

**Kép- és beszédfeldolgozás** c. tantárgy (GEVAU122M)  
előadásának és gyakorlatának ütemterve

<i>Tárgynév:</i>	<b>Képfeldolgozás</b>		
<i>Rövid név:</i>	Kepfel.	<i>Kód</i>	<b>GEVAU122M</b>
<i>Angol név:</i>	Image processing		
<i>Intézet:</i>	Automatizálási és Infokommunikációs Intézet		
<i>Tárgyfelelős:</i>	Dr. Czap László egy. docens (e-mail: czap@uni-miskolc.hu)		
<i>Előtanulmányok:</i>	nincs		
<i>Kredit:</i>	5	<i>Követelmény:</i>	kollokvium
<i>Heti óraszámok</i>	<i>Előadás: 2</i>	<i>Gyakorlat: 2</i>	
<i>Oktatási cél:</i>	Digitális képfeldolgozással kapcsolatos ismeretek elsajátítása		
<i>Tárgy tartalom:</i>	A számítógépes képfeldolgozás eszközei. Emberi látás, színlátás, műveletek a képtartományban. Színelmélet, színrendszerek. A gépi látás alapfogalmai, sztereo- és 3D látás. Geometriai transzformációk. Hisztogram műveletek. Konvolúció, zajsűrés, élkiemelés. Laplace, Roberts, Prewitt, Sobel operátorok. Medián szűrés. A síkfrekvencia értelmezése, kétdimenziós Fourier transzformáció, képjavítás a síkfrekvencia tartományban. Egy- és kétdimenziós diszkrét koszinusz transzformáció. Veszteséges és veszteségmentes képtömörítés, JPEG. Képmorfológiai műveletek. Alakzat felismerés, optikai karakterfelismerés. A hangképzés mechanizmusa, beszédhangok osztályozása. Hangtömörítés. Beszédszintézis. Beszédfelismerés. Audiovizuális beszédfeldolgozás.		
<i>Ajánlott Irodalom</i>	<ol style="list-style-type: none"> <li>1. Dr. Czap László: Képfeldolgozás (online elérhető jegyzet), 2008. <a href="http://www.gepesz.uni-miskolc.hu/hefop/letolt.php?dwn=1kepfeldolgozas">http://www.gepesz.uni-miskolc.hu/hefop/letolt.php?dwn=1kepfeldolgozas</a></li> <li>2. Olaszky, Németh (Ed.) A magyar beszéd. Akadémiai kiadó 201</li> <li>3. Berke J., Hegedűs Gy. Cs., Kelemen D., Szabó J.: Digitális képfeldolgozás és alkalmazásai. V. E. Georgikon M. K., PICTRON, 2001.</li> <li>4. R. G. Gonzales, R. E. Woods: Digital Image Processing. Prentice Hall, 2002.</li> <li>5. I. Pitas: Digital Image Processing Algorithms and Applications. Wiley, 2000.</li> </ol>		
<i>Jellemző oktatási módok</i>			
<i>Oktatási nyelv:</i>	magyar		
<i>Előadás:</i>	tábla, számítógép, projektor		
<i>Gyakorlat:</i>	számítógép, projektor		
<i>Évközi feladatok, zárthelyik:</i>	3 db beadandó feladat		
<i>Lezárási feltételek:</i>	A Tanulmányi és Vizsgaszabályzat szerint. Az Előadások látogatása, a gyakorlatokon való aktív részvétel, a kiadott évközi gyakorlati feladatok elfogadható szinten való elkészítése. A lezáráshoz az aláírás megszerzését követően írásbeli / szóbeli vizsgát kell tenni. Az évközi teljesítmény		

beszámításra kerül (40 %-ban) a tárgyat lezáró érdemjegyre.	
<i>Előadás és gyakorlat ütemterve</i>	
1. alkalom	EA: A számítógépes képfeldolgozás eszközei Gyak: Bevezetés, képfeldolgozó eszközök és műveletek
2. alkalom	EA: Emberi látás, színlátás, műveletek a képtartományban Gyak: Optikai illúziók., felbontás, képfeldolgozó módszerek
3. alkalom	EA: Színelmélet, színrendszerek Gyak: Színrendszerek, grafikus fájlformátumok
4. alkalom	EA: A gépi látás alapfogalmai, sztereo- és 3D látás. Geometriai transzformációk. Gyak: Pixelgrafikus képszerkesztés, képkorrekciók, képtranszformációk
5. alkalom	EA.: Pont-pont, lokális és globális műveletek Gyak: Képfeldolgozási műveletek, egyéni feladatok kiadása
6. alkalom	EA: A síkfrekvencia értelmezése, kétdimenziós Fourier transzformáció, képjavítás a síkfrekvencia tartományban. Egy- és kétdimenziós diszkrét koszinusz transzformáció Gyak: Vektorgrafikus képszerkesztés, transzformációs eszközök, egyéni feladatok kiadása
7. alkalom	EA: Egy- és kétdimenziós diszkrét koszinusz transzformáció Gyak: Háromdimenziós képszerkesztés alapjai, animációk
8. alkalom	EA: Veszteséges és veszteségmentes képtömörítés, JPEG Gyak: Képtömörítő eljárások, egyéni feladatok bemutatása
9. alkalom	EA: Képmorfológia, alakzat felismerés, optikai karakterfelismerés Gyak: Képmorfológiai műveletek, videóformátumok, mozgóképszerkesztés, egyéni feladatok pótolási lehetősége
10. alkalom	EA: A hangképzés mechanizmusa, beszédhangok osztályozása Gyak: Lineáris predikció
11. alkalom	EA: Handtömörítés Gyak: Cepstrum transzformáció, MFCC
12. alkalom	EA: Beszédszintézis Gyak: Profovox
13. alkalom	EA: Beszédfelismerés Gyak: Rejtett Markov modell
14. alkalom	EA: Audiovizuális beszédfeldolgozás Gyak: A Beszédasszisztens rendszer

Miskolc, 2019. szeptember 09.

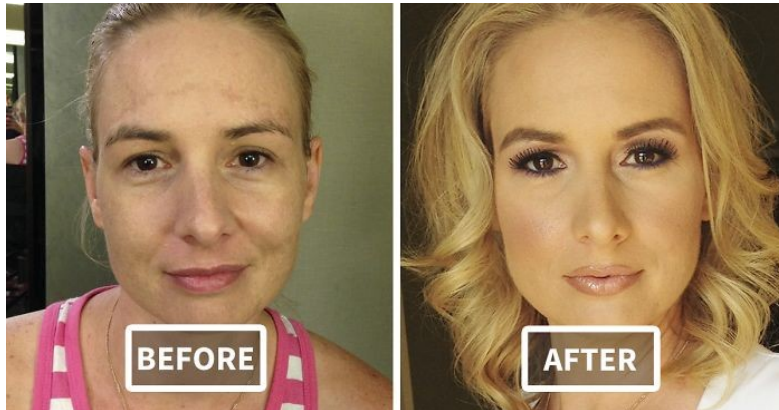
Dr. Trohák Attila  
intézetigazgató, egyetemi docens

Dr. Czap László  
egyetemi docens, tárgyjegyző

**MINTA FÉLÉVKÖZI FELADATOK**  
**Képfeldolgozás (GEVAU509B) c. tantárgyból**

1.) Mutassa be a gyakorlati órákon tanult pixelgrafikus képszerkesztő, képkorrekciós és képtranszformációs műveletek egy gyakorlati példán keresztül.

*Például:*



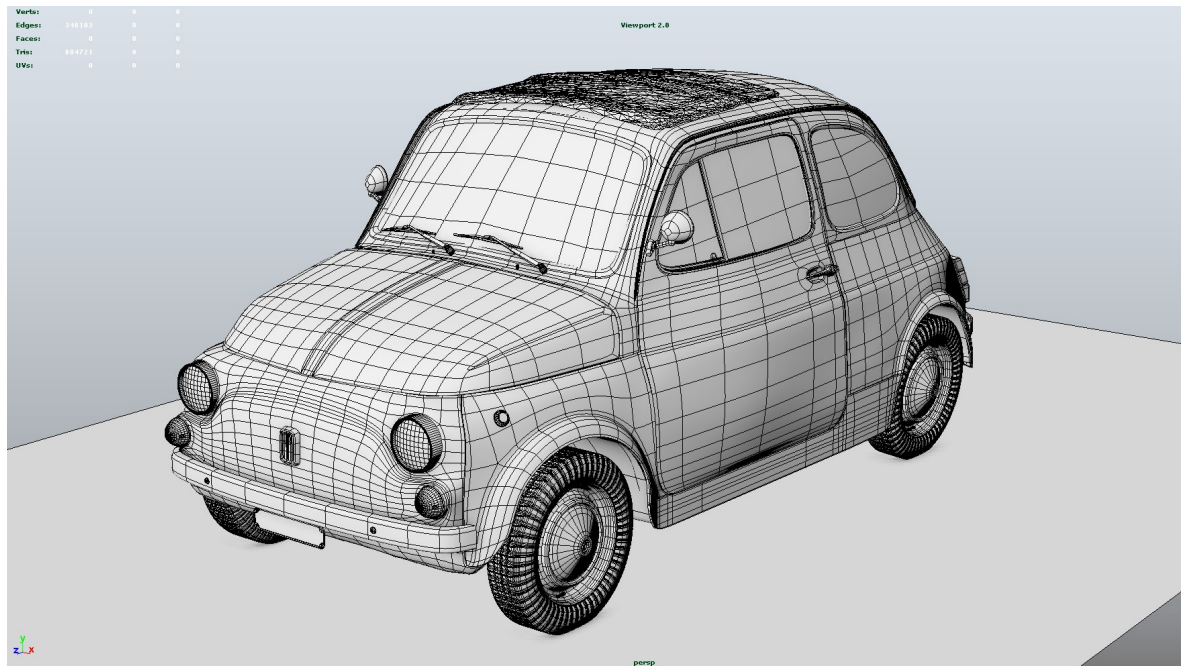
2.) Mutassa be a gyakorlati órákon tanult vektorgrafikus képszerkesztő, képkorrekciós és képtranszformációs műveleteket egy gyakorlati példán keresztül.

*Például (pixelgrafikus logó vektoros kiszerkesztése):*



3.) Mutassa be a gyakorlati órákon tanult 3D grafikai képszerkesztő műveleteket egy gyakorlati példán keresztül.

*Például (Autodesk Maya 3D-ben):*



## MINTA VIZSGA ZÁRTHELYI DOLGOZAT MEGOLDÁSAL Képfeldolgozás (GEVAU509BL) c. tantárgyból

---

### 1.) Mutassa be az RGB színrendszert!

A szemünkben a színek érzékeléséért felelős háromféle csap érzékenységi görbének maximuma környékén keltett ingerületek összeadódnak és színérzetet keltenek. A szemünkbe jutó fény komponensei összegződnek (additív színkeverés). Az additív színrendszer alapszínei a vörös (R), a zöld (G) és a kék (B). A komponensek kezdőbetűiből ered a színrendszer másik elterjedt elnevezése: RGB. A színrendszer elemeinek szabványos hullámhosszát az előállíthatóságuk figyelembevételével állapították meg: R: 700 nm, G: 546 nm, B: 435 nm. A szín megadása a háromdimenziós tér egy pontjának kijelölését jelenti. A számítógépes számábrázolás elterjedése miatt a korábbi 1 helyett 255-re szokás normálni az értékeket, tehát az egyes alapszínek a 0-255 tartományt foglalják el. A három alapszínből mintegy 16 millió szín és árnyalat állítható elő a számítógépes képfeldolgozás során, amelynek szemünk csak töredékét képes megkülönböztetni.

Ha egy adott színt akarunk ábrázolni a háromdimenziós RGB térben, a színt a koordinátaival azonosítjuk. Az R tengelyen az origótól (fekete) a pirosig, a G tengelyen a zöldig, a B tengelyen a kékig terjed a skála.

Az RG síkban a kék összetevő nulla. A négyzet origóval szemközi csúcán a sárgát találjuk. Hasonlóan a GB síkban az origótól legtávolabb a kékeszöld helyezkedik el. Az RB síkban a két alapszín maximuma a bíbort jelöli ki. Ha az RB sík átlójából kiindulva a G tengellyel párhuzamosan készítünk egy metszetet, a sík magába foglalja a testátlót, amely az origóbeli feketétől a legtávolabbi csúc fehér pontjáig a szürkeárnyalatokat fedi le.

### 2.) Mutassa be, hogy működik a 3D képalkotás a röntgensugaras tartományban!

A röntgensugaras képfelvétel célja a háromdimenziós képalkotás, amely nem csak a felületi, hanem az alkatrészek alatti és a furatokon belüli viszonyok mérésére is alkalmas. A röntgensugaras megvilágítás mellett készített képsorozat – a kamera és a képfelvevő eszköz mozgatása révén – más-más pozícióból ábrázolja az átvilágított nyomtatott áramköri lemezt. Az így kapott képekből megalkotható a vizsgált terület 3D modellje. Ennek alapján az alkatrészek alatti forrasztások minősége is ellenőrizhető. Fontos jellemző a furatkitöltés, a furatokba felszívódó forrasztanyag magassága, amit más módszerrel nemigen lehet ellenőrizni. A röntgensugaras 3D tesztet a gyártási folyamat végső fázisában célszerű végezni, kiegészítve a látható fény tartományában végzett optikai ellenőrzést.

### **3.) Csoportosítsa a képfeldolgozási műveleteket!**

Több szempont alapján osztályozhatjuk a képfeldolgozási műveleteket. Egyik szokásos felosztás aszerint csoportosítja a műveleteket, hogy a feldolgozás alatt álló pont milyen nagyságú környezete befolyásolja a pont új világosságát. (A példákban rendszerint szürke árnyalatú képekkel fogunk dolgozni és világosság transzformációról beszélünk. Színes képeknél általában háromszor kell elvégezni a műveletet.)

- Pont-pont műveletek: Az adott képpont új világosságát csak ennek a pontnak az eredeti világossága befolyásolja.
- Lokális műveletek: Az adott képpont új világosságára a képpont megadott (pl.: 3\*3 pixeles) környezetének van hatása.
- Globális transzformációk: Az adott képpont új világossága a teljes képtől függ.

### **4.) Röviden mutassa be a szintrevágást!**

Képszegmentálási feladatoknál el kell választani az objektumot a háttértől. Egyik lehetséges megoldás a világosság alapján elvégzett osztályozás. Egy alkalmas világosság küszöb alapján bináris képpé alakíthatjuk a szürke árnyalatú képet. A küszöb kijelölését segíti a hisztogram. Vizuálisan kijelölhetjük az alkalmas vágási szintet. Ennél pontosabb kijelölést tesz lehetővé a hisztogram matematikai elemzése, amellyel pl. Bayes döntési algoritmussal minimalizálhatjuk a hibás osztályba került pontok számát.

### **5.) Mutassa be az erózió és a dilatació műveletét!**

#### ***Erózió***

Az erózió az objektum körbenyírását jelenti, a strukturáló elem sugarával csökken minden irányban az objektum mérete. Nevezik még hámozásnak és fogyasztásnak is. Az erózió végrehajtása során az objektumnak azok a képpontjai maradnak meg továbbra is az objektum részeként, amelyekre ráhelyezve a strukturáló elemet, annak minden pontja az objektumhoz tartozó képpontot fed. Más szavakkal: Az objektum azon pontjait, amelyekre ráhelyezve a strukturáló elemet, annak valamely pontja a háttérhez tartozó képpontot takar, a háttérhez kell sorolni. Az erózió hatékonyan tünteti el az apró objektumokat.

Az erózió alkalmas az összeolvadt objektumok szétválasztására. A strukturáló elem határozza meg, hogy vízszintesen és függőlegesen hány pixelt farag le az objektumból. Az oszlopszerű strukturáló elem az egymás melletti, a vízszintes pálcikaszerű elem az egymás alatti objektumokat választja el hatékonyan.

Az erózió visszafordíthatatlan művelet. Az objektumok sarkai lekerekednek, ezeket nem lehet visszaépíteni. A vékony alakzatok el is tűnhetnek, ezeket lehetetlen visszanyerni.

### ***Dilatáció***

A dilatáció az erózió ellentéte, az objektumot bővítjük a strukturáló elem sugarával. Hizlalásnak is szoktuk nevezni.

A dilatáció eredményeként a háttér azon pontjai, amelyekre a strukturáló elemet ráhelyezve az objektumhoz tartozó képponttal is fedésbe kerül, az objektumhoz fognak tartozni. A dilatáció alkalmas az objektumon támadt lyukak befoltozására. A hizlalás eredményeképpen a szétválasztott, közeli objektumok újra összeolvadhatnak.

Ha a dilatáció után kapott képből kivonjuk az eredeti képet, az objektum körvonalát kapjuk, kívülről körberajzolva. Ha az eredeti képből vonjuk ki az erózió eredményét, a belső körvonalat kapjuk a strukturáló elem sugarának megfelelő vastagságban. Egyvonalas körvonalból megkaphatjuk a terület becslését, ha összegezzük a körvonal, mint objektum pontjainak a számát. Ennél pontosabb mérésre ad lehetőséget az objektum határvonalára fektetett görbe hossza.

### **6.) Lineáris predikció**

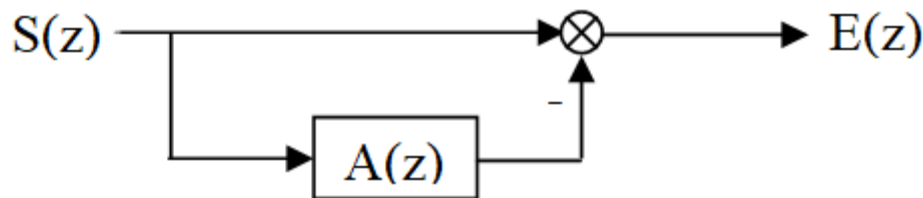
Egy mintasorozat következő mintáját megkíséreljük a megelőző minták lineáris kombinációjával megbecsülni. A megfelelő együtthatók megtalálását az teszi lehetővé, hogy a becslés helyességét az eddigi mintákon ellenőrizhetjük. A hangképző szervek modellezése révén az eljárás fizikai háttérrel is rendelkezik.

A következő minta becslése:  $\tilde{s}(n) = \sum_{i=1}^p a(i)s(n - i)$

A becslés hibája:  $e(n) = s(n) - \tilde{s}(n) = s(n) - \sum_{i=1}^p a(i)s(n - i)$ ,

A rendezett számsorok műveleteit a z-transzformálttal jelölve:

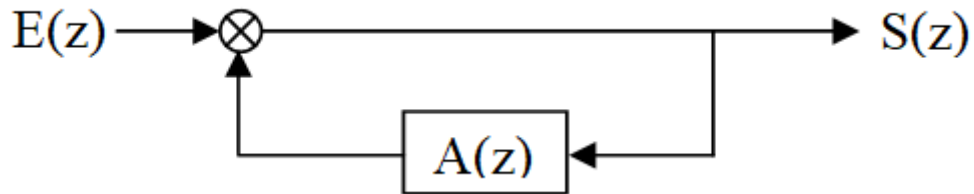
$$E(z) = S(z) - S(z)A(z).$$



A lineáris predikció analízis modellje

Ez az egyenlet írja le a lineáris predikció analízis modelljét. Ha átrendezzük az egyenletet, a lineáris predikció szintézis modelljéhez jutunk, amely lehetővé teszi az együtthatók ismeretében a hibajelből a jel visszanyerését.

$$S(z) = E(z) + S(z)A(z).$$



A lineáris predikció szintézis modellje

6.