

Reprezentációs és felismerési módszerek párhuzamos és nem szokványos gépeken

Stubendek Attila

PhD disszertációi tézisei



Témavezetők: Dr. Szolgay Péter, dr. Karacs Kristóf

Pázmány Péter Katolikus Egyetem

Információs Technológiai és Bionikai Kar

Roska Tamás Műszaki és Természettudományi Doktori Iskola

Budapest, 2018.

1. Bevezetés

Megtervezni és elkészíteni a gépet, hogy segítse az embert, civilizációnk egyik legfontosabb kihívása és eredménye. Évezredek óta a mozgó és a mozgató eszközök fejlesztésével egyre kevesebb terhet kellett az embernek a vállán viselnie, több időt szentelhetett a művészetnek és a gondolkodásnak, hogy aztán megszülethessenek a számolásban, a logikában, sőt a következtetésben is segítséget nyújtó gépek. Civilizációnk fejlettségét mára már nem csak az mutatja, hogy mire képes az ember, hanem hogy mire képesek az általa létrehozott és megtanított gépek.

A mai mérnöki tervezési feladatok az emberi gondolkodás, felismerés, tanulás képességének utánzását tűzik ki célul, mindazt, ami minket a teremtés koronájává tudott emelni. Mindeközben a tudományos társadalom újra és újra találkozik azzal a felismeréssel, hogy az élet, a természet mintái és szerveződése, Isten csodálatos munkája a tökéletes inspiráció a legegyszerűbb és legbonyolultabb eszközök elkészítésére egyaránt.

A látás az ember legalapvetőbb érzékelése, így a külvilággal minimálisan kölcsönhatásba lépő gépek esetében is szükség van a látottak értelmezésére. A gépi látásnak számos olyan alkalmazása van, amely megkönnyít az ember mindennapi életét: Az egészségügyben segíti az orvosok munkáját, olyan helyek kutatását teszi lehetővé, ahová ember nem juthat el, vagy veszélynek lenne kitéve, az iparban gyorsítja a termelést és automatizáltan végzi a minőség ellenőrzését, önvezető autókkal segít tisztává tenni környezetünket. A gépi látásra építve az életminőséget nagyságrendekkel javító segédeszközök hozhatók létre a látássérültek számára.

A *gépi képtelmezés* több tudományterület módszereit is alkalmazza. A hatékony működés alapja a képfeldolgozás, a kép – az aktuális feladatnak megfelelő – előfeldolgozása és a matematikai tulajdonságainak a kinyerése, a *képleírás*, objektumreprezentáció előállítása. A leírás szolgál a későbbiekben arra, hogy bizonyos tulajdonságokat elemezni lehessen, és össze lehessen hasonlítani a különböző képek vagy objektumok reprezentációit egymással. Ezért a leírót úgy kell megválasztani, hogy azon tulajdonságokat fejezze ki, amelyek az elvégzendő feladat szempontjából alkalmas a különbségek kiemelésére, illetve a nem releváns információk elhanyagolására.

Az *objektum-felismerés* eredménye matematikai megfogalmazásban egy néhány dimenziós vektor, amely lehet egy *vizuális kérdés*re adott válasz, így egy objektum jelenléte, pozíciója vagy részletes típusa. A feladatban megfogalmazott kérdés *döntést* igényel, amelyet a reprezentáció alapján lehet meghozni. A következtetés történhet a leírás elemzése alapján, vagy ismert reprezentációkkal való *összehasonlítás* alapján.

A gépi látás folyamatában alkalmazott algoritmusok és architektúrák egyaránt tudatosan közelítenek az emberi látásban és információfeldolgozásban megfigyelt módszerekhez. A klasszikus számítási architektúrák ugyan univerzálisak, ugyanakkor az a konkrét feladat igényeinek megfelelő, gyors működéshez és az alacsony fogyasztáshoz speciális, dedikált architektúrák szükségesek, sok számítási egységgel és a *lokalitás precedenciájával*. A tervezés során ezért azt is figyelembe kell venni, milyen architektúra áll rendelkezésre, és milyen egyéb igények vannak a válaszidőre és a fogyasztásra, illetve a súlyra és a méretekre vonatkozólag.

Munkám során az általános használhatóságra törekedtem, ugyanakkor egy konkrét alkalmazási terület, egy, a látássérülteket segítő algoritmus- és alkalmazásgyűjtemény, a Bionikus Szemüveg elkészítése, mint motiváció és

tesztelési lehetőség, végigkísérte az algoritmusok fejlesztését. A Bionikus Szemüveg a Magyar Bionikus Látásközpont egyik kutatása projektje. Az algoritmusokkal szemben támasztott követelményeket, célokat és kényszereket jól behatárolták egyrészt a későbbi felhasználók által igényelt funkciók, másrészt a rendelkezésre álló hardverkörnyezet. A felismerendő objektumok, piktogramok, bankjegyek mind kétdimenziósak, és merevek, így az algoritmusokat is az ilyen típusú alakzatok felismerésére koncentráltam.

A kép vagy képrészlet meghatározó tulajdonsága annak *alakja*. A háromdimenziós objektum kétdimenziós képre *képeződik le*, ezáltal az eredmény függ az objektum és a kamera pozíciójától. Egyrészt a háromdimenziós objektum különböző oldalai eltérőek lehetnek. Másrészt a látószögtől és a kamera, illetve az objektum relatív pozíciójától is függ a leképezés. A kétdimenziós – „lapos” – objektumok esetén az előbbi probléma ugyan nem áll fenn, viszont a leképezés ez esetben továbbra is függ a kölcsönös pozíciótól. Ezen felül az objektumok bizonyos részleteinek alakja változhat, illetve lehetnek eltérések egy objektumosztályon belül is – gondoljunk csak a cseresznye és a gesztenyefa leveleire, ugyan eltérő alakúak, de egyaránt a falevelek osztályba tartoznak. Ezért az alakfelismerés ritkán korlátozódhat sablonszerű illesztésre, bizonyos részletek jelenlétének vizsgálatára. Az ilyen típusú eltérések, illetve a diszkriminatív tulajdonságok megfelelő kiemelése a leírás feladata.

Az alak legegyszerűbb *matematikai leírói*, mint például a kerület, terület, excentricitás, csupán egy vagy néhány skalárt tartalmaznak, amivel ugyan nagyon tömören fejezik ki az alak tulajdonságát az adott aspektusból, de finomabb különbségek kimutatására ezek az adatok alkalmatlanok. A *komplex alakleírók* alapvetően két családba sorolhatók. Az élalapú leírók az alakzat kontúrját írják le, a terület-alapú deskriptorok az alak összes pontját veszik

számításba. A két módszer egy komplex leíróban való használata ritka, ugyanakkor a szakirodalom említ néhány összetett deskriptort.

A fenti felosztáson túl megkülönböztetünk globális és lokális leírókat, aszerint, hogy a leíró egy-egy értékére a bemenet egy része, vagy egésze van hatással.

Az emberi mindennapok során látott objektumok alakjai összetettek, szinte lehetetlen olyan egységes nézőpontot kiválasztani, amely szerint az összes alakzat egyértelműen bekegategorizálható lenne. Ezért célszerű egyszerre *több modalitást* is figyelembe venni, hogy ezáltal robosztusabb felismerést lehessen elérni.

Munkámban *összetett deskriptor* megalkotására tettem kísérletet, ahol a primitív, majd egyszerű, területalapú leírók alapján hozzávetőlegesen is össze lehet hasonlítani az alakzatokat, majd részletes következtetést az élalapú tulajdonságokból lehet levonni.

Előzetes mintákból gyűjtött tapasztalat alapján megalkotott modell használata esetén *példányalapú tanulásról* beszélünk. Amennyiben a rendelkezésre álló minták esetében ismertek a megállapítani kívánt értékek, *felügyelt tanulásról* van szó.

Példányalapú döntési modell például a legközelebbi-szomszéd modell, a kernelmódszerek, illetve a neurális hálózatok. Abban az esetben, amikor az előrejelzendő adat diszkrét, és szemantikus információval bír, *osztályozásról* beszélünk. A legegyszerűbb osztályozás bináris, ilyen esetben a kimenet értékészlete igaz vagy hamis, illetve ezekkel bijektív bármilyen kételemű tér. Ez jelentheti egy konkrét tulajdonság meglétét vagy hiányát, de reprezentálhat olyan teret is, amelyben a vizsgálni kívánt szempont szerint kétféle elem létezik.

Az osztályozás értékkészlete lehet több osztály is. Feladattól függően az osztályok száma nem feltétlenül korlátos. Ez esetben megkülönböztetünk a feladat szempontjából *releváns és nem releváns osztályokat*. Ez a szempont azért különösen fontos, mert minőségileg más feladatról van szó, ha biztosan tudott, hogy a bemenet néhány adott osztályból való, illetve, ha a várt osztályokon kívül még bármilyen más osztályhoz is tartozhat a bemenet. Valós képek alapján történő objektum-felismerés esetében a bemenet tulajdonságtere a legkritikább esetben korlátos. Dolgozatom második részében a legközelebbi-szomszéd osztályozás legfőbb hiányosságainak pótlását tűztem ki célul, amelynek köszönhetően nem releváns osztályokat tartalmazó térben egyaránt használhatóvá válik a módszer, illetve lényegesebben gyorsabb működés érhető el a tulajdonságok fokozatos összehasonlításával.

Mind a leíró, mind az osztályozó tervezésekor az egyik fő motiváció a *célhardvereken* való hatékony futtatás volt. A feladatokból következően célul tűztem ki, hogy az algoritmusok *kiloprocesszoros architektúrákon* nagyon alacsony fogyasztás mellett működjenek, ugyanakkor kevés magos, számítógép vagy telefon digitális architektúráján futtatva is elfogadható legyen a feldolgozás sebessége.

A processzorokban alkalmazott digitális számítási építőelemek sebessége közelít a fizikai határukhoz. Az architektúrák fejlődésére vonatkozó tapasztalati Moore-törvénnyel lépést tartva több processzort alkalmazó, párhuzamos architektúrák fejlesztése kezdődött az elmúlt évtizedben. Emellett viszont megnőtt az érdeklődés az alternatív technológiák iránt is, a korábban alkalmazott műszaki keretek lebontására. Az egyik ilyen fő kutatási ág a *nem booli alapú*, nem digitális számítási egységek fejlesztése, azon belül pedig a *nanomágnesek hálózatának* kutatása. A mágnesek árammal gerjesztve *oszcillátorként* működnek, amelyek kölcsönhatnak egymással, a rendszer állapota az egyes

oszcillátorok dinamikájától, a gerjesztés mintázatától, illetve az oszcillátorok topológiájától függ. Munkámban az ilyen oszcillátorhálózatok számítási egységként való használatára tettem kísérletet, definiálva az adat, illetve a program fogalmait, emellett kísérleteket hajtottam végre arra vonatkozólag, hogy milyen típusú számítások esetén bizonyul hatékonnak az oszcillátorhálózatokból felépülő architektúra.

2. Módszertan

Az algoritmusokat, az azokhoz kapcsolódó tesztek és kísérleteket legtöbbször Matlab¹ környezetben implementáltam. A vizuális feldolgozás alacsony szintjein (mátrixműveletek, konvolúció, morfológia – propgen, nyomatékszámítás) a Matlab beépített feldolgozóit is alkalmaztam, magasabb szinteken azonban saját implementációkat készítettem a szélesebb körű megismerés és változtathatóság végett.

Munkám alapjaihoz, az alapvető módszerek alkalmazhatóságának kipróbálására hasznos segédeszköznek bizonyult a WEKA² implementáció-gyűjtemény, későbbiekben pedig az előbbit is magába foglaló RapidMiner³ gyűjtemény illetve tesztkörnyezet. Az eredmények alapján saját implementációt készítettem legközelebbi-szomszéd, neurális hálózat döntési módszerekhez illetve általános genetikus algoritmushoz.

A Bionikus Szemüveg bankjegyfelismerő applikációját kollégáimmal látássérültekkel teszteltük. A tesztek elején a látássérültek tájékoztatást kaptak az eszköz működéséről, majd segítség nélkül, csak az eszköz válaszára támaszkodva kellett bankjegyeket azonosítaniuk. A tesztek során az összes feldolgozott képkockát és a hozzá tartozó döntéseket eltároltuk, így a méréseket később szimulációs környezetben meg tudtuk ismételni, mérve ezzel az algoritmusok fejlődését. Az élő tesztek során kézi feljegyzéseket is készítettünk, amelyekkel egyéb, a programra, a környezetre, a használhatóságra vonatkozó tapasztalatokat gyűjtöttük össze.

¹ <http://www.mathworks.com/products/matlab/>

² <http://www.cs.waikato.ac.nz/ml/weka/>

³ <https://rapidminer.com/>

3. Új tudományos eredmények

1. Tézis

Megterveztem egy új általános alakleírót, a Globális Statisztikai és Vetített Fő Élleírót (Global Statistical and Projected Principal Edge Distribution), amely több független jellemző leírásán alapul: ötvözi a globális matematikai tulajdonságok és az élalapú leíró előnyeit. A deskriptort implementáltam és hatékonyságát kísérletileg igazoltam.

Kapcsolódó publikációk: [A1][A3][A4] [A7]

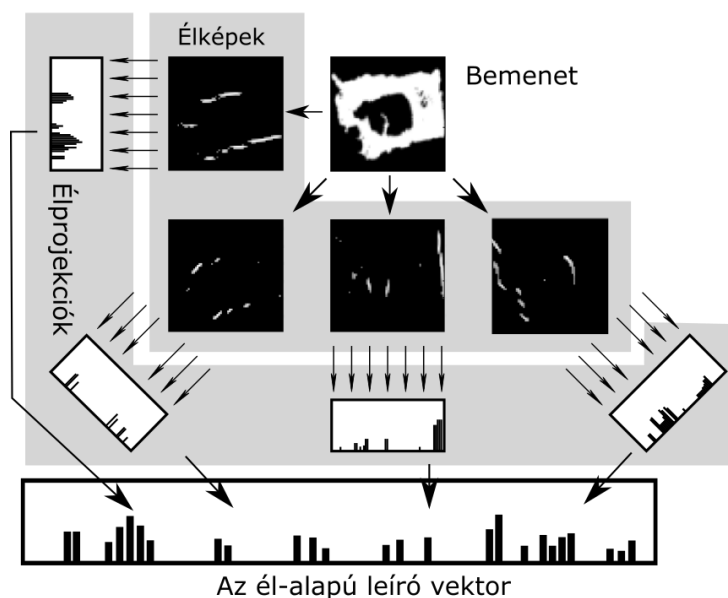
A leíró globális matematikai alakleírókból és egy élalapú, lokális leíróból áll. A többféle alakleíró egy leíróba való ötvözése azt a célt szolgálja, hogy az alak különböző jellemzőit – belső terület, határvonal – egymástól függetlenül, a lehető legteljesebben le lehessen írni, és a döntést meggyorsítani.

A tulajdonság-vektor fejlécét, majd az azt követő nyolc elemét matematikai és statisztikai leírókból állítottam össze, amelyek az összehasonlítás alapján történő hasonlóságmérés során gyorsabb döntést és jobb fedést eredményeznek.

Az első tulajdonság az alakzat excentricitása, amely a fő és melléktengely hosszának arányát fejezi ki, és az alakzat elnyúltságát jellemzi egy skalár értékkel. A második tulajdonság az alakzat határoló téglalapjának és az alakzat területének aránya. A következő nyolc jellemző az alakzat két tengely szerinti hisztogramjának 4-4 statisztikai nyomatéka.

A leíró második része élalapú alakleíró (Extended Projected Principal Shape Edge Distribution, EPPSED), és az alak lokális információit fejt ki. Az algoritmust a PPED képleíró inspirálta a principális élek kiválasztását és leképezését illetően. (Ábra 1.)

A felismerő bemenete egy, az alakzatot tartalmazó bináris kép, amelynek az előállítása egy alakkifejtő (szegmentáció, mintakifejtő) algoritmus feladata.



Ábra 1. Az EPPSED él-alapú leíró vázлата. A bináris bemenetből konvolúcióval négy élkép képeződik. Az élképek a generálásukkal megegyező irányban levetítésre kerülnek, majd normalizálva, egymás után fűzve alkotják az él-alapú leíró vektort.

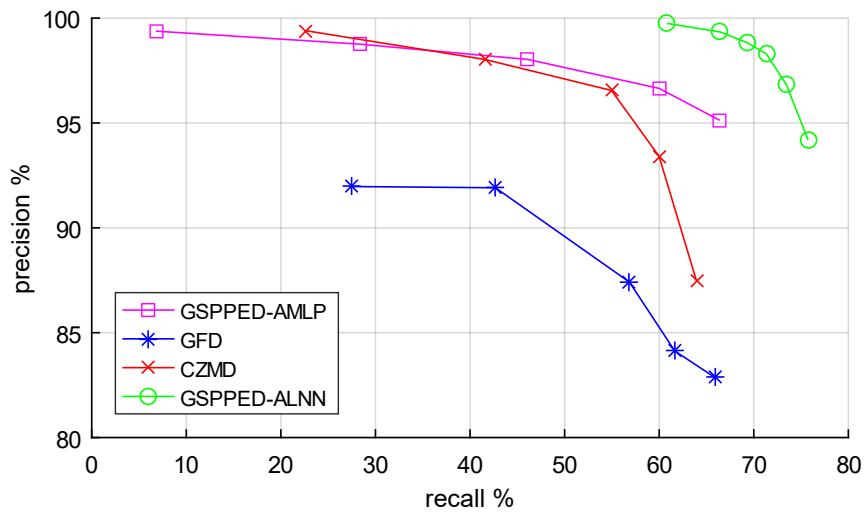
Az alakzataból 5x5-ös konvolúciók segítségével, négy irányban négy élkép készül, amelyeken minden pixelpozícióban kiemeljük a legfontosabb éleket a négy irány közül. A kiemelés lépcsőfüggvény helyett soft-threshold függvénnyel történik, amelynek alkalmazása robusztussá teszi az algoritmust: egyrészt csökkenti a túllillesztést, másrészt segíti az átjárhatóságot különböző architektúrák, illetve azok eltérő számreprezentációi között. Kísérletileg igazoltam, hogy a folytonos függvény használata javít a felismerés eredményén.

A forgásinvarianciát a folyamat során két ponton lehet biztosítani. Az egyik lehetőség a leíró generálásának módjában rejlik, ahol vagy közvetlenül a

leíró független az aktuális orientációtól, vagy olyan előfeldolgozás hajtódik végre, amelynek eredményeképpen normalizált orientációba kerül az alak. A másik lehetőség a forgásinvariancia osztályozás során történő biztosítása, ahol hasonló, de különbözőképpen elforgatott egyedek azonos osztályba kerülnek. Ez történhet adekvált metrika vagy döntés alkalmazásával vagy a forgatásokat tekintve redundáns tanítóhalmazzal. A PPED családjába tartozó leírók önmagukban nem forgásinvariánsak, ezért redundáns adatbázissal oldják meg a problémát. Habár ez a felfogás közelebb áll a humán érzékeléshez, átláthatatlanná nagyságrenddel nagyobbá tenné az adatbázist, ezért én a forgásnormalizálás megoldását választottam. Az alak nyomatékaiból kiszámítható orientációt veszem alapul, és az alakot úgy forgatom, hogy a főtengelye vízszintes legyen. A 180 fokos redundanciát úgy kezelem, hogy a súlypontot balra helyezem.

A skálainvarianciát a vetítési keret alapjául szolgáló 64×64 pixeles képméretre történő átméretezéssel, normalizálással érem el. A pozícióinvarianciát az egyik dimenzióban az alak középre helyezésével érem el, úgy, hogy a másik dimenzióban az alak érinti a keret mindkét szélét.

A principális éleket jellemző leíró generáló algoritmus alkalmas dedikált architektúrán vagy celluláris hullámszámítógépen való párhuzamos, nagy sebességű futásra.



Ábra 2. A GSPPED alakleíró összehasonlítása más alakleírókkal.

Az általam tervezett alakleírót összehasonlítottam több, széleskörűen alkalmazott alakleíróval magyar Forint bankjegyekről kinyert, illetve zajjal szennyezett alakzatokon. Az eredmények azt mutatják, hogy a GSPPED túlteljesíti mind a komplex Zernike nyomatékok alapján összeállított alakleírót, mind az általános kétdimenziós Fourier-transzformáción alakuló deskriptort. Az eredményeket a 2. Ábrán foglalom össze.

2. Tézis

Megterveztem és implementáltam egy kétszintű, reprezentáns egyedekkel történő, dedikált metrikával számolt összehasonlításon alapuló, nem korlátos osztályszámú osztályozási algoritmust több független algoritmusból származó tulajdonságvektorok alapján történő döntésre.

Kapcsolódó publikációk: [A1][A3][A4]

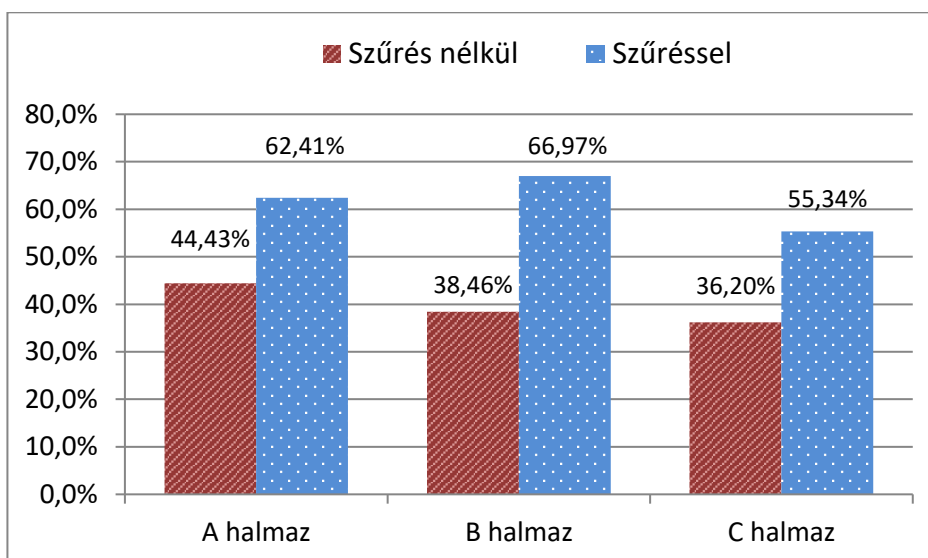
A többlépcsős döntés igazodik a leíró szemantikájához és a leíró részeinek diszkriminatív erejéhez. Az osztályozási folyamatban előbb a matematikai és statisztikai alakleírók alapján történő szűrés történik, majd korlátos legközelebbi szomszéd algoritmussal (Adaptive Limited Nearest Neighborhood, ALNN) az élelapi tulajdonságok kerülnek összehasonlításra, amelyben az egyedek közötti távolság dedikált metrika alapján kerül meghatározásra.

Módszerem lényege a többrészes leíró többlépcsős felhasználása, amely a döntés gyorsítását és fedésének növelését is eredményezi. A reprezentatív egyedekkel való összehasonlítás során először az alapvető matematikai tulajdonságok kerülnek összehasonlításra, amely részletekről egyértelműen kiderül, hasonló objektumról van-e szó. Biztos döntés ugyanakkor az alapvető tulajdonságok alapján nem várható, viszont a módszer alkalmazásával csupán néhány tulajdonság alapján történő összehasonlítás elegendő lehet arra, hogy a nyilvánvalóan különböző reprezentációs halmazbéli elemeket elutasítsuk. Ezért a matematikai leírók közti különbségeket a tanítóhalmaz alapján megtervezett küszöbszinthez hasonlítva értékelem ki, utasítom el, illetve folytatom a következő lépcsővel.

2.1 Altézis

Megmutattam, hogy többlépcsős, összehasonlításon alapuló osztályozás esetén, ha a mintaösszehasonlítás második fázisának célja a pontosság maximalizálása, az első lépésben történő, hozzávetőleges tulajdonságok alapján megvalósuló szűrés gyorsítja az osztályozást, illetve jelentősen növeli az osztályozás fedését.

A közvetlenül a minták reprezentációjához való összehasonlításon alapuló döntések esetében a minták számával egyenesen arányosan növekszik a kiértékelés komplexitása. Ezért kézenfekvő, hogy amennyiben a különbség meghalad egy szintet, az összehasonlítás folyamata le legyen állítva. A módszer annál hatékonyabb, minél kevesebb összehasonlítás után minél több egyértelmű elutasítás hajtható végre. A GSPPED leíró első néhány tulajdonsága tömören, egy-egy skalárral fejezi ki az alak excentricitását, illetve terület-arányát, amely megfelelő szűrő-paraméternek bizonyult.



Ábra 3. A fedés a döntés szűrés fázisa használatának függvényében, három különböző teszhalmazon mérve.

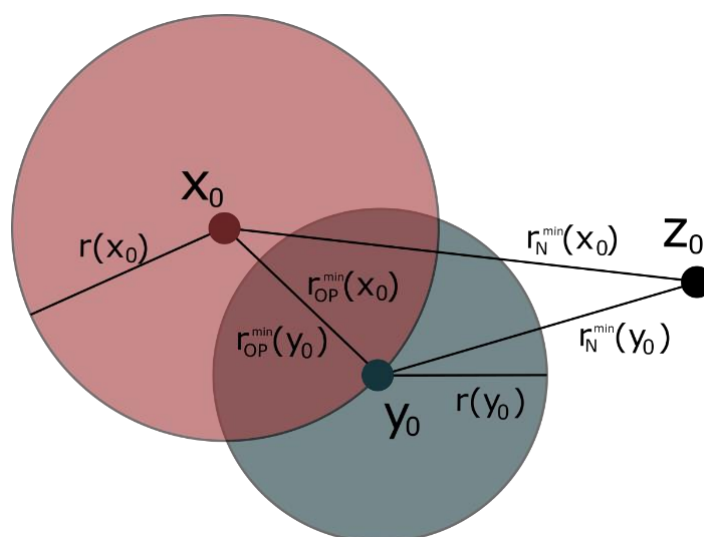
A szűrés alapján egy döntés átlagos ideje a szűrés nélkül szükséges idő 12%-ra esik vissza. A gyorsításon kívül azonban a szűrésnek van egy másodlagos hatása is, a felismerés fedése átlagosan 17%-kal nagyobb, ha szűrést alkalmazunk (3. ábra). A látszólag paradox jelenség az alábbiakkal magyarázható: A szűrésre használt tulajdonságok információelméleti értelemben merőlegeselek a második fázisban, a részletes összehasonlításra szolgáló tulajdonságoktól. A második fázisban a pontosság maximalizálásának érdekében az elfogadási sugár mértéke a legközelebbi, különböző osztályú egyed távolságától függ. A szűrés során tehát olyan elemek is kiszűrésre kerülnek, amelyek a második fázisban közel kerülnének a reprezentáns elemhez, így a szűrés növeli az elfogadási régiók nagyságát. A magyarázatot méréssel is igazoltam, amelyből kitűnik, hogy az előszűrés hatására az elfogadási régiók mintegy 25%-kal nagyobbak, mint szűrés nélkül. Mivel a második fázisban pontosságra optimalizált tanulás történik, a pontosság a szűréstől függetlenül változatlan, szinte 100%.

2.2 Altézis

Megterveztem és implementáltam a legközelebbi szomszéd osztályozó családjába tartozó Adaptív Korlátos Legközelebbi-Szomszéd (Adaptive Limited Nearest Neighborhood) osztályozót. A döntési módszer alkalmas arra, hogy a nem releváns (nulla osztályú) elemeket elutasítsa, amely a standard módszer egyik legfontosabb hátránya.

A legközelebbi szomszéd osztályozás az osztályozandó bemenetet a reprezentatív mintahalmaz előre megcímkézett elemeihez hasonlítja. A módszer hátrányát, a nem releváns elemek elutasításának hiányát kiküszöbölendő az algoritmust olyan módon bővítettem ki, hogy minden tárolt elemhez egy korlátot állítok be.

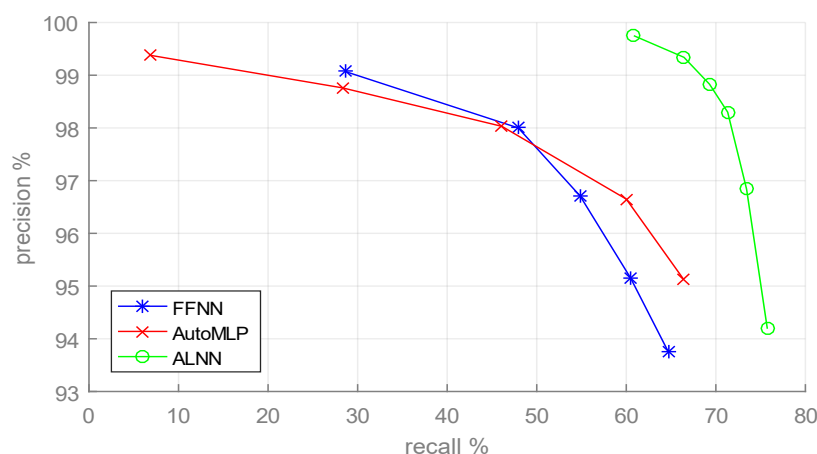
A korlátok beállítása minden reprezentatív halmazban lévő egyedre egyedien, a tanítóhalmazban lévő nem releváns és osztályhoz tartozó egyedek alapján történik. Ehhez olyan tanítóhalmazt alkalmazását vezettem be, amelyben nem csak a felismerendő osztályokhoz sorolható egyedek találhatóak, de ettől eltérő, semelyik osztályba sem tartozó, úgynevezett nem releváns minták is. Ezáltal lehetőség nyílik tartalmazási döntésre (ahol csak egy osztály van és a rajta kívül álló egyedek) és több osztály közötti döntésre is. A jelenlegi munka részét ugyan nem képezi, de a modell lehetőséget biztosít egyedek szintjén átfedő osztályok közötti döntésre is.



Ábra 4. Az Adaptív Korlátos Legközelebbi Szomszéd modell építése során elfogadási régiókat jelölünk ki, amelyek a tanítóhalmaz legközelebbi elemei alapján kerülnek meghatározásra (a piros illetve kék körök az x_0 és az y_0 elemek elfogadási régiói, z_0 pedig egy nem releváns elem – így nem rendelkezik elfogadási régióval.) y_0 elfogadási régiója ($r(y_0)$) a legközelebbi nem releváns elem távolsága alapján van meghatározva ($r_N^{\min}(y_0)$). x_0 elfogadási régióját ($r(x_0)$) a legközelebbi más osztályú, releváns elem határozza meg ($r_{OP}^{\min}(x_0)$)

A modell építése a következő folyamat szerint történik. A tanítóhalmaz címkével (ismert kimeneti értékkel) ellátott egyedeket tartalmaz, köztük nem

releváns, vagy „nulla osztályú” egyedeket is. A tanítóhalmaz releváns osztályú egyedeihez a környezetében elhelyezkedő más osztályú egyedek alapján korlátot rendelünk. Ha a szűrőkörnyezeten belül csak egy nem releváns egyed tartózkodik, akkor a korlát a legközelebbi nulla osztályú egyed és a tanítandó egyed közti távolság felével lesz egyenlő. Ha a szűrőkörnyezetben csak más valódi osztályú egyed található, akkor a korlát a legközelebbi ilyen egyed és a tanítandó egyed közti távolsággal lesz egyenlő. Ha csupán a tanítandó egyeddel megegyező osztályú egyed van a szűrőkörnyezetben, akkor a korlát a legmesszebb lévő ilyen egyed távolsága lesz. A legvalószínűbb ugyanakkor, hogy a tanítandó egyed közelében több osztály is megtalálható, ilyenkor a három módszerrel megállapítható korlát minimumát vesszük (lásd 4. ábra). Ugyanakkor, ha nincs más egyed a szűrőkörnyezetben, akkor ezt az egyedet, mivel semmi nem tudható a környezetéről, elhagyjuk. A valódi osztályú, egyedien beállított korláttal rendelkező egyedek összessége a reprezentatív halmaz. A korlát értékét, és így a fedést és a pontosságot egy vigilancia-paraméterrel tudjuk hangolni.



Ábra 5. Az ALNN algoritmus hatékonysága összevetve a ma leginkább használatos és eredményes osztályozókkal. Az ábrán nem szerepel az SVM és a k-NN osztályozó, mivel a többinél nagyságrenddel rosszabb eredményt tudtam elérni a használatukkal.

A módszert összevettem több, általánosan használt, eredményesen alkalmazott döntési algoritmussal. A GSPPED által előállított alakleíró osztályoztam k-NN, előrecsatolt neurális hálózatok és SVM algoritmusokkal, illetve az ALNN használatával. A kísérletekből egyértelműen kitűnik az ALNN hatékonysága. Az elért eredményeket a 5. ábrán szemléltetem.

2.3 Altézis

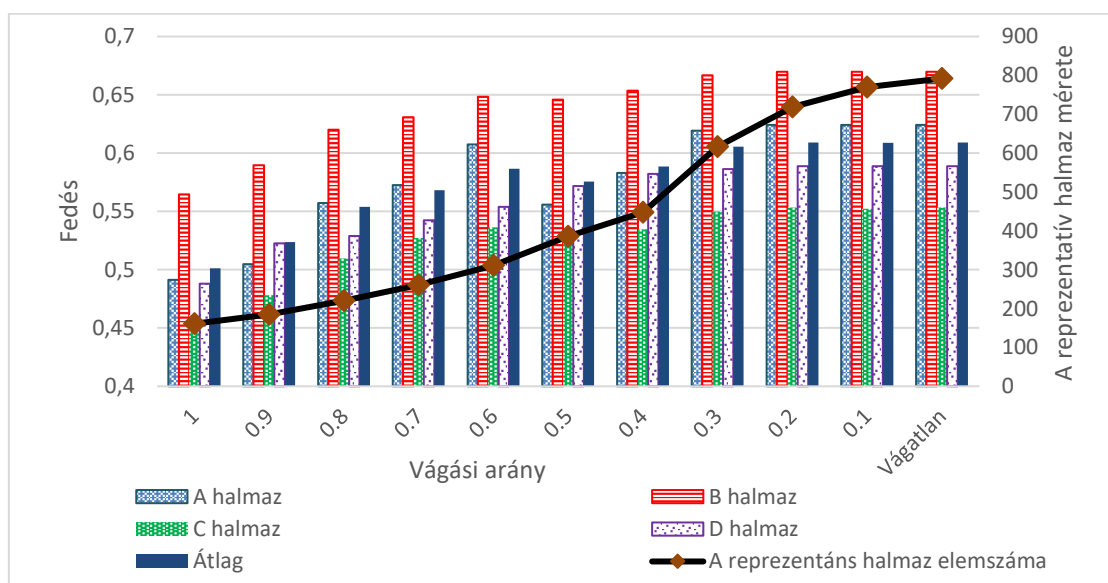
A reprezentációs halmaz adaptív tanítására és optimalizációjára bővítő és szűkítő módszert dolgoztam ki, amelyekkel elérhető, hogy a kiértékelés során új egyedek a tanítóhalmazba való beemelésével javítható legyen a modell, illetve, hogy a reprezentáns halmaz mérete optimális legyen a felesleges túlreprezentáltság elkerülésével.

A döntési modellek felhasználhatósága jelentősen növekszik azáltal, ha a mérések, tesztek során az osztályozási kimenet alapján félig vagy teljesen automatikusan javítható a tanító vagy reprezentáns halmaz. Ennek speciális esete az a felhasználási eset, amikor a kezdeti tanítóhalmaz nagyon kevés számú, kézzel összegyűjtött mintából áll, a teljes tanítóhalmaz pedig néhány tesztmérés során alakul ki.

Az adaptív korlátos legközelebbi szomszéd (ALNN) döntés során nem csak a bemenet osztályát lehet megállapítani, de a legközelebbi elemtől mért távolság és az elfogadási régió alapján a döntés bizonyosságát, illetve folytonos tulajdonságtér esetén azt is, mennyire helyezkedik el a bemenet az adott osztály határvonalán. Ha az osztályozott bemenet a meglévő osztályrégió (vagy egy belső lefedetlen terület) szélén található, azt is jelenti, hogy a tanítóhalmazhoz hozzáadva növelhető annak fedése, illetve a régióhatárvonal pontossága. A két feltétel egymással ellentétes, ezért a bemenetek időbeli viszonyát, egymásutánosságát használtam fel a bizonyosság igazolására az osztályozási perem közeli bemenetek esetében. A reprezentáns halmaz így biztosan olyan

egyedekkel bővül, amelyek fedésnövekedést biztosítanak, és – a teszteredmények fényében – biztosan az adott osztályhoz tartoznak.

Az elfogadási régiók metszetét becsülve a túlzottan sűrű fedésű reprezentatív halmaz redundáns elemei is kiszűrhetőek. Az optimalizálási algoritmus tartalmazási frekvencia alapján rendezi a reprezentáns egyedeket. Ezek után egy iteratív algoritmussal szűrjük a tanítóhalmaz elemeit: A sor elején lévő egyedeket eltávolítjuk a sorból, és az általa lefedett egyedeket kivesszük a reprezentáns halmazból. A lépést addig folytatjuk, amíg a sor végére nem érünk. Az optimalizálás során a reprezentáns halmaz méretét 30-50%-kal csökkentettük, miközben a fedés csupán 5-15%-ot csökkent (a pontos értékeket a 6. ábra mutatja.)



Ábra 6. A reprezentációs halmaz optimalizáció hatása a fedésre.

3. Tézis

A síkon elhelyezett, gyengén csatolt oszcillátorok által alkotott rendszert számítási egységként formalizáltam, definiálva a program, az adat, a bemenet és a kimenet fogalmát a hálózaton. Kísérletileg igazoltam, hogy az oszcillátorhálózat alkalmas arra, hogy a gerjesztés formájában kapott – statikus és dinamikus – bemenet alapján a feladatnak megfelelő kimenetet kapjunk, tehát az eszköz használható, mint programokat végrehajtó számítási egység.

Kapcsolódó publikációk: [A2][A3][A6]

Az oszcillátorok gerjesztés hatására a dinamikájuktól és a csillapítás mértékétől függő rezgésbe jönnek. Amennyiben több oszcillátor is jelen van, amelyek minimális, de mégis érzékelhető mértékben kölcsönhatnak egymással, gyengén csatolt oszcillátorhálózatról beszélünk. A nem-booli számítási egységek terén végzett kutatások egyik fókuszpontjába a Spin Torque Oszcillátorok (STO) és az általuk alkotott hálózat viselkedése és alkalmazása került. A fizikai megvalósítás előtt a legnagyobb hangsúly egyelőre a lehetséges alkalmazási módokra, illetve azok szimulációjára helyeződik. Az általam használt modellben az oszcillátorok egy miliszekundum nagyságrendű idő elteltével szinkronizálódnak, azonban nem azonos fázisban. A számítási egységet úgy formalizáltam, hogy az állapot, tehát az adat az aktuális fázisok különbségeinek vektorából áll össze. A program a topológia és az oszcillátorok dinamikájának összessége. A bemenet a bemenetként címkézett oszcillátorokra gyakorolt gerjesztés, a kimenet pedig a kimeneti oszcillátorok közti fáziskülönbség.

Tértranszformáció STO hálózat segítségével

A klasszikus kétlépcsős objektumfelismerési modellben a sokdimenziós bemenetből előbb egy vagy több leírás készül, amelyek dimenziószáma

lényegesen kisebb a bemenet dimenziószámánál, a végső döntés a leírók alapján történik. A taníthatóság, így a döntés helyessége nagymértékben függ attól, hogy a leíró mennyire diszkriminatív a kimeneti osztályok vonatkozásában, tehát hogy a tulajdonságtérbe mennyire kompakt módon képezi le az egy osztályba tartozó bemeneteket, és mennyire különíti el a különböző bemeneteket.

Kísérletemben a tértranszformáció demonstrálására gépi tanulási eredmények javítását tűztem ki célul oly módon, hogy a leíró-generálás és az osztályozás közé egy új elemet, egy OCNN tömböt helyeztem, amelynek a bemenete a reprezentáns vektor, kimenete pedig az úgynevezett szignatúra, ami alapján történik az osztályozás. Az OCNN topológia tervezésére, ez esetben a láncstruktúrában lévő oszcillátorok távolságának beállítására genetikus algoritmust terveztem, amely célfüggvénye a teljes tanítás helyessége volt, amely szoros összefüggésben áll a már említett kompaktsággal, tehát a minták alapján meghatározott tértranszformáció működésével. Tanítási mintaként szándékosan gyengített EPPSED alakleírót és H-MAX képleírót használtunk.

A kísérlet során az összes tanítóhalmaz esetén sikerült javulást elérni. Az EPPSED leírón alkalmazott tanulás eredményében bekövetkezett javulást az X. ábra mutatja. A H-MAX leíró vektorain a tulajdonság- illetve szignatúratér kompaktságát is kimértük az osztályokon belüli átlagos távolság, illetve az osztályok közötti átlagos távolság által. A mérések alapján a kompaktság terén akár hússzoros javulást is el tudtunk érni.

STO-hálózaton alapuló osztályozó implementálása statikus és dinamikus bemenetre

Az STO hálózatok számítási egységként való használatának létjogosultságát másodszorra osztályozóként való alkalmazásban teszteltem. Az osztályozás feladatát úgy is fel lehet fogni, mint egy magas dimenziós bemenet leképezését egy nagyon alacsony, jellemzően egy dimenziós kimenetre. Az

osztályozás – a legegyszerűbb, bináris osztályozásnál maradván – a tértranszformáción túl egy szeparációt is magában foglal, ezért komplex feladatnak tekinthető. Kísérleteimben ugyanakkor nem csak statikus, hanem dinamikus, időben változó bemenetet is használtam.

A bináris osztályozó kimenete egy adat, ami oszcillátorhálózat esetén két szomszédos szélső oszcillátor fáziskülönbsége. A bemeneti oszcillátorok száma a bemenet dimenziószámától függ. Így egy olyan topológia megépítését tűztem ki célul, amely többrétegű: Az alsó rétegben a bemenet dimenziószámával megegyező oszcillátor található. A következő rétegekben egy-egy oszcillátorral kevesebb, egészen a legfelső rétegig, ahol két oszcillátor zárja a topológiát.

A topológiát többféle térbeli és téridőbeli bemenettel teszteltem. Statikus mintaként előbb összegzési képességet vizsgáltam, majd XOR-típusú logikai képességet. Dinamikus bemenetnek pedig mozgó lépcsőfüggvényt használtam, változó irányokban és zajjal. A kísérletek mindegyikében sikerült igazolnom a topológia taníthatóságát.

4. Az eredmények alkalmazási területei

Az első és a második tézis a Bionikus Szemüveg projekttel együtt fejlődött, aktívan hasznosítva az élő tesztek során gyűjtött tapasztalatokat. Az alakleíró és osztályozó alkalmazásával egy összetett osztályozó épült, amely alkalmas Forint bankjegyek felismerésére. Az alkalmazás több jellemző egyidejű elemzésével állapítja meg a bankjegy címletét, és az osztályozott objektumok kölcsönös viszonyát is ellenőrizve magasan megbízható eredményt szolgáltat. Mivel téves végső döntés csak úgy születhet, ha több osztályozó egyidejűleg több képkockára is téves választ ad, a végső döntés szintű hiba mértéke elhanyagolható.

A rendszer a bankjegyeken kívül természetesen más vizuális feladatokhoz is használható, amelyek merev, lapos alakzatok osztályozását igénylik. A Bionikus Szemüveg projekt keretében tovább lehet hasznosítani az algoritmusokat egyrészt más bankjegyek felismerésére, illetve piktogramok, LED kijelzők leolvasására, információs táblák azonosítására.

A kifejlesztett leíró és osztályozó ugyanakkor más területen is alkalmazható. A módszerek implementálásra kerültek FPGA-n is, lehetővé téve a gyorsabb feldolgozást. (Az FPGA-n történő implementáció nem volt része a kutatásomnak.)

Az oszcillátorhálózatok a jövőben a mai klasszikus számítástechnikai alkotóelemek egy részét válthatják fel. Alkalmazási területeiket számos kutatóintézetben vizsgálják, jelen dolgozat is egy ilyen folyamat része.

5. Köszönetnyilvánítás

Elsősorban köszönöm témavezetőmnek, Roska Tamásnak, hogy befogadott a tudományos élet családjába, és hogy tudásával, alázatával és lelkesedésével magával ragadott a tudományos pályán. Köszönöm, hogy lehetővé tette, hogy Tokyoba, illetve Portlandba utazhassak.

Köszönet illeti Karacs Kristófot, hogy türelmes mentorom volt az évek során. Hálás vagyok neki, hogy ötleteivel és rendkívül aprólékos, ugyanakkor építő kritikáival segítette munkámat.

Hálás vagyok Pázmány Péter Katolikus Egyetem Információs Technológiai és Bionikai Kara, illetve a doktori iskola volt és jelenlegi vezetőinek, Szolgay Péternek és Nyékiné Dr. Gaizler Juditnak, hogy egy magas színvonalú és értékközpontú kutatóközpontot teremtettek meg.

Köszönet illeti tokyoi témavezetőmet, Tadashi Shibát, hogy előremozdította kutatásomat, és hozzájárult látóköröm szélesítéséhez.

Köszönettel tartozom családomnak, hogy gyerekkorom óta az értékek tiszteletére neveltek, és a tudás folyamatos bővítésére ösztönöztek.

Hálával tartozom menyasszonyomnak, Timinek támogatásáért, türelméért és szeretetéért.

Köszönöm Bérczesné Novák Ágnesnek, hogy egyetemi tanulmányaim kezdetén segítette a hallgatók közösségének kialakulását.

Köszönöm közvetlen kollégáimnak, barátaimnak, Horváth Andrásnak, Radványi Mihálynak, Koller Miklósnak, Zsedrovits Tamásnak, hogy mind szakmai, mind baráti társak voltak doktoranduszi tanulmányaim során.

6. Publikációk

- [A1] A. Stubendek, K. Karacs, “Shape Recognition Based on Projected Edges and Global Statistical Features”, *Mathematical Problems in Engineering*, vol. 2018, Article ID 4763050, 18 p, 2018.
- [A2] A. Horváth, M. Koller, A. Stubendek, T. Roska, “Spatial-temporal Event Detection via Frameless Cellular Wave Computing - a Review”, *Nonlinear Theory and Its Applications*, IEICE, 2014, vol 5 (3), pp. 391-408
- [A3] A. Stubendek, K. Karacs, T. Roska, „Shape Description Based on Projected Edges and Global Statistical Features”, *International Symposium on Nonlinear Theory and its Applications (NOLTA 2014)*, Luzern, Switzerland, 2014
- [A4] K. Karacs, M. Radványi, A. Stubendek and B. Bezányi, “Learning Hierarchical Spatial Semantics for Visual Orientation Devices,” in *Proc. of the IEEE/CAS-EMB 2014 Biomedical Circuits and Systems Conference (BIOCAS 2014)*, Lausanne, Switzerland, pp. 141–144, Oct. 2014
- [A5] A. Horváth, A. Stubendek, T. Roska, “One dimensional cellular array of in-plane spin torque oscillators,” *13th IEEE International Conference on Nanotechnology*, Beijing, China, 2013
- [A6] Z. Solymar, A. Stubendek, M. Radvanyi, K. Karacs, "Banknote Recognition for Visually Impaired" in *Proc. of the European Conference on Circuit Theory and Design (ECCTD'11)*, Linköping, Sweden, Aug 2011.
- [A7] M. Radvanyi, Z. Solymar, A. Stubendek, K. Karacs, „Mobile Banknote Recognition – Topological Models in Scene Understanding”, in *Proc. of the 4th International Symposium on Applied Sciences and Communication Technologies (ISABEL 2011)*, Barcelona, Spain, 2011.
- [A8] T. Roska, A. Horvath, A. Stubendek, F. Corinto, Gy. Csaba, W. Porod, T. Shibata, G. Bourianoff, “An Associative Memory with Oscillatory CNN Arrays using Spin-Torque Oscillator Cells and Spin-Wave Interactions Architecture and End-to-end Simulator,” *13th IEEE international workshop on cellular nanoscale networks and applications*, Turin, Italy, 2012.

A témához kapcsolódó irodalom

- [1] D. Poole, A. Mackworth, “Artificial Intelligence: Foundations of Computational Agents”, Cambridge University Press, ISBN 978-0-521-51900-7, 2010.
- [2] K. P. Murphy, “Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning series)”, The MIT Press, ISBN-13: 978-0262018029, 2006.
- [3] T. Hastie, R. Tibshirani, J. Friedman, “The Elements of Statistical Learning – Data Mining, Inference, and Prediction”, Springer Series in Statistics, ISBN: 978-0-387-84857-0, 2011.
- [4] Ch. M. Bishop, “Pattern Recognition and Machine Learning”, Springer Science + Business Media, LLC, ISBN: 978-0387-31073-2, 2009.
- [5] S. Russel, P. Norwig, “Artificial Intelligence: A modern approach”, Pearson Education, ISBN 963545411-2, 2010.
- [6] W. Ertel, “Introduction to Artificial Intelligence”, Springer Science + Business Media, LLC, ISBN: 978-085729-298-8, 2011.
- [7] M. Kantrardzic, “Data Mining: Concepts, Models, Methods, and Algorithms”, Institute of Electrical and Electronic Engineers, 2nd ed., ISBN: 978-0-470-89045-5, 2011.
- [8] I. Witten, E. Frank, “Data Mining: Practical Machine Learning Tools and Techniques”, 2nd ed., Morgan Kaufmann series in data management systems, ISBN: 0-12-088407-0, 2005.
- [9] P. Flach, “Machine Learning: The Art and Science of Algorithms that Make Sense of Data”, Cambridge University Press, ISBN: 978-1-107-09639-4, 2012.
- [10] S. Haykin, “Neural Networks and learning machines”, 3rd edition, Pearson Education Inc, ISBN: 978-0-13-129376-2, 2009
- [11] M. L. Minsky, S. A. Papert, “Perceptrons: An Introduction to Computational Geometry”, MIT Press, Cambridge, ISBN: 0-262-63022-2, 1969.
- [12] S. Marshland, “Machine Learning: An Algorithmic Perspective”, Chapman & Hall / CRC machine learning & pattern recognition series, ISBN: 978-1-4200-6718-7, 2009.

-
- [13] J. W. Rudy, "The neurobiology of Learning and Memory", Sinauer Associates, ISBN: 978-0-87893-669-4, 2008
- [14] R. O. Duda, P. E. Hart, D. G. Stork, "Pattern Classification", A Wiley-Interscience Publication, John Wiley & Sons, 2nd edition, ISBN: 0-471-05669-3, 2001.
- [15] E. Alpaydin, "Introduction to Machine Learning", MIT Press, Cambridge, ISBN: 9780262012119, 2004.
- [16] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading, ISBN: 0-201-15767-5, 1989.
- [17] M. Mitchell, "An Introduction to Genetic Algorithms" Cambridge, MA: MIT Press, ISBN 9780585030944., 1996.
- [18] A. Andreopoulos, John K. Tsotsos, "50 Years of object recognition: Directions forward", Computer Vision and Image Understanding, vol. 117 (8), pp. 827-891, 2013.
- [19] W. K. Pratt, "Digital Image Processing", John Wiley & Sons, ISBN: 0-471-37407-5, 2001.
- [20] I. Biederman, G. Ju, "Surface versus edge-based determinants of visual recognition", Cognitive Psychology, vol. 20, pp. 38-64, 1988.
- [21] L. K. Samuelson, L. B. Smith, "They call it like they see it: Spontaneous naming and attention to shape", Developmental Science, vol 8 (2), pp. 182-198, 2005.
- [22] D. K. Prasad, "Survey of the problem of object detection in real images." International Journal of Image Processing (IJIP), vol. 6.6, pp. 441. 2012.
- [23] S. L. Tanimoto, "An Interdisciplinary Introduction to Image Processing", Massachusetts Institute of Technology, ISBN: 978-0-262-01716-9, 2012.
- [24] Jähne, "Digital Image Processing: concepts, algorithms, and scientific applications", Springer-Verlag Berlin Heidelberg, ISBN: 3-540-56941-3, 1991.
- [25] J. W. H. Tangelder, R. C. Velkamp, "A survey of content based 3D shape retrieval methods," Proceedings Shape Modeling Applications, 2004., Genova, Italy, 2004, pp. 145-156., 2004.
- [26] S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques.", Optics and lasers in engineering, vol 48.2, pp.149-158, 2010.
- [27] J. C. Russ, "The Image Processing Handbook", CRC Press, 1995, ISBN: 0-8493-2516.
-

-
- [28] S. A. Dudani, K. J. Breeding, R. B. McGhee, "Aircraft identification by moment invariants," *IEEE Transactions on Computers*, vol. 26 (1), pp. 39–46, 1977.
- [29] L. Gupta and M. D. Srinath, "Contour sequence moments for the classification of closed planar shapes," *Pattern Recognition*, vol. 20, no. 3, pp. 267–272, 1987.
- [30] C. Chang; D. J. Buehrer; S. M. Hwang, "A shape recognition scheme based on relative distances of feature points from the centroid", *Pattern Recognition*, Vol: 24, Issue: 11, pp 1053-1063, 1991.
- [31] R. Davies, "Machine Vision: Theory, Algorithms, Practicalities", vol. 54, Academic Press, New York, NY, USA, 1991.
- [32] P. J. van Otterloo, "A Contour-Oriented Approach to Shape Analysis", Prentice-Hall International (UK) Ltd, New Jersey, NJ, USA, 1991.
- [33] M. Sonka, V. Hlavac, and R. Boyle, "Image Processing, Analysis and Machine Vision", Chapman and Hall Computing, Boca Raton, Fla, USA, ISBN: 978-0-412-45570-4, 1993.
- [34] Q. M. Tieng and W. W. Boles, "Recognition of 2D object contours using the wavelet transform zero-crossing representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 8, pp. 910–916, 1997.
- [35] Freeman, "On the encoding of arbitrary geometric configurations," *IRE Transactions on Electronic Computers*, vol. EC-10, no. 2, pp. 260–268, 1961
- [36] D. Zhang, G. Lu, "Review of shape representation and description techniques", *Pattern Recognition* 37, pp.1 – 19, 2004
- [37] M. K. Hu, "Visual pattern recognition by moment invariant," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.
- [38] A. Khotanzad and Y.H. Hong, "Invariant image recognition by zernike moments", *IEEE Trans. on Pattern Anal. and Machine Intell.*, 1990.
- [39] R. B. Yadava, N. K. Nishchalb, A. K. Guptaa, V. K. Rastogi, "Retrieval and classification of objects using generic Fourier, Legendre moment, and wavelet Zernike moment descriptors and recognition using joint transform correlator", *Optics & Laser Technology*, vol. 40 (3), pp. 517-527, 2008
- [40] C.-H. Teh and R. T. Chin, "On Image Analysis by the Methods of Moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 496-513, 1988.
-

-
- [41] P. Rosin, J. Zunic, "Handbook of Applied Algorithms: Solving Scientific, Engineering, and Practical Problems, 2D Shape Measures for Computer Vision", Wiley-IEEE Press, ISBN: 978-0-470-04492-6, 2008.
- [42] Flusser, T. Suk, B. Zitova, "Moments and Moment Invariants in Pattern Recognition", John Wiley & Sons, ISBN: 978-0-470-69987-4, 2009.
- [43] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing: Image Communication*, vol. 17, no. 10, pp. 825–848, 2002.
- [44] D. Zhang, G. Lu. "Generic Fourier descriptor for shape-based image retrieval." *ICME* (1). 2002.
- [45] S. Li, M.-C. Lee, and C.-M. Pun. "Complex Zernike moments features for shape-based image retrieval", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol 39. (1) , pp. 227-237., 2009.
- [46] M. Vorobyov, "Shape classification using zernike moments", Technical Report, iCamp-University of California Irvine, 2011.
- [47] E. R. Kandel, J. H. Schwartz, T. M. Jessel, S. A. Siegelbaum, A. J. Hudspeth, (Ed.), "Principles of Neural Science", E. R. Kandel (Ed), McGraw-Hill Companies, ISBN: 987-0-07-139011-8, 2013.
- [48] M. Yagi, T. Shibata, "An Image Representation Algorithm Compatible with Neural-Associative-Processor-Based Hardware Recognition Systems," *IEEE Trans. Neural Networks*, Vol. 14, No. 5, pp. 1144-1161, 2003.
- [49] D. Geer, "Chip makers turn to multicore processors", *Computer, IEEE*, vol 38. (5), pp 11-13. 2005.
- [50] T. Roska, P. Giovanni, W. W. Chai. "Cellular wave computing via nanoscale chip architectures." *International Journal of Circuit Theory and Applications* 40.12, pp. 1187-1189. 2012.
- [51] Y. Zhang, W. Zhao, J. O. Klein, W. Kang, D. Querlioz, Y. Zhang, "Spintronics for low-power computing." *Proceedings of the conference on Design, Automation & Test in Europe. European Design and Automation Association*, pp. 303, 2014.
- [52] U. Ingard, "Fundamentals of Waves & Oscillations", Cambridge University Press, ISBN: 0-521-32734-2, 1988.
- [53] D. Landau, E. M. Lifshitz, "Theory of the dispersion of magnetic permeability in ferromagnetic bodies", *Phys. Z. Sowietunion*, 8, 153, 1935.
-

-
- [54] A. Slavin: “Microwave sources: Spin-torque oscillators get in phase”, *Nature Nanotechnology* 4, pp. 479-480. 2009.
- [55] S. E. Russek, W. H. Rippard, T. Cecil, R. Heindl, “Spin-transfer nano-oscillators“, *Handbook of Nanophysics: Functional Nanomaterials*, 2010.
- [56] G. Bourianoff, “Towards a Bayesian processor implemented with oscillatory nanoelectronic arrays“, Keynote lecture at IEEE International Workshop on Cellular Nanoscale Networks and Applications, Turin, Italy, 2012.
- [57] Gy. Csaba, Á. Papp, W. Porod, “Perspectives of using spin waves for computing and signal processing“, *Physics Letters A*, vol. 381 (17), 2017.
- [58] A. Horváth, F. Corinto, Gy. Csaba, W. Porod, T. Roska, “Synchronization in Cellular Spin Torque Oscillator Arrays,” 13th IEEE International Workshop on Cellular Nanoscale Networks and Applications, Turin, Italy, 2012.
- [59] O. Chua and T. Roska, *Cellular Neural Networks and visual computing*. Cambridge, UK: Cambridge University Press, 2002.
- [60] T. Roska and L. O. Chua, “The CNN universal machine: an analogic array computer,” *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 163–173, Mar. 1993.
- [61] Kék, K. Karacs, and T. Roska. (2007) Cellular wave computing library, version 2.1 (templates, algorithms and programs). [Online]. Available: <http://cnn-technology.itk.ppke.hu/Library/v2.1b.pdf> visited on 02-12-2009.
- [62] K. Karacs, A. Lázár, R. Wagner, D. Bálya, T. Roska, and M. Szuhaj, “Bionic Eyeglass: an Audio Guide for Visually Impaired,” in *Proc. of the First IEEE Biomedical Circuits and Systems Conference (BIOCAS 2006)*, London, UK, Dec. 2006, pp. 190 – 193.
- [63] K. Karacs, A. Lázár, R. Wagner, B. Bálint, T. Roska, and M. Szuhaj, “Bionic Eyeglass: The First Prototype, A Personal Navigation Device for Visually Impaired,” in *Proc. of First Int’l Symp. on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2008)*, Aalborg, Denmark, 2008
- [64] K. Karacs, Á. Kusnyerik, M. Szuhaj, T. Roska, “Situation-specific Scene Interpretation in a Bionic Navigation Device for Visually Impaired, *International Conference on Vision in 3D Environments*, Toronto, Canada, Jul, 2009.
- [65] K. Karacs, M. Radványi, Á. Kusnyerik, T. Roska, “Basic Scene Understanding and Navigation with a Bionic Camera,” *International Conference on The Eye and The Auto*, Detroit, MI, Sept, 2009
-

-
- [66] K. Karacs, Á. Kusnyerik, M. Radványi, T. Roska, M. Szuhaj, "Towards a Mobile Visual Navigation Device", CNNA2010, 2010.
- [67] Radvanyi, G. E. Paziienza, and K. Karacs, "Crosswalk Recognition through CNNs for the Bionic Camera: Manual vs. Automatic Design," in Proc. of the European Conference on Circuit Theory and Design (ECCTD'09), Antalya, Turkey, Aug, 2009.
- [68] K. Karacs, M. Radványi, M. Görög, T. Roska, "A Mobile Visual Navigation Device: New algorithms for crosswalk and pictogram recognition," in Proc. of 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2009), Bratislava, Slovakia, Nov, 2009.
- [69] M. Radvanyi, K.Karacs, "Navigation through Crosswalks with the Bionic Eyeglass" in Proc. of 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010), Rome, Italy, 2010.
- [70] M. Radványi, K. Karacs, "Autonomous Detection of Information Patterns Through Hierarchical Peeling," in Proc. of the 21st European Conference on Circuit Theory and Design (ECCTD 2013), Dresden, Germany, Sept. 2013.
- [71] M. Fritz, B. Schiele, "Towards Integration of Different Paradigms in Modeling, Representation, and Learning of Visual Categories", Object Categorization: computer and human vision perspectives, edited by S. J. Dickinson, A. Leonardis, B. Schiele, M. J. Tarr, Cambridge University Press, ISBN: 978-0-521-88738-0, 2009.
- [72] E. Pekalska, R. P. W. Duin, "The Dissimilarity Representation for Pattern Recognition: Foundations and Applications", World Scientific Publishing Co. Pte., ISBN: 981-256-530-2, 2005
- [73] S. Khanam, S. Jang, W. Paik, "Shape Retrieval Combining Interior and Contour Descriptors", International Conference FGCV, Jeju Island, pp.120-128, 2011.
- [74] T. G. Dietterich, "An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization", Article Machine Learning, vol. 40, (2), pp 139-157, 2000.
- [75] T. G. Dietterich, "Multiple Classifier Systems", Chapter Ensemble Methods in Machine Learning, Volume 1857 of the series Lecture Notes in Computer Science pp 1-15, 2000.
-

-
- [76] D. Opitz, R. Maclin, "Popular Ensemble Methods: An Empirical Study", vol. 11, pp. 169-198, 1999.
- [77] M. Yang, Kidiyo Kpalma, Joseph Ronsin. "A Survey of Shape Feature Extraction Techniques", Peng-Yeng Yin. Pattern Recognition, IN-TECH, pp.43-90, 2008.
- [78] G. Carneiro, A. B. Chan, P. J. Moreno, N. Vasconcelos, "Supervised Learning of Semantic Classes for Image Annotation and Retrieval", IEEE T. on Pattern Analysis and Machine Intelligence, 2007.
- [79] B. Kotsiantis, Sotiris, I. D. Zaharakis, P. E. Pintelas. "Supervised machine learning: A review of classification techniques." vol. 3-24, 2007.
- [80] R. C. Gonzalez, R. E. Woods. "Digital Image Processing". Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, ISBN: 978-0131687288, 2001.
- [81] M. Riesenhuber, T. Poggio, "Hierarchical models of object recognition in cortex," Nature Neuroscience, vol. 2, no. 11, pp. 1019-1025, 1999.