A FUZZY LOGIKA ALKALMAZÁSA A MULTI-MODÁLIS BIOMETRIKUS AZONOSÍTÁSBAN

APPLICATION METHODS OF FUZZY LOGIC IN THE MULTIMODAL BIOMETRICAL IDENTIFICATION

WERNER GÁBOR PhD aspiráns
Óbudai Egyetem Biztonságtudományi Doktori Iskola
DR. HANKA LÁSZLÓ adjunktus
Óbudai Egyetem Bánki Donát Gépész- és Biztonságtechnikai Mérnöki Kar

Abstract

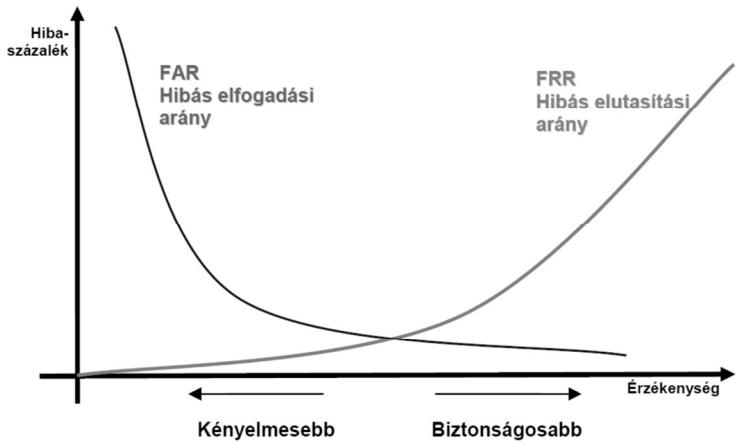
In this paper we presented an algorithm which can be applied in biometrical identification, especially wherever the high level of false rejection is significant. As the biometrics spreads, the basic difficulties of the controller algorithms are leading us to invent an appropriate technique, which is able to adapt and enough flexible to deal with the daily fluctuations. These daily changes may derive from the routines of the usage, the environmental impacts or even the natural lesions and changes of the user's individual biological identification patterns. The fuzzy logic as a soft computing method applies the artificial intelligence, which based on human thinking and behavior. The fuzzy logic involves the linguistic variables instead of exact numerical equations, and makes implications just like human logic. To highlight the results of our experiments we compared that with the classical mean value calculations.

1. Bevezetés

Korunkban a személyek azonosításának igénye számos helyen megjelenik. Habár az azonosítás maga is egy összetett, több igényt (sebesség, költség, pontosság, kinyert információ, stb.) kiszolgáló folyamat, nem feledkezhetünk meg annak egyik legfontosabb eleméről, az azonosítandó személyről. Ismeretes, hogy az úgynevezett humán faktor jelentős szereppel bír minden olyan rendszerben, ahol az ember-gép kapcsolat interakciója hatással van a kimenetre. A biometrikus azonosítás esetében különösen nagy kihívást jelent többek között a pontosság és ennek két szokásos mérőszáma a hibás elfogadások (False Acception Rate) és a hibás elutasítások aránya (False Rejection Rate), hiszen több olyan – nem biometrikus – azonosítási rendszer is létezik (tulajdon vagy tudás alapú) amellyel jól összemérhetőek ezen jellemző értékek.[1]

Vizsgálataink szempontjából mi elsősorban a hibás visszautasítás arányára fókuszáltunk, mert ennek értéke esetenként nem elhanyagolhatóan kicsi és igen erősen oszcillál. Ezzel együtt fontos megjegyezni, hogy a hibás elutasítás és a hibás elfogadás aránya általában fordított korrelációs viszonyban van, és az alkalmazott eszközre, valamint az azt használó populációra jellemző.

1. ábra. Hibás elfogadás és hibás elutasítás aránya



(forrás: [1])

Általánosan elmondható, hogy a hibás elfogadások aránya nem lehet szignifikánsabban nagyobb, mint más jogosultság ellenőrzési eljárásokban, viszont ebben az esetben a biometrikus eszközök általában igen magas hibás elutasítási tartományban dolgoznak. Ennek káros következménye jó esetben csak ingerültséget és torlódást okoz, rossz esetben az eszköz kizárását is magával vonzza.[2]

Korábbi vizsgálatainkban megnéztük, hogy a hibás visszautasítások száma hogyan vonatkoztatható a biometrikus azonosítás egyes elemi lépéseiből. Ennek ismertetéséhez fontosnak tartjuk röviden bemutatni a biometrikus azonosítás általános modelljét. Ahogy más azonosítási eljárásokban, az identifikációt meg kell, hogy előzze az adatbázis felépítése, de az, hogy, ebben milyen jellegű információ szerepel a modell tekintetében másodlagos kérdés. A tényleges azonosítás során első és kritikus lépésként az ember és gép közti interakciónak kell lezajlania, ami önmagában is potenciális hibákkal erősen terhelt. A továbbiakban már kvázi automatikusan történik a beolvasott mintából származó szükséges információ extrakciója és összevetése az adatbázisban tárolt sablonnal. Ezen automatizált lépések egy jól felépített algoritmust követnek, ahol a hiba lehetősége egyáltalán nem zárható ki, de bekövetkezési valószínűsége szignifikánsabban alacsonyabb, mint a mintavételi fázisban.[3]

A mintából kiolvasott információ minden esetben valamilyen matematikailag egzaktul megfogalmazható vektor egyes komponenseit kódolja. Ezen vektorokat nevezzük egyedi azonosító jegyeknek. Tulajdonképpen a klasszikus biometrikus azonosítás célja az, hogy megtaláljuk a kellő számú meggyező egyedi azonosító jegyet az adatbázisban tárol sablonokhoz képest. Biometrikus azonosítás módjától függetlenül is sokféle algoritmus létezik, és nem minden megoldásban lehet ennyire absztrakt módon megfogalmazni az egyedi azonosító jegyek vektoros leképezését, de a további számításaink bizonyításához fogadjuk el ezt a megközelítést.

A hibás elutasítás problematikája valahol ott kezdődik, amikor a felhasználóhoz tartozó egyedi azonosító jegy vektorjának komponenseit nem sikerült pontosan leolvasni, de akár visszavezethető egy olyan sablonra is, ami a későbbiekben beolvasott mintákkal általában eltérő, így ebben az esetben nem a minta beolvasása terhelt, hanem már maga a sablon eredendően hibás. Cikkünkben bemutatjuk, hogy miként lehet megbecsülni a hibás eseteket számát és milyen módon lehet csökkenteni a hibás visszautasítások arányát.

2.1. Hibás esetek meghatározása

A matematikai statisztikából ismerjük, hogy amennyiben egy [*n*] elemű mintában meg akarjuk határozni bizonyos hibás [*x*] esetek előfordulási valószínűségét [*P(x)*], akkor a binomiális tétel és a hibázás bekövetkezési valószínűségének ismeretében ezt egzakt módon kiszámolhatjuk.

$$\boldsymbol{P_{(x)} = \binom{n}{x} \cdot p^{x} \cdot (1-p)^{n-x} = f(x|n,p)}$$

Azonban amint szeretnénk meghatározni [*p*] hibázás bekövetkezésének valószínűségét nehézségbe fogunk ütközni, mert ez a [*p*] paraméter nemhogy nem konstans, hanem további két paraméterrel [$\boldsymbol{\alpha, \beta}$] jellemezhető eloszlást követ. Érdekes, hogy mind egyetlen felhasználó esetében, mind egy adott populáció tekintetében alkalmas ez, az úgynevezett béta-binomiális eloszlással számított modell. Így tehát meghatározhatóak az [$\boldsymbol{\alpha, \beta}$] paraméterek egyetlen felhasználóra az egyes azonosítások során bekövetkező elemi hibák és ennek következtében a tévesen komponált vektorok ismeretében, valamint egy teljes populációra, munkakörnyezetre az egyéni FRR értékek tükrében. A [*p*] valószínűségi paraméter sűrűség függvénye az alábbi:

$$\boldsymbol{p(\alpha, \beta) = \frac{1}{B(\alpha, \beta)} \cdot p^{\alpha-1} \cdot (1-p)^{\beta-1} = f(p|\alpha, \beta)}$$

Ahol a Béta-függvény a gamma függvény segítségével az alábbi módon fejezzük ki:

$$\boldsymbol{B(\alpha, \beta) = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)}}$$

Végezetül pedig, ahogy azt korábbi cikkünkben bizonyítottuk egy kvázi-Newton módszerrel és az Armijo-Goldstein feltételekkel meghatároztuk a poszteriori Béta-binomiális eloszlását az adott hibás esetek előfordulásának esetére:

$$\boldsymbol{P(x|n, \alpha, \beta) = \binom{n}{x} \cdot \frac{B(\alpha + x; \beta + n - x)}{B(\alpha, \beta)}}$$

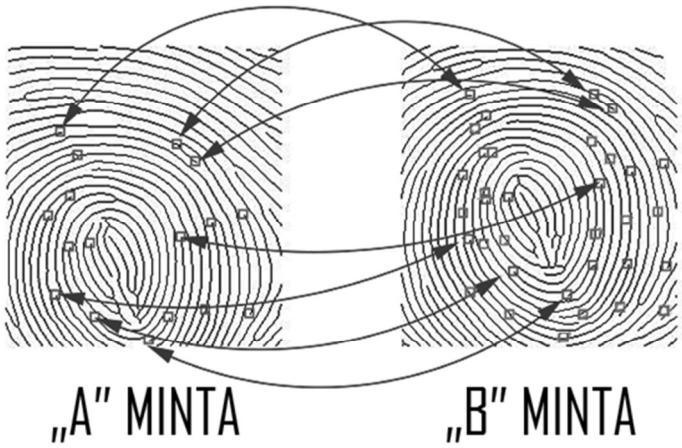
A módszer bevezetését és egy automatikus algoritmus fejlesztését eredendően az indokolta, hogy a piaci szereplők kezébe kerülhessen egy olyan metodika, amivel objektív módon hasonlíthatóak össze az egyes biometrikus azonosító eszközök az adott feladat és populáció tekintetében. Emellett viszont, sikerült megadni egy olyan módszert is, amivel az egyedi nem szisztematikus hibákkal terhelt összetett rendszerek, – így a biometrikus azonosító eszközök is – nap, mint nap küzdenek. A hiba előre becslési algoritmus időben akár folyamatosan is alkalmazható, így lehetőségünk nyílhat dinamikus folyamatok megismerésére és a hiba forrásának pontosabb körülhatárolására.[4]

2.2. Hibák szuperpozíciója

A béta-binomiális hibabecsléssel elvégzett számítások azt mutatták, hogy a hibák előfordulása általában vagy sztochasztikus, vagy visszavezethető valamilyen determinisztikus hatással bíró forrásra. Következésképpen az elemi hibák előfordulása egyes esetekben – amelyek száma nem infinitezimálisan kicsi – nem független egymástól. Az egyes mintákban az elemi hibák mint az egyedi azonosító jegyeket kódoló vektorok komponenseinek biológiai sajátossága, együttesen térnek el a sablonban szereplő értéktől. Így tulajdonképpen a vektor abszolút hibája ezen elemi hibák szuperponáládásból fakad. A jelenség a gyakorlatban azért okoz problémát, mert a mérnökök igyekeznek a biometrikus azonosító algoritmusokban meghagyni egy bizonyos rugalmasságot, növelve a hibatoleranciát, azonban abban az esetben, ha több elemi hiba szimultán fordul elő, akkor az ezek szuperpozíciójaként keletkező hiba már olyan mértékű, hogy a rendszer –tévesen ugyan–, de el fogja utasítani a felhasználót.

A béta-binomiális eloszlással tehát bizonyítható, hogy a az adott mintában előforduló hibák száma függ a [p] valószínűségi paramétertől, és amennyiben az elemi hibák értéke szignifikáns, akkor azok szimultán fognak fellépni, és halmazati hatásuk eredményeként erősen rontják az eszközök hatásfokát. Következésképpen a további a cél az, hogy amennyiben jelentkeznek is hibák az azonosítás során – természeti rendszer révén pedig ez nem zárható ki – azok értéke nem szuperponálódjon. Ez pedig úgy biztosítható, hogy [p] valószínűségi paraméter eloszlását megpróbáljuk variálni. Ahogy azt az előzőekben írtuk a béta eloszlás paraméterei elsősorban egy adott minta forrást vagy populációt jellemeznek, tehát ha biztosítani szeretnénk a hibák halmozódásának minimalizálását, meg kell keresni azokat a mintaforrásokat és/vagy módszereket amelyekkel különböző [$\boldsymbol{\alpha, \beta}$] paraméterű béta eloszlást kapunk.

2. ábra. A hibák szuperpozíciója



(forrás: saját szerkesztés)

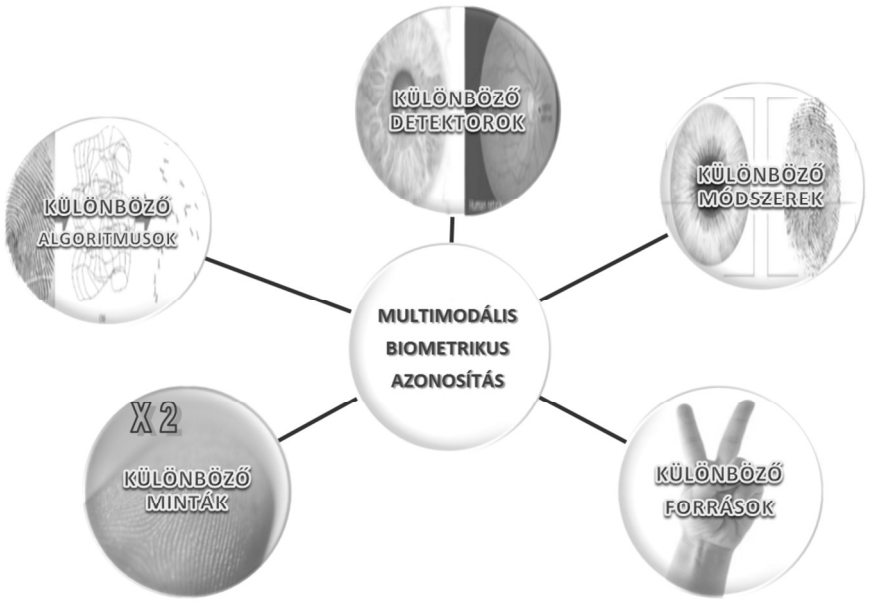
Ahogy a fenti ábrán is látható, az „A” mintához képest eltolt és kissé elforgatott „B” minta összes egyedi azonosító jegye valamilyen torzítást szenvedett. Egy ilyen egyszerű torzítást természetesen minden algoritmus képes tolerálni, amíg a torzítás hatása a mintát globálisan érinti. Amennyiben viszont a hatás lokális, illetve a torzítás háttere kissé összetettebb (dehidratáció, sérülés, koszolódás, stb.) akkor hibatoleranciát nehezebb megvalósítani.

2.3. Hibacsökkentés kombinációs módszerekkel

Beláttuk, hogy amennyiben az egyetlen mintából érkező információt vesszük alapul, akkor az abban rejlő esetleges hibák nem függetlenek egymástól, tehát egyetlen minta elemzése erősen korlátozza a rendszereink robusztusságát és hibatoleranciáját. Az azonosítási eljárások ma már egyre szélesebb körben alkalmazzák a multi-faktoros azonosítás, amely nem keverendő össze a multi-modális biometrikus azonosítással. Az előbbi elsősorban egy logikai biztonsági rendszer, miszerint a ismert, a birtokolt és a viselt alapú kódok valamilyen kombinációja hivatott gondoskodni a magasabb védelem fenntartásáért, míg utóbbi kizárólag a viselt alapú kódok, tehát a biometrikus azonosítási eljárások valamilyen kombinációját jelenti.[5]

A multi-modális azonosítás során általában öt kategóriába sorolhatjuk a tipikus megoldási módszereket. Amely kategóriák önkényesek ugyan és nem minden megoldás sorolható egyértelműen egyik vagy másik csoportba, de segítik a megértést és az alkalmazást. Ezen kategóriák az alábbiak:

3. ábra. Multi-modális biometrikus azonosítás módszerei



(forrás: saját szerkesztés)

A kombinációs lehetőségek során választható, hogy milyen testrészről, milyen típusú detektorral és milyen algoritmussal vegyünk mintát. Az, hogy melyik megoldási módszert válasszuk több tényező együttes figyelembe vételét igényli, úgy mint a beléptetésre rendelkezésre álló időkeret, munkakörülmények, és természetesen az anyagi források.

A multi-modális azonosítás azonban jelenleg csak korlátozottan használható, mert az egyes azonosítási módok algoritmusai és adatai nehezen fuzionálhatóak. A fúziónak az OSI modellhez hasonlóan több szintjéről tehetünk említést, de minden esetben végső soron meg kell találni azt a platformot, ahol a különböző információk összefésülhetőek és együttesen kezelhetőek. Könnyen belátható, hogy két párhuzamosan alkalmazott biometrikus eszköz jelentősen javítja a védelmi szintet, de egyáltalán nincs hatással a hibatoleranciára, hiszen a köztük lévő logikai kapcsolat kizárólag a kettős értékű, Boole féle logikát követi.[6, 7]

Ennek megfelelően a hatékonyabb működés érdekében olyan algoritmust és mögöttes matematikát kell alkalmazni, ami kellőképen tudja fuzionálni a különböző detektorok adatait. A gyakorlati megvalósítás tekintetében azonban továbbra is nehézséget jelent, hogy az egyes biometrikus azonosító eszközök adatait és forrásfájlait általában nem tudjuk kinyerni, így a továbbiakban bemutatott modellben az egyéni azonosítójegyek egy általános azonosítási metodikáját vettük alapul.

3.1. Fuzzy logika alapjai

A lágy számítási módszerek (soft computing) és így a mesterséges intelligenciát megvalósító matematika közé sorolt fuzzy logika mintegy ötven éve, Lotfi A. Zadeh 1965-ös „Fuzzy Sets" című cikke óta van jelen a számítástechnikában és a robotikában. Az elmosódott halmazok logikájának alapja az, hogy nem binárisan kódolt (igen/nem) módon értelmezi a változókat, hanem valamilyen tulajdonságot egy részekre bontott halmazrendszer ír le aszerint, hogy az adott tulajdonság egyes jellemzői mennyire tekinthetőek igaznak. Zadeh gondolata abból indult ki, hogy például relatíve könnyen meg tudjuk mondani pillanatnyi hőérzetünk alapján mennyire tekintjük hidegnek, melegnek, vagy akár forrónak a környezetünket, és ezt nagyjából időben állandó pontossággal, szemben azzal, mintha egzaktul a Celsius skálán kellene megbecsülni a hőmérsékletet.[8, 9]

Olyan összetett rendszerekben, ahol több változó, például több biometrikus eszköz és nem lineáris következtetési metodika (halmozott béta-binomiális eloszlású hibák) vannak jelen, ott a fuzzy logika alkalmazása minden bizonnyal hamarabb és pontosabb megoldásra vezet, mint más algebrai vagy klasszikus statisztikus modell.[10]

3.2. Fuzzy logika alapú multi-modális biometrikus modell

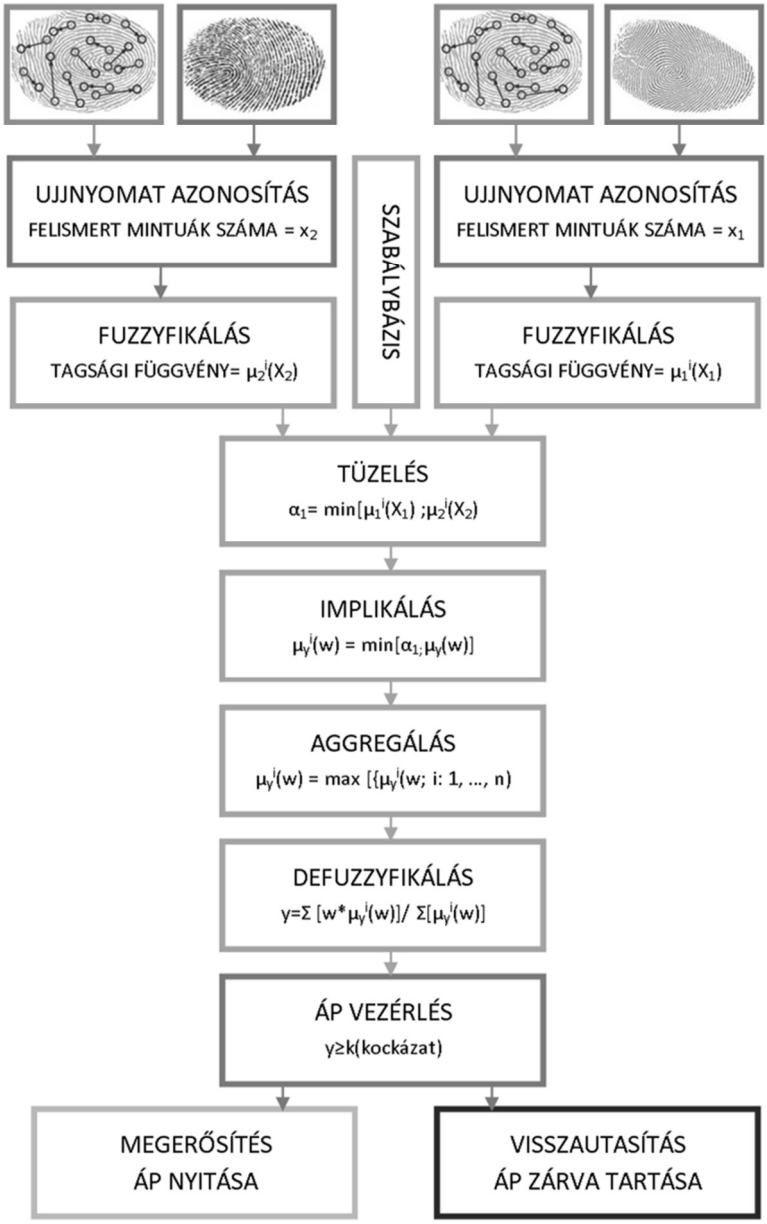
Az alább ismertetett modellben olyan fuzzy logikai vezérlőt alkottunk meg, ami két ujjnyomat olvasó eszközből kiolvassa a helyesen azonosított egyéni azonosítójegyek számát, és a korábban optimalizált szabálybázis alapján automatikusan dönt az eredmény elfogadásáról vagy elutasításáról.

A megalkotott algoritmus részletes matematikai jellemzése helyett annak funkcióit érdemes kiemelni. Így például fontosnak tarjuk megjegyezni, hogy a fuzzy logikai vezérlő (FLC) számos olyan beállítással rendelkezik, amelyekkel szükségképpen optimalizálható a működés. Egy egyszerű FLC esetén ezen beállításokat tapasztalati úton, az irodalmi adatokra támaszkodva és folyamatos, kvázi mechanikus programozással kell megtalálni, de összetettebb vezérlő esetén vannak módszerek az automatikus optimalizálás megvalósítására is, például a szélesebb körben ismert ANFIS vezérlés (Adaptive Neuro-Fuzzy Inference System).[11]

Az FLC vezérlőben meghatározhatjuk az úgynevezett tagsági függvények számát és alakját, ami az egyes tulajdonságok pontosab meghatározását segíti elő, azonban az egyik leglényegesebb beállítás maga a szabálybázis és az implikáció szabálya. Többféle vezérlés ismert, köztük a leghíresebbek a Takagi-Sugeno és a Mamdani típusúak. Mi jelen esetben az utóbbi választottuk, ami magában foglalja, hogy milyen t-normákat és t-konormákat kell alkalmazni. A szabálybázis által kódolt kimeneti függvényeket valamilyen módon összegezni kell, és végezetül defuzzyfikálni, amelyre szintén több módszer ismeretes.

Az általunk alkalmazott modell blokksémája és az egyes matematikus műveletek az alábbiak:

4. ábra. Fuzzy logikai vezérlő modellje

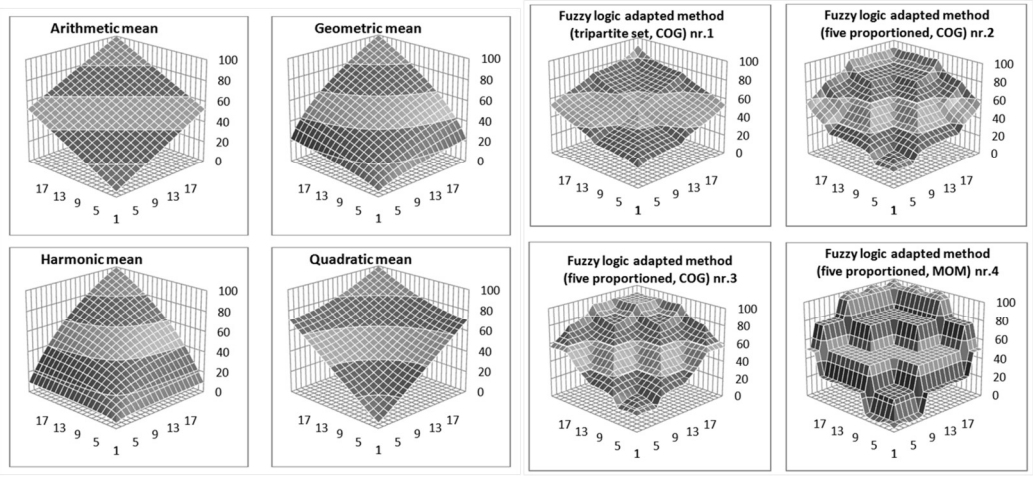


(forrás: saját szerkesztés)

4. Eredmények

Az FLC vezérlési algoritmust lefutattuk az összes lehetséges bemeneti érték kombinációjára, hogy összehasonlíthassuk a kapott eredményt más matematikai módszerekkel. Ha ábrázoljuk a legfeljebb 20-20 felismerhető egyéni azonosító jegy összesített értékelését, akkor egy összetett felületet kapunk. Ezt a felületet összevetve a klasszikus statisztikai matematikai középértékekkel számolt felületekkel jól kivehetőek a különbségek:

5. ábra. A klasszikus statisztikai középekkel és az FLC vezérlővel kapott felületek



(forrás: saját szerk.)

Míg a középértékekkel kapott felületek simák és folyamatosak, addig a jobb oldalon látható minták arról tanúskodnak, hogy az egyes esetek jól elválaszthatóak egymástól, és ezek az elválások jól programozhatóak. A szabálybázis és egyes beállítások optimális, akár automatikus, adaptív változtatásával elérhetővé válik a nagyobb hibatolerancia a hibával jobban terhelt detektorral, mintákkal, vagy az adott biometrikus azonosítási móddal szemben, úgy hogy a biztonsági szint eközben nem csökken. Következésképpen bizonyosodott, hogy az FLC vezérlők a multi-modális biometrikus azonosításban képesek a biometrikus azonosítás eredendő problémájára – miszerint a hibatolerancia és pontosság fordítottan arányos – egy lehetséges megoldást nyújtani.

Jegyzetek

1. M. I. O. C. Kovács Tibor, A Biztonságtudomány Biometrai Aspektusai, Pécs: Pécsi Határőr Tudományos Közlemények, XIII. kötet, HU ISSN 1589-1674, 2012.
2. A. R. Anil K. Jain, Multibiometrics Systems, COMMUNICATIONS OF THE ACM, 2004/Vol. 47, No. 1 , 2004.
3. O. Csaba, Classification of Biometric Access Controll Systems Based on real-time Throughput, Pozsony, Szlovákia: REVIEWED PROCEEDINGS Fifth International Scientific Video-conference of Scientists and PhD. students or candidates: Trends and Innovations in E-business, Education and Security. 129 p. ISBN 978-80-225-4191-6, 2015.
4. H. L. Werner Gábor, Using the Beta-Binomial Distribution for the Analysis of Biometric Identification, Subotica, Serbia: SISY 2015 IEEE 13th International Symposium on Intelligent Systems and Informatics ISBN 978-1-4673-9388-1, 2015.
5. A. C. D. S. S. Vaithyasubramanian, TWO FACTOR AUTHENTICATIONS FOR SECURED LOGIN IN SUPPORT OF EFFECTIVE INFORMATION PRESERVATION AND NETWORK SECURITY, India: ARPN Journal of Engineering and Applied Sciences, vol. 10, no. 5, ISSN 1819-6608, 2015.
6. F. P. J. C. X. Anne M.P. Canuto, Investigating fusion approaches in multi-biometric cancellable recognition, Brazil: Elsevier, Expert Systems with Applications 40 (2013) ISSN 1971–1980, 2012.
7. J. Y. Xuhua Liu, Fuzzy Boolean Algebra and its Properties, China: Jilin University, Changchun.

8. L. A. Zadeh, Fuzzy Sets, ELSEVIER, Information and Control, 8 (3) 338-353., doi:10.1016/S0019-9958(65)90241-X, 1965.
9. D. T. L. Kóczy, Fuzzy Rendszerek, Budapest: Typotex, ISBN: 978-963-2797-09-0, 2012.
10. Uo.
11. J.-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, USA: Dept. of Electr. Eng.& Compit. Sci., California Univ, Berkeley, p. 665-685, DOI: 10.1109/21.256541 , 2002.

Felhasznált irodalom

A. C. D. S. S. Vaithyasubramanian, TWO FACTOR AUTHENTICATIONS FOR SECURED LOGIN IN SUPPORT OF EFFECTIVE INFORMATION PRESERVATION AND NETWORK SECURITY, India: ARPN Journal of Engineering and Applied Sciences, vol. 10, no. 5, ISSN 1819-6608, 2015.

A. R. Anil K. Jain, Multibiometrics Systems, COMMUNICATIONS OF THE ACM, 2004/Vol. 47, No. 1 , 2004.

D. T. L. Kóczy, Fuzzy Rendszerek, Budapest: Typotex, ISBN: 978-963-2797-09-0, 2012.

F. P. J. C. X. Anne M.P. Canuto, Investigating fusion approaches in multi-biometric cancellable recognition, Brazil: Elsevier, Expert Systems with Applications 40 (2013) ISSN 1971–1980, 2012.

H. L. Werner Gábor, Using the Beta-Binomial Distribution for the Analysis of Biometric Identification, Subotica, Serbia: SISY 2015 IEEE 13th International Symposium on Intelligent Systems and Informatics ISBN 978-1-4673-9388-1, 2015.

J. Y. Xuhua Liu, Fuzzy Boolean Algebra and its Properties, China: Jilin University, Changchun.

J.-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, USA: Dept. of Electr. Eng.& Compit. Sci., California Univ, Berkeley, p. 665-685, DOI: 10.1109/21.256541 , 2002.

L. A. Zadeh, Fuzzy Sets, ELSEVIER, Information and Control, 8 (3) 338-353., doi:10.1016/S0019-9958(65)90241-X, 1965.

M. I. O. C. Kovács Tibor, A Biztonságtudomány Biometrai Aspektusai, Pécs: Pécsi Határőr Tudományos Közlemények, XIII. kötet, HU ISSN 1589-1674, 2012.

O. Csaba, Classification of Biometric Access Controll Systems Based on real-time Throughput, Pozsony, Szlovákia: REVIEWED PROCEEDINGS Fifth International Scientific Video-conference of Scientists and PhD. students or candidates: Trends and Innovations in E-business, Education and Security. 129 p. ISBN 978-80-225-4191-6, 2015.