ISO/IEC 19795-1:2021 (E)

Information technology — Biometric performance testing and reporting — Part 1: Principles and framework

Contents

Foreword

Introduction

1 Scope

2 Normative references

3 Terms and definitions

4 Abbreviated terms

5 Conformance

6 General biometric system

6.1 Conceptual representation of general biometric system
6.2 Conceptual components of a general biometric system
6.2.1 Data capture subsystem
6.2.2 Transmission subsystem
6.2.3 Signal processing subsystem
6.2.4 Data storage subsystem
6.2.5 Comparison subsystem
6.2.6 Decision subsystem
6.2.7 Administration subsystem
6.2.8 Interface to external application
6.3 Functions of general biometric system
6.3.1 Enrolment
6.3.2 Verification of a positive biometric claim
6.3.3 Identification
6.4 Enrolment, verification and identification transactions
6.5 Performance measures
6.5.1 Error rates
6.5.2 Throughput rates
6.5.3 Types of performance testing

7 Planning the evaluation

7.1 General
7.2 Determine information about the system
7.3 Controlling factors that influence performance
7.4 Test subject selection
7.5 Test size
7.5.1 General
7.5.2 Collecting multiple recognition transactions per test subject per system
7.5.3 Requirements on test size
7.6 Multiple tests

8 Data collection

8.1 Avoidance of data collection errors
8.2 Data and details collected
8.3 Enrolments
8.3.1 Enrolment transactions
8.3.2 Enrolment conditions
8.3.3 Enrolment failures and presentation errors

8.4 One-to-one comparison trials
8.4.1 General
8.4.2 Collection conditions
8.4.3 Frequency of use
8.4.4 Systems performing optimization based on enrolled references
8.4.5 Systems performing reference adaptation
8.4.6 Processes for data entry errors and system misuse
8.4.7 Failures to acquire
8.4.8 Adding test data to the corpus
8.4.9 Online comparison trials
8.4.10 Offline comparison trials
8.4.11 Offline non-mated comparison trials when references are dependent
8.4.12 Offline non-mated comparison trials based on comparison of references
8.4.13 Use of samples from multi-capture comparison transactions
8.5 Identification trials
8.5.1 General
8.5.2 Identification testing with non-enrolled test subjects
8.5.3 Use of jack-knife approach for identification testing

9 Analyses

9.1 General
9.2 Performance of biometric enrolment
9.2.1 Failure-to-enrol rate
9.2.2 Enrolment transaction duration
9.3 Performance of biometric acquisition
9.3.1 Failure-to-acquire rate
9.3.2 Acquisition process duration
9.3.3 Other aspects of acquisition performance
9.4 One-to-one comparison performance
9.4.1 False non-match rate
9.4.2 False match rate
9.5 Verification system performance metrics
9.5.1 General
9.5.2 False reject rate
9.5.3 False accept rate
9.5.4 Verification transaction duration
9.5.5 Generalized false reject rate and generalized false accept rate
9.6 Identification system performance metrics
9.6.1 General
9.6.2 False-negative identification rate
9.6.3 False-positive identification rate
9.6.4 Generalized false-negative identification rate and generalized false-positive identification rate
9.6.5 Selectivity
9.6.6 Closed-set test of identification performance
9.6.7 Estimation of identification error rates from one-to-one comparison results
9.6.8 Predicting identification error rates in larger populations
9.7 Analysis of performance across controlled experimental factors
9.7.1 Longitudinal analyses
9.7.2 Pairwise analyses
9.8 Detection error trade-off
9.9 Transaction durations
9.10 Computational workload
9.11 Uncertainty of estimates

10 Graphical presentation of results

10.1 Score distributions
10.1.1 General
10.1.2 Boxplots
10.2 Error rate vs threshold plot
10.3 DET plot
10.4 CMC plot / FNIR over rank plot
10.5 FNIR over number of enrolees plot
10.6 Heat maps

11 Record keeping

12 Reporting performance results

12.1 Reporting test details
12.2 Summary statistics
12.3 Reporting enrolment performance
12.4 Reporting acquisition performance
12.5 Reporting one-to-one comparison performance
12.6 Reporting verification system performance
12.7 Reporting identification system performance
12.8 Reporting performance across factors

Annex A (informative) Differences between evaluation types

Annex B (Informative) Test size and random uncertainty

B.1 Confidence intervals and test size assuming independent identically distributed comparisons
B.1.1 Rule of 3
B.1.2 Rule of 30
B.1.3 Number of comparisons to support a claimed error rate
B.2 Variance of performance measures as a function of test size
B.3 Estimates for variance of performance measures
B.3.1 General
B.3.2 Variance of observed false non-match rate
B.3.2.1 False non-match rate — Single attempt per test subject
B.3.2.2 False non-match rate — Multiple attempts per test subject
B.3.2.3 False non-match rate — Unequal numbers of attempts per test subject
B.3.3 Variance of observed false match rate
B.4 Estimating confidence intervals
B.4.1 General
B.4.2 Bootstrap estimates of the variance and confidence intervals
B.4.3 Subset sampling

Annex C (informative) Factors influencing performance

C.1 General
C.2 List of factors
C.2.1 Population demographics
C.2.2 Application
C.2.3 Capture subject anatomy
C.2.4 Capture subject behaviour
C.2.5 Capture subject appearance
C.2.6 Environmental influences
C.2.7 Sensor and hardware
C.2.8 User interface
C.3 Examples for reporting
C.3.1 Finger position
C.3.2 Change of illumination
C.3.3 Illumination reflections
C.3.4 Glasses
C.3.5 Dirt on platen
C.3.6 Weather

Annex D (Informative) Pre-selection algorithm performance

Annex E (informative) Identification performance as a function of database size

Annex F (informative) Algorithms for generating DET and CMC

F.1 Algorithm for generating DET data points
F.1.1 MATLAB code for DET calculation
F.1.2 Python code for DET calculation
F.1.3 R code for DET calculation
F.2 Algorithm for generating CMC data points

Annex G (informative) DET properties and interpretation