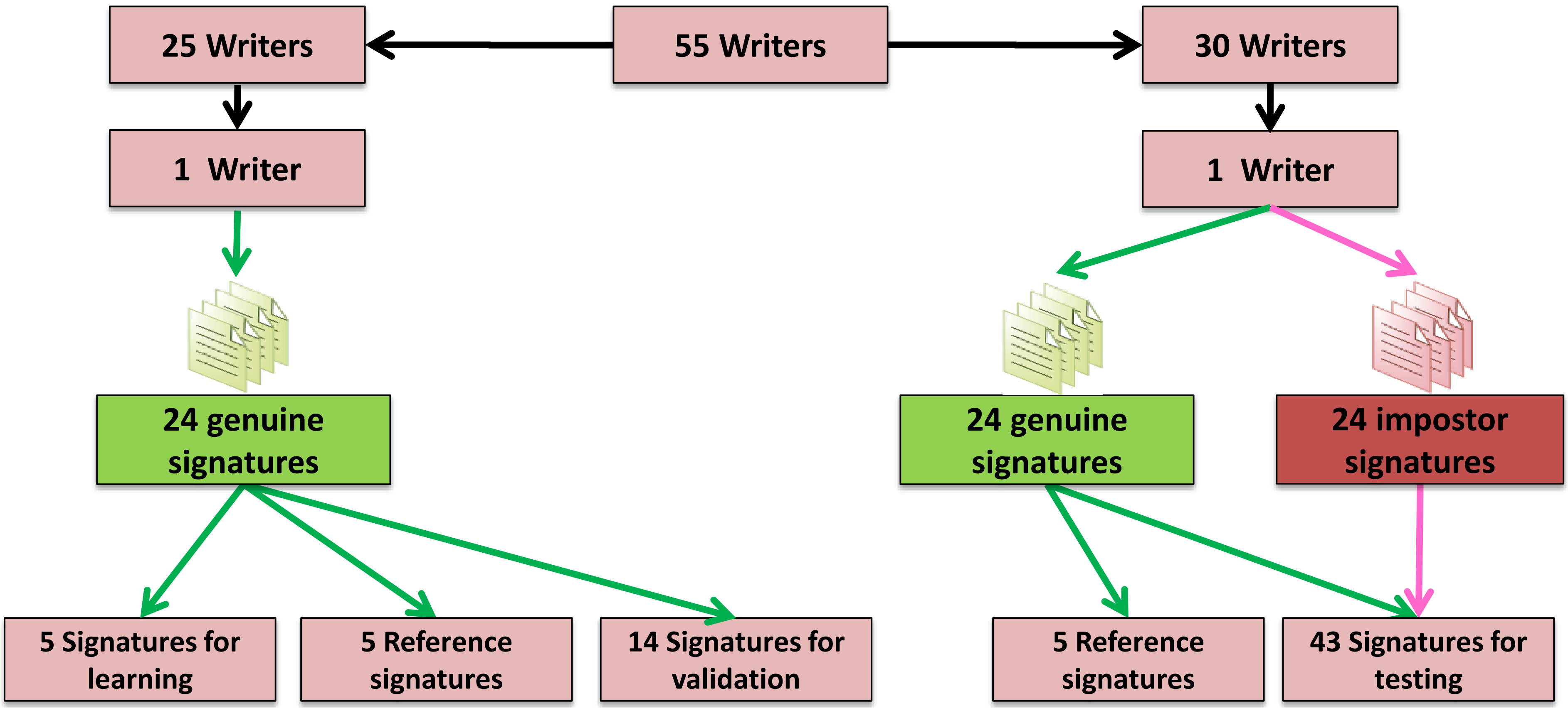
Decision making

$$\text{Decision} = \begin{cases} \text{Accepted if } \min\{m_{test}(\theta_1), m_{test}(\theta_2)\} \geq t_{mass_{opt}} \\ \text{Rejected} \qquad \text{otherwise} \end{cases}$$

4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Case study: Combining two Off-Line HSV Systems (1)

Partitioning of the CEDAR database:



4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Case study: Combining two Off-Line HSV Systems (2)

Feature generation: Simple features are generated from each off-line signature image, which are:

- Discrete cosine transform (DCT) based features
- Curvelet transform (CT) based features

Advantage of both transforms :

- DCT: Two important properties: Decorrelation and energy compaction
- CT: Analyzing local line or curve singularities

Sources of information: Two sources are considered

- Source 1: DCT based descriptor
- Source 2: CT based descriptor

Performance criteria: Three popular errors are considered

- False Rejection Rate (FRR)
- False Acceptance Rate (FAR)
- Average Error Rate (AER)

4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Case study: Combining two Off-Line HSV Systems (3)

Comparative analysis:

Algorithm	Optimal Threshold	Verification Error Rates (%)		
		FRR	FAR	AER
OC-SVM classifier 1 (DCT)	-0.060712	28.7719	44.0278	37.2868
OC-SVM classifier 2 (CT)	-0.419880	9.6491	0.0000	4.2636
Max combination rule	-0.060710	17.5439	44.0278	32.3256
Sum combination rule	-0.480590	6.8421	44.0278	27.5969
Min combination rule	-0.419880	9.6491	0.0000	4.2636

Table 3. Experimental results of proposed individual systems and classical combination algorithms

4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Case study: Combining two Off-Line HSV Systems (4)

Conflict managing in DSmT framework:

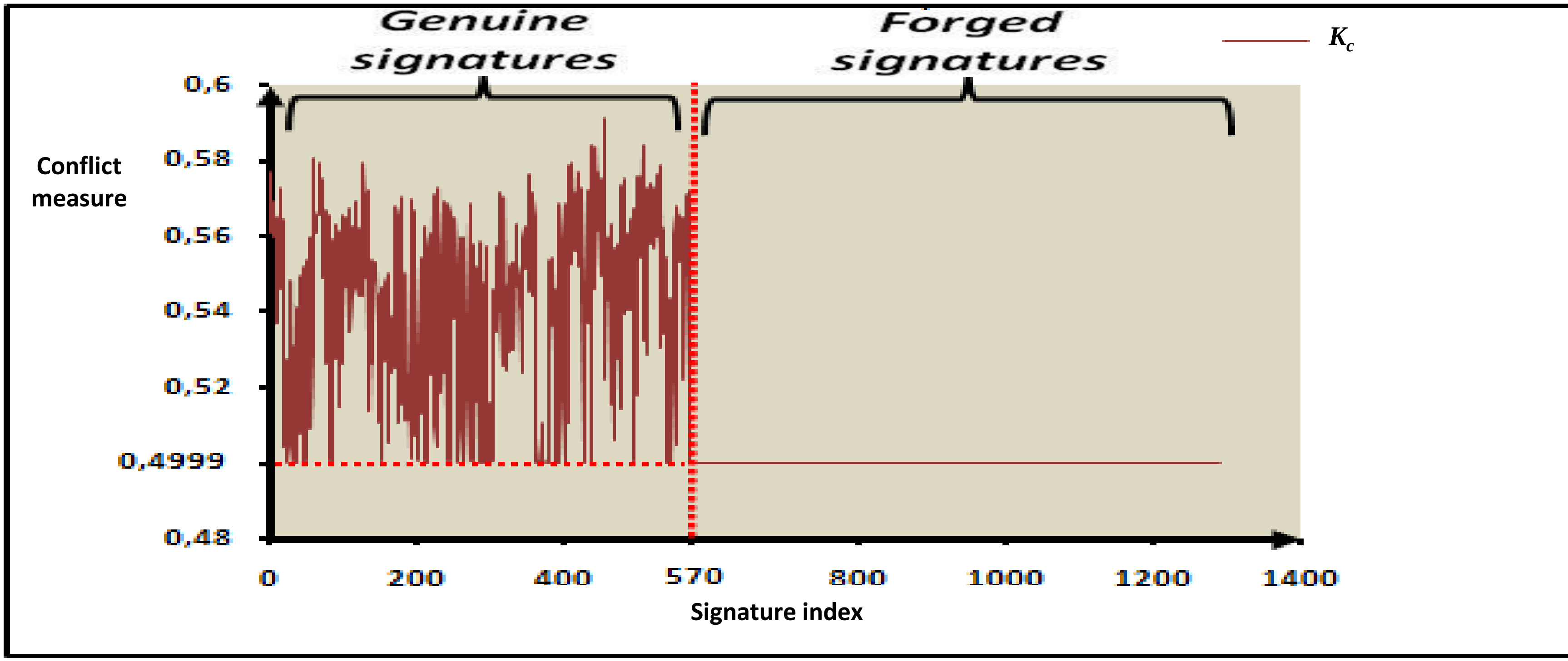


Figure 1. Conflict between both OC-SVM classifiers using DCT and CT-based descriptors for testing signatures

4. Proposed Combination Scheme for Handwritten Signature Verification: *Writer-Independent Handwritten Signature Verification*

Case study: Combining two Off-Line HSV Systems (5)

Comparative analysis:

Algorithm	Optimal Threshold	Verification Error Rates (%)		
		FRR	FAR	AER
OC-SVM classifier 1 (DCT)	-0.060712	28.7719	44.0278	37.2868
OC-SVM classifier 2 (CT)	-0.419880	9.6491	0.0000	4.2636
Max combination rule	-0.060710	17.5439	44.0278	32.3256
Sum combination rule	-0.480590	6.8421	44.0278	27.5969
Min combination rule	-0.419880	9.6491	0.0000	4.2636
DS combination rule	0.334200	0.0000	6.3158	2.7907
PCR6 combination rule	0.267100	0.0000	6.1404	2.7132

Table 4. Experimental results of proposed algorithms

6. Conclusion and futur work

Conclusion

- Proposed combination scheme with PCR6 rule yields the best verification accuracy compared to the statistical match score combination algorithms and DS theory-based combination algorithm even when the individual writer-independent off-line HSV systems provide conflicting outputs.

Futur works

- Adapt the use of the evidence supporting measure of similarity (ESMS) criteria to select complementary sources of information using the same proposed combination scheme in order to attempt to improve the FRR.
- Replace the OC-SVM classifier by the "Histogram Symbolic Representation" (SHR) – based one class classifier.

Many thanks for your attention

Questions…

Florentin SMARANDACHE
smarand@unm.edu

[2]Department of Mathematics
University of New Mexico
Gallup, NM 87301, U.S.A.

THE UNIVERSITY of NEW MEXICO

UNM GALLUP