Biometryczna Identyfikacja Tożsamości

Wykład 1: Biometria – znaczenie pojęcia

Adam Czajka

Wykład na Wydziale Elektroniki i Technik Informacyjnych Politechniki Warszawskiej

Semestr zimowy 2015/16

Biometria języka ...

Wykład 1: Biometria – znaczenie pojęcia

- Informacje podstawowe
- (Krótka) historia biometrii
- Modalności biometryczne
- Pożądane własności biometrii
- System biometryczny i jego decyzje

Wykład 1: Biometria – znaczenie pojęcia

Biometria w dwóch znaczeniach

1. Biometria w szerszym znaczeniu
   - etymologia: pomiar własności istot żywych (grec. *bios* = "życie", *metron* = "pomiar")
   - niesprecyzowany cel pomiaru (np. diagnostyka medyczna)

Biometria w dwóch znaczeniach

2. Biometria jako dział informatyki
   - pomiar własności anatomicznych lub własności zachowania człowieka
   - określony cel pomiaru: automatyczne rozpoznanie tożsamości

Zastosowanie własności anatomicznych lub własności zachowania człowieka do automatycznego rozpoznania jego tożsamości

Biometria jako dział informatyki

Zastosowanie własności anatomicznych lub własności zachowania człowieka do automatycznego rozpoznania jego tożsamości, czyli:

- nie decyduje ekspert, ale
- metodyka rozpoznawania może częściowo wykorzystywać doświadczenie eksperta,
- możliwość braku nadzoru, konieczna duża szybkość, powtarzalność i przewidywalność przetwarzania danych.

Biometria jako dział informatyki

Zastosowanie własności anatomicznych lub własności zachowania człowieka do automatycznego rozpoznania jego tożsamości, czyli:

- "jacy jesteśmy", "co nas charakteryzuje"
- nie "coś, co wiemy" (hasło, PIN)
- nie "coś, co mamy" (klucz, karta)
- czasem zamiast "anatomiczne": biologiczne lub fizjologiczne (ale to zbyt ogólne)
- czasem zamiast "własności zachowania": własności behawioralne (ale to niepotrzebna konotacja z teorią behawioru)

Biometria jako dział informatyki

Zastosowanie własności anatomicznych lub własności zachowania człowieka do automatycznego rozpoznania jego tożsamości, czyli:

- przetwarzane dane muszą być wynikiem właściwego pomiaru własności żywej osoby
- urządzenia (sensory, czytniki) biometryczne muszą dostarczać właściwe próbki biometryczne
- badanie żywotności niezbędne aby system można było nazywać biometrycznym

Rozpoznanie biometryczne

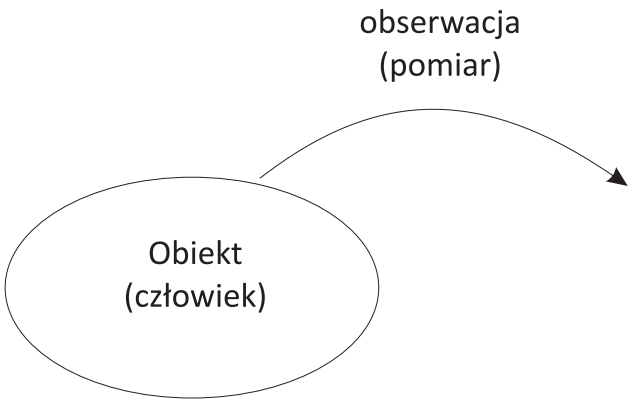
Wayman, Jain, Maltoni, Maio, 2005

1. Pozytywne rozpoznanie
   (ang. *positive recognition*)
   $\Rightarrow$ weryfikacja hipotezy: próbka pochodzi od osoby znanej systemowi (wcześniej zarejestrowanej)
2. Negatywne rozpoznanie
   (ang. *negative recognition*)
   $\Rightarrow$ weryfikacja hipotezy: próbka pochodzi od osoby nieznanej systemowi (wcześniej niezarejestrowanej)

Schematy uwierzytelniania

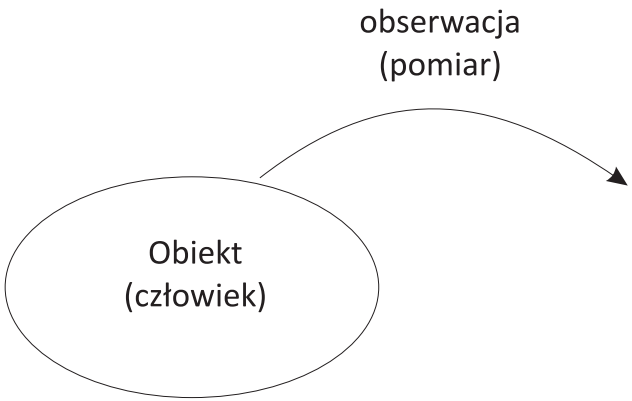
1. Klasyczne
   - ustalenie tożsamości
     (identyfikacja, często określane jako porównanie 1:N)
   - potwierdzenie tożsamości
     (weryfikacja, często określane jako porównanie 1:1)
2. Nowe (wynikające z możliwości biometrii)
   - negatywne uwierzytelnienie
     - negatywna identyfikacja: nie należę do grupy X
     - negatywna weryfikacja: nie jestem osobnikiem X
   - eliminacja "wielokrotnych tożsamości"

Podstawowe pojęcia



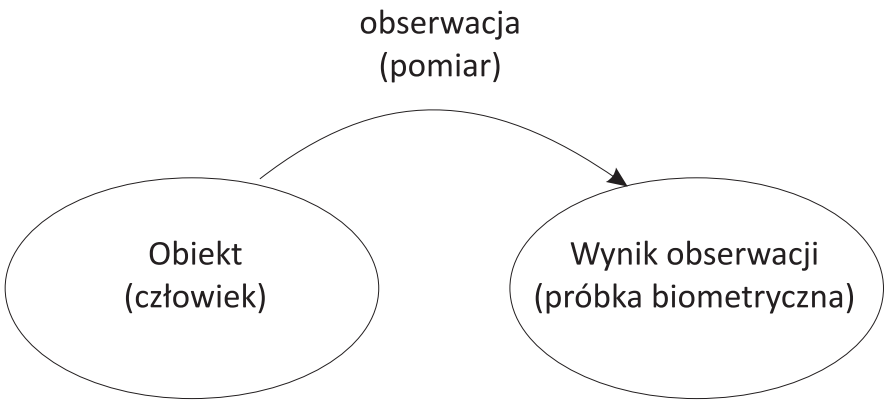
1. Obiekt (człowiek)
2. Obserwacja obiektu (pomiar)
3. **Charakterystyka lub właściwość biometryczna** (ang. *biometric characteristic*)
   - dziedzina obserwacji
   - przykłady: wygląd twarzy, kształt dłoni, naczynia krwionośne dłoni

Podstawowe pojęcia



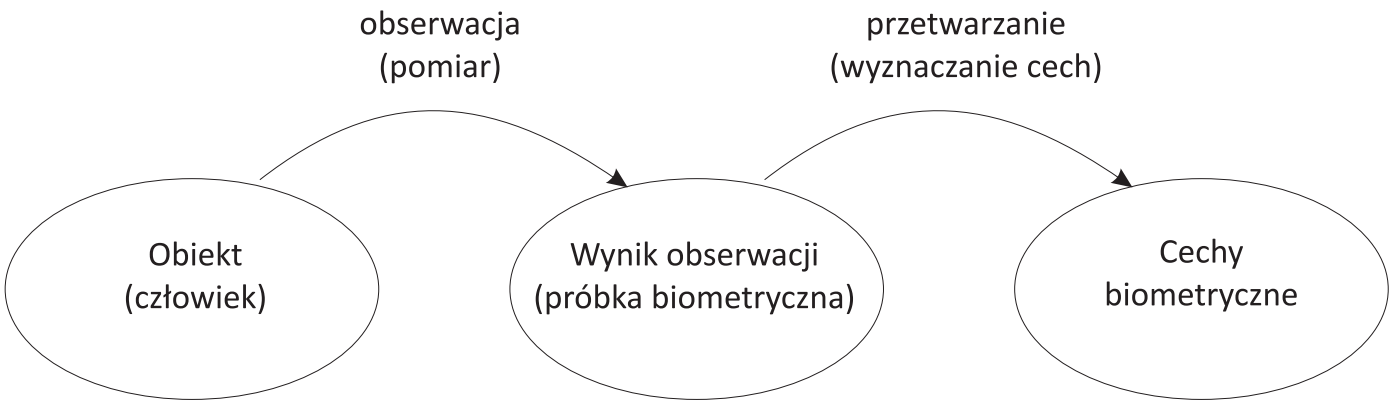
4. **Modalność biometryczna** (ang. *biometric mode*)
   - kombinacja charakterystyki biometrycznej oraz sposobów pomiaru i przetwarzania
   - przykłady: wygląd twarzy 3D, cechy termiczne dłoni w świetle podczerwonym, geometria dłoni w świetle podczerwonym

Podstawowe pojęcia



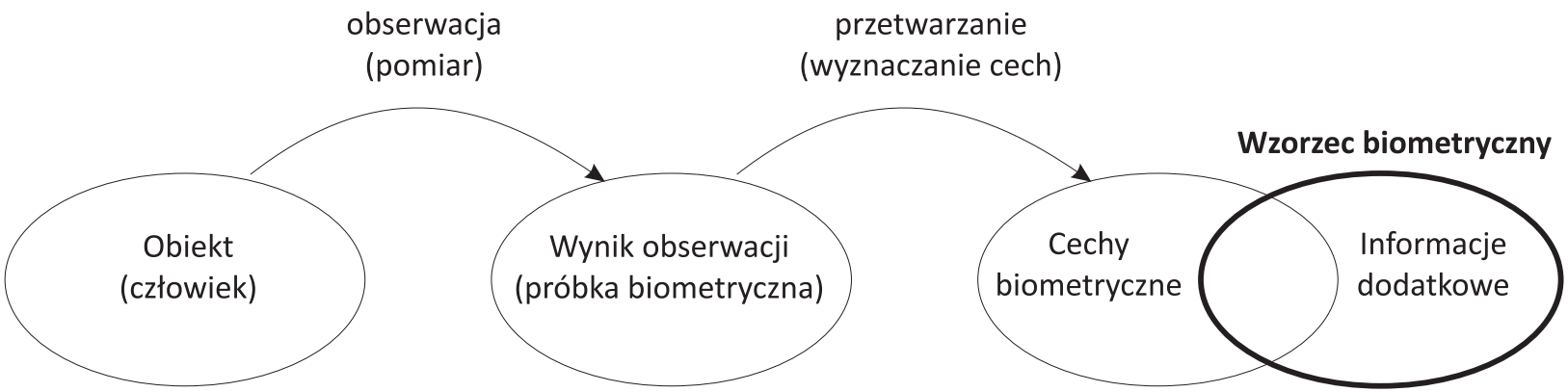
5. **Próbka biometryczna** (ang. *biometric sample*)
   - wynik obserwacji; w praktyce wyniki pomiaru (surowe lub wstępnie przetworzone)
   - przykłady: obraz twarzy 3D, obraz dłoni 2D, obraz naczyń krwionośnych

Podstawowe pojęcia



6. **Cecha biometryczna** (ang. *biometric feature*)
   - reprezentacja (zwykle skrócona) próbki biometrycznej
   - przykłady: odległości między punktami charakterystycznymi twarzy, długości palców dłoni, szerokości naczyń krwionośnych

Podstawowe pojęcia



7. **Wzorzec biometryczny** (ang. *biometric reference*)
   - dane referencyjne, które zachowujemy w bazie danych na potrzeby rozpoznawania biometrycznego
   - próbki biometryczne, przetworzone próbki biometryczne, wyselekcjonowane cechy biometryczne lub dodatkowo informacje niezbędne do uwierzytelnienia obiektu

Wykład 1: Biometria – znaczenie pojęcia

Historia: pierwsze metody nieautomatyczne

Odciski palców znalezione na wyspie Gavrinis we Francji (datowane na ok. 2000 lat p.n.e.)

źródło: A. Jain, "Introduction to biometrics", w: "Biometrics: Personal Identification in Networked Society", Kluwer, 1998

1. 31 000 lat p.n.e.
   - jaskinia Chauvet, najstarsze odciski dłoni, hipoteza (niepotwierdzona) o wykorzystaniu odcisków w identyfikacji twórców malowideł
2. VII w. p.n.e.
   - starożytne Chiny, Babilonia i Asyria
   - powiązanie osób z transakcjami i wydarzeniami
   - tabliczki gliniane z odciskami palców
3. Stary Testament
   - weryfikacja przynależności do grupy: wymowa słowa "Shibboleth"
   - oszustwo związane w "weryfikacją biometryczną": imitacja skóry dłoni i szyi za pomocą skórek z koźlęcia

Historia: systemy automatyczne

1. Lata 60-te XX wieku
   - systemy automatycznej identyfikacji tożsamości wykorzystujące odcisk palca i rozpoznawanie głosu
2. Lata 70-te XX wieku
   - systemy oparte na geometrii dłoni
3. Lata 80-te XX wieku
   - systemy wykorzystujące obraz siatkówki oka oraz podpis odręczny
4. Lata 90-te XX wieku
   - systemy biometrii tęczówki
5. XXI wiek
   - zastosowania biometrii na masową skalę
   - biometryczne urządzenia przenośne
   - brak wymagania kooperacji z użytkownikiem

Wykład 1: Biometria – znaczenie pojęcia

Podstawowy podział modalności

1. Związane z właściwościami anatomicznymi
   - zwykle obserwacja chwilowa, pomiar statyczny
   - ewentualne zależności czasowe w wynikach pomiaru nie uwzględniane w uwierzytelnieniu
2. Związane z naszym zachowaniem
   - obserwacja akcji (najczęściej świadomej) wykonywanej przez użytkownika
   - obserwacja w czasie, pomiar dynamiczny
   - zależności czasowe w wynikach pomiaru podstawą uwierzytelniania

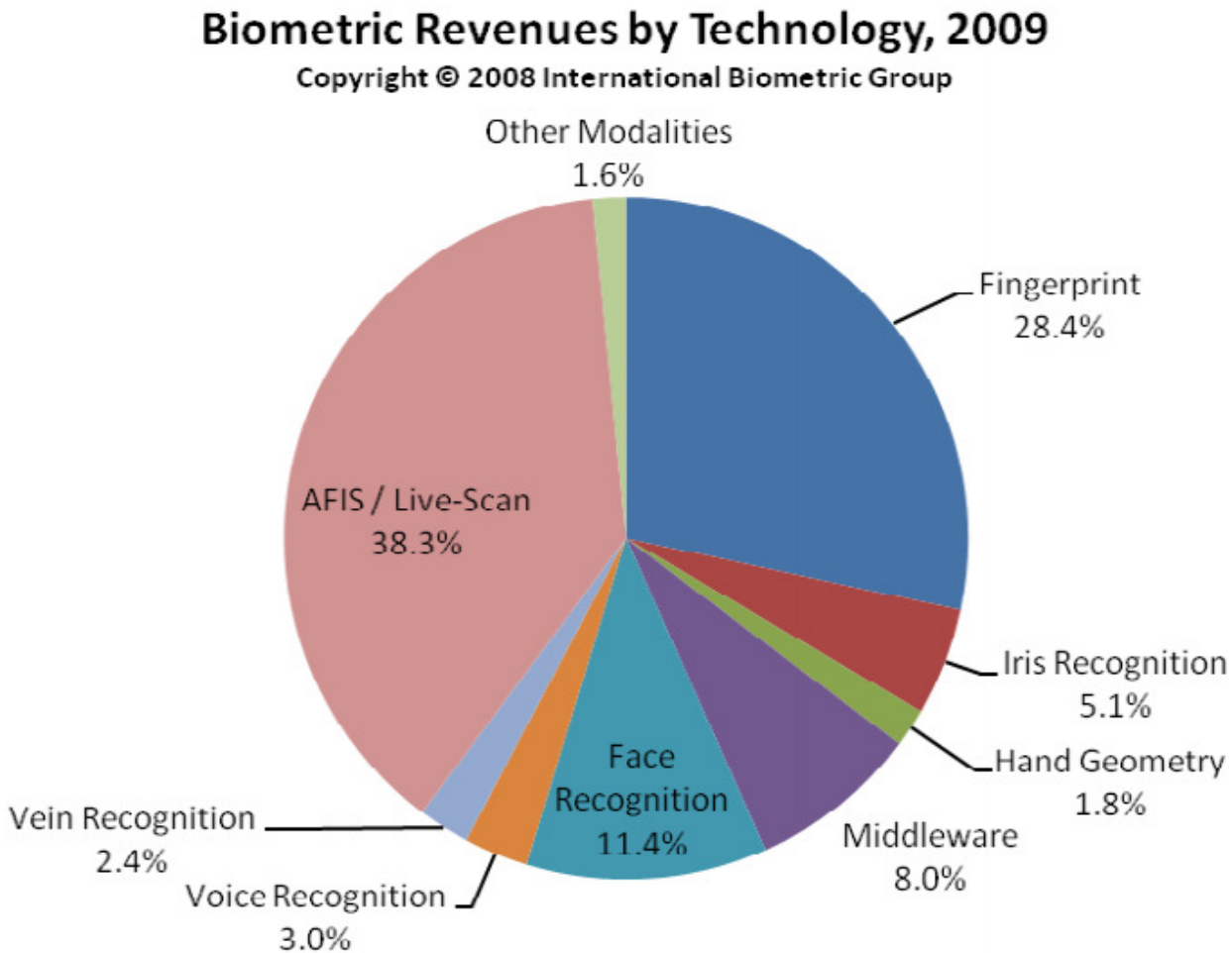
Modalności oparte o właściwości anatomiczne

1. Stosowane w praktyce
   - palec: odcisk, układ żył
   - dłoń: geometria 2D/3D, termika, układ żył, odcisk
   - twarz: geometria 2D/3D, termika
   - oko: tęczówka, układ żył w naczyniówce (popularna choć niewłaściwa nazwa: "biometria siatkówki") lub w białkówce (naczynia widoczne gołym okiem)
2. Możliwe do wykorzystania w praktyce
   - DNA
   - zapach
   - ucho: geometria, termika
   - palec: termika, geometria, struktura kanałów pod paznokciem

Modalności oparte o cechy naszego zachowania

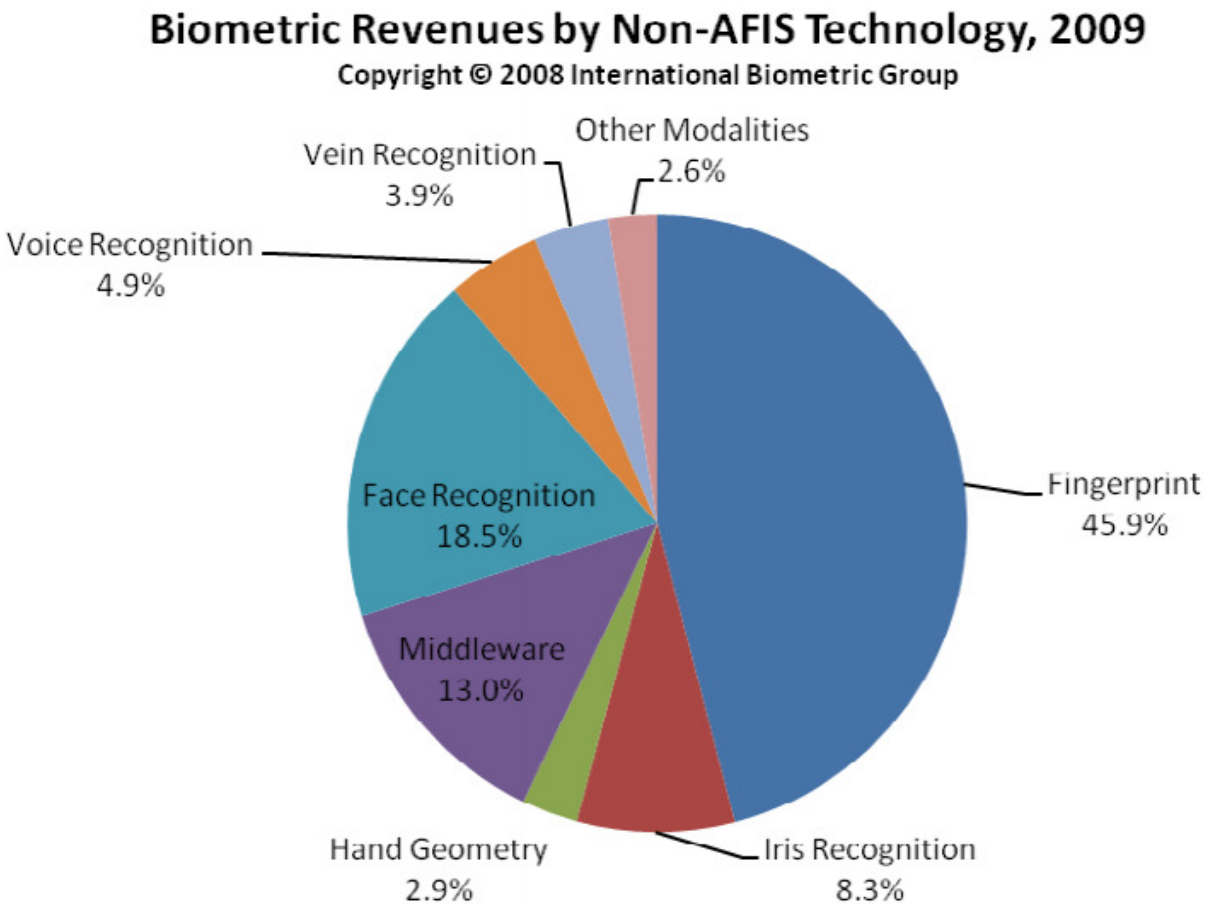
1. Stosowane w praktyce
   - podpis odręczny
   - głos: rozpoznawanie mówiącego (ale nie mowy)
   - rytm uderzania w klawisze
2. Możliwe do wykorzystania w praktyce
   - fale mózgowe EEG
   - pismo ręczne
   - oko: dynamika gałki ocznej, dynamika źrenicy
   - sposób chodzenia
   - ruch warg
   - palec: rezonans opuszków palców
   - skojarzenia

Modalności biometryczne: udział w rynku



Źródło: IBG Biometrics Market and Industry Report 2009-2014

Modalności biometryczne: udział w rynku

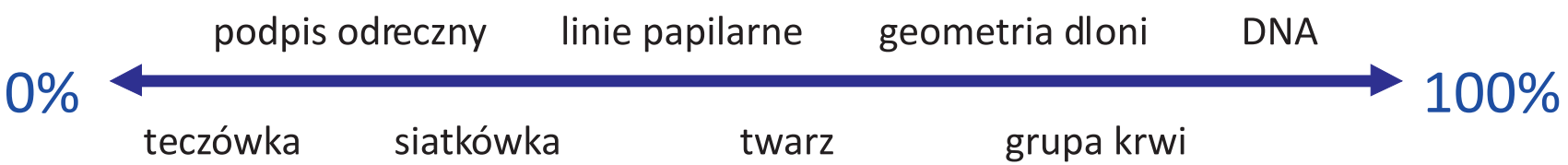


Źródło: IBG Biometrics Market and Industry Report 2009-2014

Wykład 1: Biometria – znaczenie pojęcia

Pożądane własności biometrii

1. Wysoka zawartość informacyjna zapewniająca wysoki stopień zróżnicowania
   - genotyp i penetracja genetyczna
   - bliźnięta 1 na 80 urodzin, bliźnięta jednojajowe 1 na 240 urodzin (co najmniej 0.8% osób błędnie rozpoznanych)
   - hipoteza o unikalności cech uzasadniona tylko eksperymentalnie



Zależność wybranych modalności od genotypu

Pożądane własności biometrii

2. Niezmienność w czasie
   - odporność na zmiany powodowane chorobami
   - problem starzenia się wzorców, ang. *template ageing*
     - aktualne badania dostarczają sprzecznych wniosków nt. zależności dokładności rozpoznawania biometrycznego w funkcji upływu czasu
     - trudność w pozyskaniu odpowiednich baz danych; przykłady baz naukowych: MORPH (twarz; dwa albumy: 515 i 4000 osób; interwał od kilku miesięcy do 29 lat dla pierwszego albumu, kilka lat dla drugiego albumu), KFRIA Ageing DB (odcisk palca; 100 osób; interwał 1 rok), BioBase II NASK/PW (twarz, tęczówka, odcisk palca, podpis, dłoó; ok. 50 osób; interwał 7 lat), CBSA OPS-XING (a), OPS-FIELD (b) (tęczówka; a) ok. 350k osób, interwał ok. 4 lat; b) ok. 610k osób, interwał ok. 4 lat)

Pożądane własności biometrii

3. Akceptacja użytkowników
   - weryfikacja stosowana w celu zwiększenia wiarygodności, często dla wygody użytkowników
   - brak wymogu współpracy osoby rozpoznawanej (możliwe nadużycia)
   - obawy społeczne, religijne, kulturowe, ochrona danych osobowych
   - obawy natury zdrowotnej (np. czy oświetlenie kamery nie zniszczy mi wzroku?)

Pożądane własności biometrii

Uwagi dot. akceptacji użytkowników

3a. Ochrona danych osobowych

- dane biometryczne są danymi osobowymi, ale
  - wg art. 6 ust. 3 ustawy z dnia 29 sierpnia 1997 r. ochronie danych osobowych "informacji nie uważa się za umożliwiającą określenie tożsamości osoby, jeżeli wymagałoby to nadmiernych kosztów, czasu lub działań"
  - brak precyzyjnych definicji "nadmiernych kosztów, czasu lub działań", pole do interpretacji

Pożądane własności biometrii

Uwagi dot. akceptacji użytkowników

3a. Ochrona danych osobowych (c.d.)

- obawa przed kradzieżą tożsamości
  - łatwa dostępność danych biometrycznych tradycyjnie służących do weryfikacji (podpis, obraz twarzy)
  - biometria odwrotna, dane syntetyczne
  - urządzenia biometryczne powinny dostarczać jedynie właściwych próbek biometrycznych
  - możliwa identyfikacja pośrednia w kryminalistyce poprzez próbki ukryte (utajone)
- udostępnianie niepożądanych danych, np. obraz tęczówki (choroby?), obraz podpisu (stan psychofizyczny?), obraz dłoni (powodzenie w miłości/biznesie?)

Pożądane własności biometrii

Uwagi dot. akceptacji użytkowników

3b. Syndrom "wielkiego brata"
- przetwarzanie danych bez wiedzy właściciela, brak prywatności
- biometryczne bazy danych, konieczność tworzenia naukowych baz danych
- wszechobecne obliczenia ("ubiquitous computing", "pervasive computing", "cloud computing", itp.)
- podglądanie osobowości

3c. Decyzje maszyny: nieufność

3d. Minimalna inwazyjność i komfort użytkowania

Komfort użytkowania i minimalna inwazyjność: identyfikacja tęczówki
(kadr z filmu *Minority report*, 2002)

Pożądane własności biometrii

4. Odporność na fałszerstwa
   - zmiana własności naszego ciała niemożliwa lub ryzykowna
   - możliwość konstrukcji efektywnych testów autentyczności (danych, obiektów)
5. Możliwość realizacji technicznej
   - łatwość obserwacji i pomiaru
   - powtarzalność obserwacji i pomiaru
   - powszechność własności biometrycznych
   - niski koszt budowy urządzeń

Wykład 1: Biometria – znaczenie pojęcia

Tryby pracy systemu biometrycznego

1. Rejestracja (ang. *enrolment* (UK), *enrollment* (US))
   - wygenerowanie i zapamiętanie wzorca biometrycznego użytkownika (oraz danych innych niż biometryczne)
2. Uwierzytelnienie (ang. *authentication*)
   - wykorzystanie zachowanego wzorca i wyznaczonych aktualnie cech biometrycznych w celu potwierdzenia lub ustalenia tożsamości

Tryby pracy systemu biometrycznego

Rejestracja

1. Nadzorowana przez człowieka
2. Wielokrotny pomiar
   - zaostrzona kontrola jakości próbek w celu pozyskania reprezentatywnych danych
3. Przetwarzanie surowych danych i utworzenie wzorca
   - kontrola spójności cech biometrycznych
   - opcjonalnie: estymacja zmienności cech biometrycznych (wynikającej ze zmienności obiektu i możliwości pomiarowych) i dobór indywidualnego poziomu zgodności cech
4. Powiązanie danych biometrycznych z pozostałymi danymi osobowymi
5. Zapis wzorców biometrycznych w bazie danych lub w indywidualnym nośniku danych (np. karcie elektronicznej)

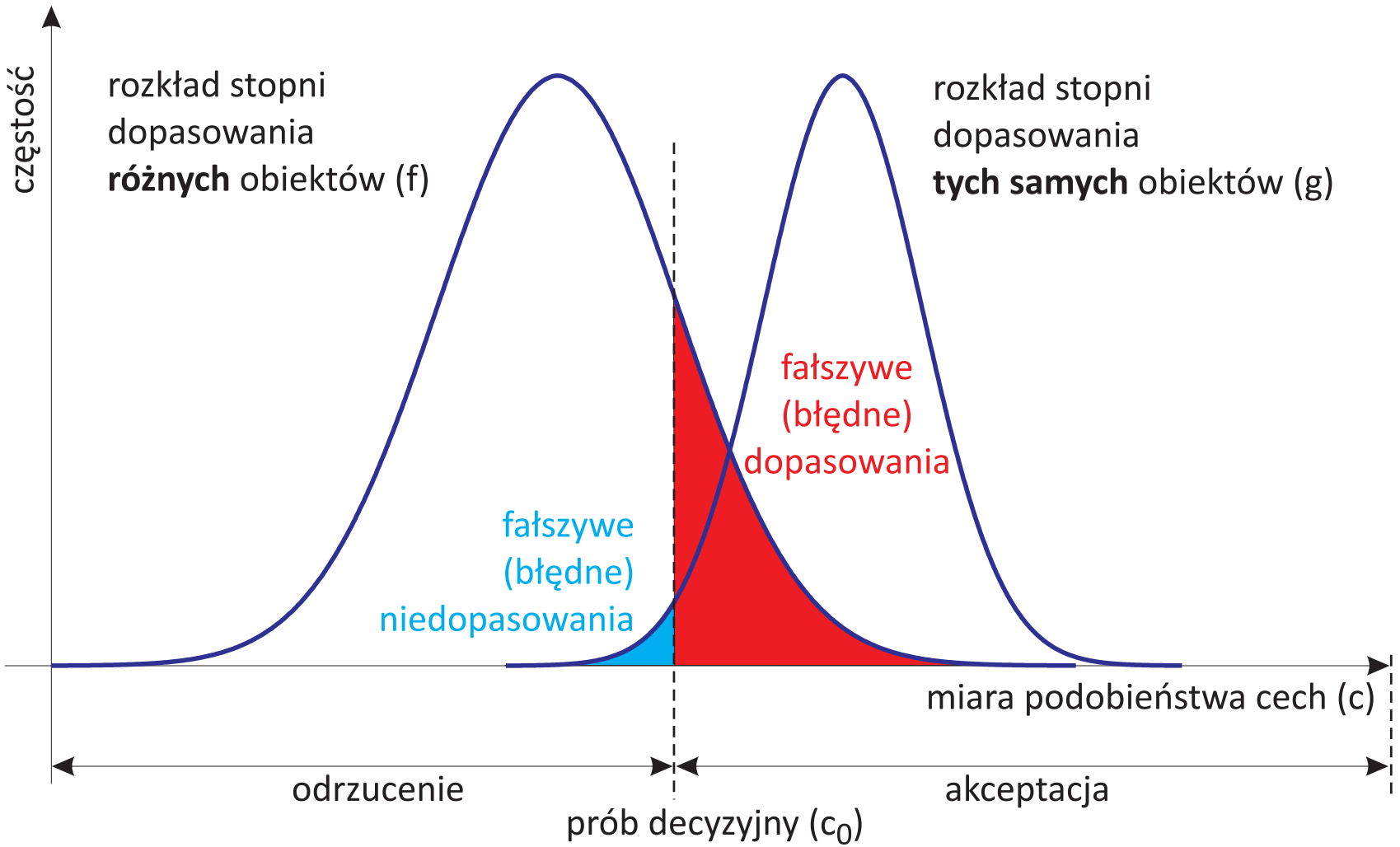
Tryby pracy systemu biometrycznego

Uwierzytelnienie

1. Bez nadzoru człowieka
2. Odczytanie wzorca biometrycznego (lub wzorców) z bazy danych (lub innego nośnika)
3. Najczęściej jednokrotny pomiar
   - kontrola jakości próbek w celu minimalizacji błędnych odrzuceń
   - test żywotności / autentyczności
4. Przetwarzanie surowych danych i wyznaczenie cech
5. Wyznaczenie dopasowania
   - uwzględnienie zmienności cech biometrycznych
6. Podjęcie decyzji
   - konfrontacja wyniku dopasowania z zapamiętanym progiem zgodności
   - opcjonalnie: uwzględnienie indywidualnego poziomu zgodności cech

1. Błędy metod biometrycznych
   - fałszywe (błędne) niedopasowanie (ang. *false non-match*)
   - fałszywe (błędne) dopasowanie (ang. *false match*)
2. Stopień dopasowania/podobieństwa
   - miara podobieństwa cech (biometrycznych) wyznaczonych dla weryfikowanej próbki i cech (biometrycznych) zawartych we wzorcu biometrycznym
3. Decyzja
   - dopasowanie (ang. *match*) lub niedopasowanie (ang. *non-match*) podejmowana z wykorzystaniem progu decyzyjnego
   - aby zaakceptowaæ próbkę stopień dopasowania musi przekraczaæ próg decyzyjny określony dla danego urządzenia/systemu


częstość
rozkład stopni dopasowania **różnych** obiektów (f)
rozkład stopni dopasowania **tych samych** obiektów (g)
miara podobieństwa cech (*c*)


częstość
rozkład stopni dopasowania **różnych** obiektów (f)
rozkład stopni dopasowania **tych samych** obiektów (g)
fałszywe (błędne) dopasowania
fałszywe (błędne) niedopasowania
miara podobieństwa cech (c)
odrzucenie
akceptacja
prób decyzyjny (c_0)

Aproksymacja prawdopodobieństw błędów

Stopień fałszywych (błędnych) niedopasowań (ang. *False Non-Match Rate – FNMR*)

Jeśli znamy g:

$$g_{\mathsf{FNM}}(c_0) = \int_{-\infty}^{c_0} g(c)dc$$

w przeciwnym wypadku:

$$\widehat{g}_{\mathsf{FNM}}(c_0) = \mathsf{FNMR}(c_0) = \frac{\text{liczba fałszywych niedopasowań dla } c_0}{\text{liczba wszystkich porównań w próbie}}$$

Aproksymacja prawdopodobieństw błędów

Stopień fałszywych (błędnych) dopasowań (ang. *False Match Rate – FMR*)

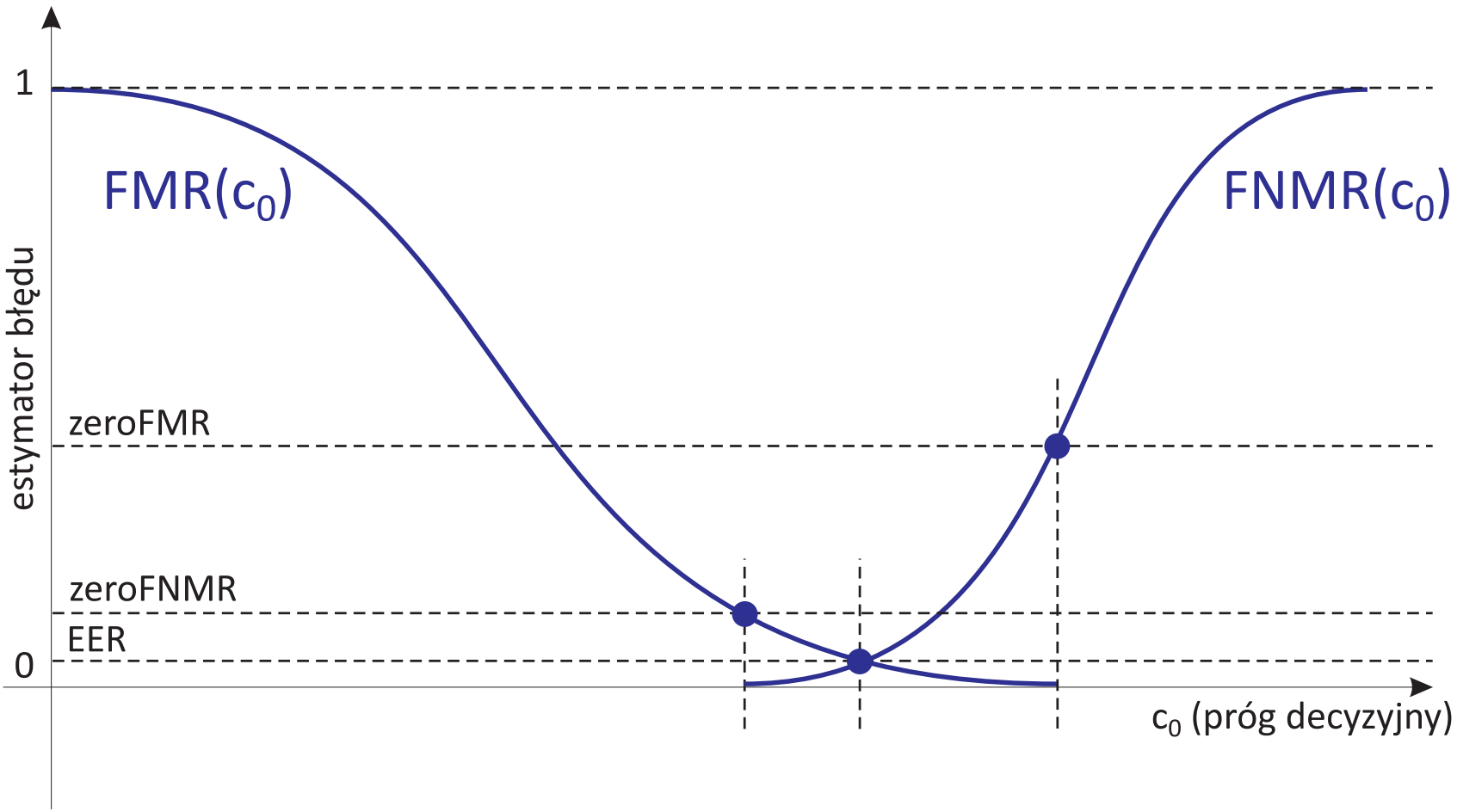
Jeśli znamy f:

$$f_{\mathsf{FM}}(c_0) = \int_{c_0}^{\infty} f(c)dc$$

w przeciwnym wypadku:

$$\widehat{f}_{\mathsf{FM}}(c_0) = \mathsf{FMR}(c_0) = \frac{\text{liczba fałszywych dopasowań dla } c_0}{\text{liczba wszystkich porównań w próbie}}$$

Podstawowe wskaźniki oceny



Podstawowe wskaźniki oceny

1. FTA: stopień nieudanych pomiarów (ang. *Failure To Acquire*)
2. FTE: Stopień odrzuconych rejestracji (ang. *Failure To Enroll*)
3. FMR/FNMR dla danego punktu pracy (progu decyzyjnego) systemu biometrycznego
4. EER: błąd zrównoważony (ang. *Equal Error Rate*)
5. ZeroFNMR: najmniejszy możliwy stopień fałszywych dopasowań (FMR) dla którego FNMR=0
6. ZeroFMR: najmniejszy możliwy stopień fałszywych niedopasowań (FNMR) dla którego FMR=0

Przykładowe pytanie egzaminacyjne

Czy DNA byłoby dobrym identyfikatorem biometrycznym w zastosowaniu na bardzo dużą skalę, np. do identyfikacji podróżnych na lotniskach? Uzasadnij swoją odpowiedź.