

XX. reál- és humántudományi Erdélyi Tudományos Diákköri Konferencia (ETDK)

Kolozsvár, 2017. május 18–21.

# Intelligens képelemzés kompozíciós szabályok alapján

**Szerző:**

**Kálmán Noémi**

Babeş-Bolyai Tudományegyetem, Matematika és Informatika Kar, Informatika szak, III. év

**Témavezető:**

**dr. Bodó Zsolt** egyetemi docens,

Babeş-Bolyai Tudományegyetem, Matematika és Informatika Kar, Magyar Matematika és Informatika  
Intézet

# Tartalomjegyzék

1. Bevezető.....	3
2. Kompozíciós szabályok.....	4
3. Beépített szabályok alkalmazása.....	6
4. Az alkalmazás felépítése, megvalósítás lépései .....	12
4.1. Felhasznált technológiák .....	12
4.2. Elemek detektálása.....	12
4.2.1. Célobjektumok azonosítása.....	13
4.2.2. Vonalak azonosítása.....	14
4.3. Az esztétikai érték kiszámítása .....	15
4.4. A legesztétikusabb képrészlet meghatározása.....	15
4.4.1. A genetikus algoritmus paramétereit.....	17
5. Esettanulmányok, észrevételek .....	18
6. Továbbfejlesztési lehetőségek.....	23
7. Könyvészet.....	25

# 1. Bevezető

A művészek a történelem során mindig is arra törekedtek, hogy esztétikát csempésszenek műveikbe. Az esztétikus alkotások hatnak az érzékeinkre, ezáltal érzelmeket idéznek elő. Egy ilyen mű megalkotása azonban kihívást jelenthet, ugyanis nincs egy egyetemesen elfogadott esztétikai értékrend. Az esztétika megítélése szubjektív, az aktuális érzelmektől és a személyes ízléstől is nagyban függ.

Az előbb említett kihívások ellenére megalakult egy tudományos irányzat, a számítógépes esztétika. Ez arra ösztönösít, hogy megjósolja, egy bizonyos műalkotás milyen érzelmeket vált ki, milyen hatással lesz az alanyra, e mellett célja, hogy erősítse, javítsa a kiváltott hatást. A számítógépes esztétika tehát olyan számítástechnikai eljárások összessége, melyek az emberi szemhez, fülhöz hasonlóan képesek jól alkalmazható esztétikai döntéseket hozni [1].

Ez az irányzat minden művészeti ággal kapcsolatba hozható, így a zene, irodalom és festészet mellett a fényképészetben is alkalmazva van. A dolgozatban a fényképészetben fellelhető esztétikai tulajdonságokkal foglalkozunk, azon kompozíciós szabályokra fektetve a hangsúlyt, melyek betartásával egy sokkal kellemesebb látványt nyújtó képet kapunk.

A kompozíciós szabályok olyan tanácsokat foglalnak magukba, melyek betartásával sokkal sikeresebb képeket lehet készíteni. Azonban az átlagember nem rendelkezik különösebb tapasztalatokkal a fényképészet terén, nincs meg a megfelelő képzettsége, ennél fogva a kompozíciós szabályokat sem ismeri. Ellenben mindenki szereti megörökíteni a számára fontos eseményeket, utazásai során látott tájakat, gyakorlatilag mindent, ami valamiért fontos számára. Hasznos lehet tehát egy olyan alkalmazás, mely képes azonosítani a kompozíciós szabályokat, javítani a kép esztétikai értékén, ezáltal segítséget nyújtani a felhasználónak. Itt jön képbe a számítógépes esztétika használata.

A dolgozatban vizsgáljuk ezen kompozíciós szabályokat, a legrepresentatívabbakat beépítjük a rendszerbe, majd ezek felhasználásával egy bemeneti képnek meghatározzuk a számszerűsített esztétikai értékét. Ezután egy kellőképpen gyors módszert keresünk a bemeneti kép egy olyan részképeinek meghatározására, amely nagyobb esztétikai értékkel rendelkezik, tehát jobban betartja a kompozíciós szabályokat.

## 2. Kompozíciós szabályok

Kompozíció alatt a fényképen található elemek pozícióját, ezek elrendezésének egyensúlyát értjük. E mellett magába foglalja a képek színezetét és textúráját is [3]. Ezen szabályok betartása különbözteti meg az amatőr felvételeket az esztétikus képektől. Egy jól összeállított kép elérése érdekében nem árt előre eltervezni, hogy az elemek pontosan hova is fognak majd kerülni a kereten belül.

Fontosságuk ellenére nehéz meghatározni a kompozíciós szabályokat. A művészet lényege ugyanis a kreativitásban, egyediségben rejlik, éppen ezért nem kell a kompozíciós szabályokat sem kőbe vésett törvényként kezelni. Szabályok helyett tehát sokkal inkább nevezhetők tanácsoknak, melyeket jó megfogadni, ha egy igazán sikeres és maradandó képet szeretnénk alkotni [3]. Előfordulhat az is, hogy két szabály ellentmond egymásnak, ezért választani kell, melyik teszi előnyösebbé a képet.

A legalapvetőbb és legszéleskörűbben ismert a harmadolás szabály. A leggyakoribb hiba az amatőr fotózásban az, hogy minden elem a kép közepére kerül. Ellenben az emberi szem hajlamos jobban felfigyelni arra a képre, mely harmadokra van osztva, és a célobjektum a harmadvonalak metszetében vagy ezen vonalak mentén helyezkedik el [2]. Előnyös például egy emberi alakot vagy fát a kép jobb vagy baloldali harmadába helyezni, illetve a látóhatárt az alsó vagy felső harmadvonal mentére pozícionálni. A legtöbb kamerában van már lehetőség harmadolt nézetben készíteni képeket, segítséget nyújtva az amatőr fényképészeknek.

Az előző szabállyal ellentmondóan néha igenis előnyös, ha a célpont a kép közepére kerül. A szimmetrikus összeállításokat érdemes a kép közepére helyezni [2]. Az utak, vagy épületek, esetleg a vizen visszatükröződő objektumok fényképezésekor hasznos alkalmazni ezt a szabályt.

A harmadvonalak mellett az átlók is esztétikailag fontos szerepet játszanak. Egy átló menti célobjektum dinamikát kölcsönöz a képnek, míg az átlók környékén található vonalak a célobjektum felé irányítják a tekintetet. A vízszintes vagy függőleges vonalak stabilitást kölcsönöznek, míg az átlók dinamikus feszültséget [2]. Ez egyedivé teheti a képet.

Az egyensúly fogalma is központi szerepet játszik a képösszeállítás harmóniájában. Egy kiegyensúlyozott képen az objektumok egyenlő távolságra vannak elosztva a kép középpontja körül. Egy nem megfelelőképpen kiegyensúlyozott kép zavart okozhat, túlsúfolt hatást kelthet [3].

Közismert, hogy a fényképek két dimenzióban örökítik meg a minket körülvevő világot. Bizonyos technikákkal azonban háromdimenziós hatást is kelthetünk. Elérhetjük ezt a közeli és távoli objektumok szétválasztásával, úgy, hogy kihangsúlyozzuk a közelebbi objektumokat, míg a távolabbi, háttérben lévő elemeket kissé összemossuk. Így mélységérzetet lehet kelteni, illetve így a figyelem központjába a közeli célobjektum kerül, nem pedig a színes, túlsúfolt háttér. Ez a mélységélesség megvalósítható hagyományos módon, DSLR gépet használva, megfelelőképpen beállítva az objektív gyújtótávolságát, illetve az élesség távolságát. E mellett mélységélességet elérhetünk megfelelő szoftver használatával, úgy, hogy a háttérben levő elemeket elhomályosítjuk [5].

A háromdimenziós hatás keltésére alkalmas még a keret a keretben alkalmazása. Ilyenkor ideális egy ablakot, ajtót, boltívet vagy faágakat fotózni, úgy, hogy ezek körülöleljék a képet, mintegy keretet alkotva a célobjektum körül.

Megfigyelt tény, hogy az emberek természetes módon vonzódnak a mintákhoz és különböző textúrákhoz. Vonzóak a szem számára és harmóniát sugallnak [4]. Az ismétlődő minták mindig jó módszert biztosítanak a kellemes képek készítésére.

Az elemek számát tekintve a páratlanság szabálya alapjául az a megfigyelés szolgál, hogy az emberi szemnek kellemesebb látványt nyújt a páratlan számú objektumot tartalmazó kép. A páros számú elemet tartalmazó elbizonytalaníthatja a megfigyelőt, elvonja a figyelmet az objektumról, az alany nem tudja mire koncentráljon. A páratlan szám sokkal természetesebb hatást kelt.

Bizonyos esetben az előnyös, ha a célobjektum kitölti a teljes keretet, néha jobb hatást érünk el, ha hagyunk a célpont körül egy kis negatív teret. Azzal, hogy kitöltjük a teljes teret segítjük a nézőt, hogy maximálisan a célpontra összpontosítson, megfigyelje a részleteket. Ezzel ellentétben az is nagyon látványos lehet, ha sok semleges helyet hagyunk a fő objektum körül [3]. Ez is az objektumra irányítja a tekintetet, s egyben ki is emeli a háttérből.

Ezen kompozíciós szabályok egytől egyig elősegítik az esztétikus képek készítését. A kihívás abban rejlik, hogy megtaláljuk az egyensúlyt a különböző szabályok alkalmazásában. Olyan szabályokat kellett tehát keresnünk, amelyek nem mondanak ellent egymásnak. Dolgozatunkban a harmadolás, vizuális egyensúly és átló dominancia szabályok kerültek alkalmazásra, ezek összetételével próbáltuk megfelelően számszerűsíteni egy kép esztétikai értékét.

### 3. Beépített szabályok alkalmazása

Ahhoz, hogy egy képnek számszerűsíteni lehessen az esztétikai értékét, fontos az előbb említett szabályok közül néhányat kiválasztani. Mivel egyes szabályok ellentmondanak egymásnak, olyan szabálycsoportot kellett kiemelnünk ezen halmazból, melyek nem csak jól megférnek egymás mellett, de együttesen egy sokkal esztétikusabb, kellemesebb képet adnak.

A képek esztétikájának megfelelő tükrözéséhez ezen kompozíciós szabályok közül hármat választottunk ki: a harmadolás szabályt, az átlók dominanciáját, illetve a vizuális egyensúlyt [6]. A harmadolás a legalapvetőbb szabály, minden sikeres műalkotásban fellelhető az alkalmazása, ezért ez az általunk kiválasztott csoportba is bekerült. Az átlós vonalak megfelelően irányítják a tekintetet, ezzel hatást gyakorolva az emberi érzékekre, a fő célpontra terelik a figyelmet, így a néző nem érezheti semmitmondónak a képet. Emellett az emberi lélek igényli az egyensúlyt, így a képekben is hajlamos az egyensúlyt részesíteni előnyben. A rendezetlen, túlszűfolt, majd lappangó területeket tartalmazó képek zavart okozhatnak, a néző elveszítheti a kép mondandóját. Ezért esett a választás, az előbb említett két szabály mellett, a vizuális egyensúly beépítésére is.



(a)



(b)



(c)



(d)

### 1. Ábra: A felhasznált kompozíciós szabályok

(Források: (a) <http://bonuspastormarr.weebly.com/>,

(b) <http://eastdigitalarts.weebly.com/photo-daily-blog>,

(c) <http://keywordsuggest.org/gallery/861757.html>,

(d) <https://pixabay.com/en/trees-summer-nature-landscape-1589190/>)

Az 1. ábrán látható képek a felhasznált szabályokat mutatják be. Az első, (a) kép a harmadolás szabályt reprezentálja. Jól látható, hogy a világítótorony a jobboldali harmadvonal mentén helyezkedik el, mindkét súlypontot magába foglalja. A tőle jobbra található (b) képen pedig a horizont az alsó harmadvonallal teljesen megegyezik. A harmadik, (c) kép az átlódominanciát mutatja, ugyanis a lépcső szélei pontosan az átlót alkotják. Az utolsó, (d) kép pedig a vizuális egyensúly szabályát tartja be, a fák egyenlő távolságra helyezkedve el a kép középpontjától.

Ahhoz, hogy a kiválasztott szabályokat megfelelően alkalmazni tudjuk a képértékelésben, bizonyos metrikákat kellett felhasználnunk. A céljainknak legmegfelelőbbnek az L. L. Renjie, C. L. Wolf és D. Cohen-Or által kidolgozott metrikák [6] bizonyultak. Kisebbs változtatásokkal ezen metrikákat implementálva alakult ki az esztétikai értéket számszerűsítő alkalmazás.

A képhez rendelt esztétikai értéket a célobjektumok és fő vonalak pozíciója valamint térbeli felépítése alapján határoztuk meg. Ehhez azonban először azonosítanunk kellett ezen számottevő objektumokat. A metrikák implementációját tehát a képen fellelhető objektumok detektálása kellett megelőzze, melyhez a jól ismert algoritmusokat alkalmaztuk.

A célobjektumok azonosításához előbb egy sor képfeldolgozó algoritmust alkalmaztunk, mely megfelelően összemossa a képet, ezzel eltüntetve a felesleges éleket, fekete fehérré

alakítottuk, majd megfelelő küszöbértéket alkalmazva 0 és 1 közé redukáltuk a pixelek értékeit. Ezután egy Canny operátort [7] alkalmazva azonosítottuk a képen található fő kontúrokat. Ha mindez megtörtént, a kontúrokat magukba foglaló téglalapok képezik a kép objektumait. Ez azonban igen sok bekeretezett elemet eredményez. Sok közülük egymással átfedésbe is kerülhet, ezáltal egy objektumot akár több téglalappal is lefedhetünk, ami felesleges. Ennek elkerülése érdekében az egymással átfedésben levő téglalapokat egybeolvasztjuk, így kevesebb darab, nagyobb területű célobjektumot kapva eredményül.

A kép túlzott összerosása miatt fontos objektumokat figyelmen kívül hagyhatunk, ami gondot okozhat. Ilyen gyakran előfordul az arcokat tartalmazó képek esetén, annak ellenére, hogy az emberi arc fontos szerepet játszik az elkészített képekben. Az ilyen hibák elkerülése végett a képen egy arcdetektálást is végeztünk, hogy ezek mindenképp bekerüljenek a talált objektumok közé. A felfedezett arcokat nagyobb súllyal ellátva adtuk hozzá a főobjektumok listájához.

A vonalak azonosítása is képfeldolgozó algoritmusok alkalmazásával valósítható meg. Első körben Canny éldetektáló operátort [7] használtunk, úgy, hogy előtte nem mostuk különösebben össze a képet. A vonalak keresésekor fontos, hogy minél élesebb képpel dolgozzunk, ugyanis összerosás esetén értékes információt veszíthetünk, ezért előfordulhat, hogy bizonyos vonalak nem lesznek megtalálva. E lépés után egy Hough-féle vonaldetektáló algoritmust [7] alkalmaztuk. Ezt megfelelően paraméterezve eredményül meg is kaptuk a képen fellelhető vonalak halmazát. Ez azonban igen sok vonalat hajlamos magába foglalni, ami az értékelő algoritmusra nézve nem előnyös, túl sok számítást vonhat maga után, ami időigényessé teheti az egész folyamatot. Ennek elkerülése érdekében a vonalak közül csak az első  $n$  darab, legnagyobb súllyal rendelkező vonalat tartottuk meg. Egy vonal minél hosszabb, annál nagyobb a súlya.

Az általános vonalak mellett számon kellett tartanunk az átlóhoz közel eső vonalakat is. Az átló közelséget nem a vonal pozíciójára értjük, hanem arra, hogy mennyire áll az átlóhoz hasonló szögben. Az átlók szögéhez viszonyítva  $10^0$ -os eltérést engedünk meg. Az átlós vonalak listáját az átlódominancia szabály vizsgálata miatt kellett eltárolnunk.

Ismerve a főobjektumokat és vonalakat, egy olyan 0 és 1 közötti értéket definiálhatunk, mely kellőképpen mutatja, hogy egy kép milyen szinten tartja be a fent említett szabályokat.



A dolgozatban használt szimbólumok jelentései a 2. ábrában fel vannak tüntetve. Az algoritmusainkban a távolság alatt a hagyományos két pont közötti távolság helyett mindig a normalizált Manhattan távolságot értjük, melyet a következőképpen definiálhatunk :

$d_M((x_1, y_1), (x_2, y_2)) = \frac{|x_1 - x_2|}{w} + \frac{|y_1 - y_2|}{h}$  [6], illetve  $d_L(L, M)$  nem más, mint két vonal, L és M közti távolság. Két vonal közötti távolságot úgy számoljuk ki, hogy L két végpontjától kiszámítjuk az M legközelebbi pontjához mért Manhattan távolságát, majd ezt átlagoljuk.

$I(S_i)$  alatt az aktuális, i-edik objektum súlyát értjük, ami nem más, mint az objektum területe, míg  $I(L_i)$  az i-edik vonal súlya, azaz ennek a hossza.  $\bar{X}$  az összes átlagos vonalat jelöli, míg  $X$  az összes átlós vonalat, azaz azon vonalakat melyek legfeljebb  $10^\circ$ -ban térnek el a  $Q_1$  és  $Q_2$  által jelölt átlók valamelyikétől.

Szimbólum	Jelentése
$w, h$	a kép szélessége és hosszúsága
$C$	a kép középpontja
$G_i, i = \overline{1,4}$	a négy súlypont
$R_i, i = \overline{1,4}$	a négy harmadvonal
$S_i, i = 1, 2, \dots, n$	a detektált objektumok
$C(S_i), I(S_i)$	az objektum középpontja és súlya
$r(S_i)$	objektum méret – képméret arány
$Q_1, Q_2$	az átlók
$L_i, i = 1, 2, \dots, n$	a detektált vonalak
$X$	átlós vonalak halmaza
$\bar{X}$	nem átlós vonalak halmaza
$I(L_i)$	a vonal súlya
$d_M$	normalizált Manhattan távolság
$d_{LM}$	két vonal közötti távolság

**2. Ábra** : a dolgozatban felhasznált szimbólumok

**A harmadolás szabály (Rule of Thirds, RT)** kiszámítása két fő komponensből tevődik össze :  $E_{Point}$  a célobjektumoknak a súlypontoktól mért távolságát méri, míg az  $E_{Line}$  a vonalaknak a harmadvonalaktól mért távolságát tükrözi. A képletben szereplő  $\gamma_{Point}$  és  $\gamma_{Line}$  súlyok. Az összképlet ezen komponensekből a következő képpen épül fel [6]:

$$E_{RT} = \gamma_{Point} E_{Point} + \gamma_{Line} E_{Line}$$

A pont értéke az összes célobjektumnak :

$$E_{Point} = \frac{1}{\sum_i I(S_i)} \sum_i I(S_i) e^{-\frac{D^2(S_i)}{2\sigma_1}}$$

ahol  $D(S_i)$  az aktuális célobjektumnak a négy, súlyponthoz mért minimális távolságát jelenti,  $D(S_i) = \min_{j=1,2,3,4} d_M(C(S_i), G_j)$ , illetve  $\sigma_1$  konstans egyenlő 0.17 .

A vonal értéke az összes megtalált, nem átlós vonalnak :

$$E_{Line} = \frac{1}{\sum_{i \in \bar{X}} I(L_i)} \sum_{i \in \bar{X}} I(L_i) e^{-\frac{D_R^2(L_i)}{2\sigma_2}}$$

ahol  $D_R(L_i)$  a minimum vonaltávolság az aktuális vonaltól a négy harmadvonalhoz mérve,  $D_R(L_i) = \min_{j=1,2,3,4} d_L(L_i, R_j)$  illetve a  $\sigma_2$  konstans egyenlő 0.17 .

Kísérletek azt bizonyították, hogy a vonalak nagyobb szerepet játszanak az esztétikai érték kialakulásában, ezért nálunk is nagyobb súlyt kap a vonalérték. Ennek megfelelően a képletben szereplő  $\gamma_{Point} = 1/3$ , míg a  $\gamma_{Line} = 2/3$  -nak választottuk meg.

**A vizuális egyensúly (Visual Balance, VB)** –ről akkor beszélhetünk, ha a célobjektumok összességének középpontja közel található a képnek a középpontjához. Ennek megfelelően a vizuális egyensúly értékének kiszámítása [6]:

$$E_{VB} = e^{-\frac{d_{VB}^2}{2\sigma_3}}$$

ahol  $d_{VB} = d_M(C, \frac{1}{\sum_i I(S_i)} \sum_i I(S_i) C(S_i))$  .

**Az átlódominancia szabály (Diagonal Dominance, DD)** értékének kiszámítása a vonalérték kiszámításához teljesen hasonló módon történik, annyi különbséggel, hogy itt csak az átlós vonalakkal dolgozunk. Ennek megfelelően [6]:

$$E_{DD} = \frac{1}{\sum_{i \in X} I(L_i)} \sum_{i \in X} I(L_i) e^{-\frac{D_Q^2(L_i)}{2\sigma_2}}$$

ahol  $D_Q(L_i) = \min(d_L(L_i, Q_1), d_L(L_i, Q_2))$ .

**Az esztétikai érték függvény (Aesthetic score function, A)** a fennebb említett szabályok által kiszámított értékek kombinációja [6]:

$$E_A = \frac{\omega_{RT}E_{RT} + \omega_{VB}E_{VB} + \omega_{DD}E_{DD}}{\omega_{RT} + \omega_{VB} + \omega_{DD}}$$

ahol  $\omega_{RT}$ ,  $\omega_{VB}$  és  $\omega_{DD}$  súlyok, és rendre 1.0, 0.3, illetve 1.0 értékűek.

**A célobjektum mérete (Salient-region Sizes, SS)** is fontos tényező az összetett esztétikai érték számolásában. Olyan esetben, amikor egy képen csak egy szabályt ismerünk fel, a fent említett összefüggés nem írja le elég pontosan az adott kép esztétikai értékét. Előfordulhat, hogy az adott kép több részképe megegyező esztétikai értéket kap, ezáltal nem lehet objektíven eldönteni azt, hogy melyik a legsikeresebb ezen képrészletek közül. Példának okáért vegyük azt az esetet, amikor a kép csak egy célobjektumot tartalmaz, akkor azt a négy súlypontra helyezve ugyanazt az esztétikai értéket kapnánk. Ez gondot okozhat. Ennek elkerülése érdekében vezettük be a célobjektum vagy célobjektumok méretét, mint értékelési szempontot.

Esztétikus képek tanulmányozása során észrevehető, hogy a célobjektumok három csoportba sorolhatóak : kis, közepes és nagy méretűek. A dolgozatban előnyben részesítjük azon objektumokat, melyek ezt az eloszlást követik. A három méretnek az alábbi arányok felelnek meg:  $r_1 = 0.1$ ,  $r_2 = 0.56$  és  $r_3 = 0.82$ . Az átlunk használt méret-függvény az ehhez konvergáló arányokat támogatja [6]:

$$E_{SS} = \sum_i \max_{j=1,2,3} e^{-\frac{(r(S_i) - r_j)^2}{2\tau_j}}$$

ahol  $\tau_1 = 0.07$ ,  $\tau_2 = 0.2$  és  $\tau_3 = 0.16$  konstansok ezen arányok előfordulását reprezentálják.

**Az összetett esztétikai érték (Combined Aesthetic Score)** a kiszámított  $E_A$  és  $E_{SS}$  értékekből tevődik össze [6]:

$$E = (1 - \omega_{SS})E_A + \omega_{SS}E_{SS}$$

ahol  $\omega_{SS} = 0.08$ , ez az érték azért olyan kicsi, mivel a hangsúly az esztétikai értéken van, a célobjektum mérete csak másodlagos, a határozatlan esetek elkerülése végett lett beépítve a rendszerbe.

Elvégezve az objektumokat azonosító lépéseket, illetve alkalmazva a fent említett képleteket, eredményül egy 0.0 és 1.0 közötti értéket kapunk. Optimális esetben a kép esztétikai értéke 1.0 lenne, ez a gyakorlatban 100% - nak felelne meg.

## **4. Az alkalmazás felépítése, megvalósítás lépései**

### **4.1. Felhasznált technológiák**

Az alkalmazás megírásához a Java programozási nyelvet választottuk, míg a képelemző algoritmusok használatához az OpenCV nyílt forráskódú könyvtárat. Az OpenCV használatára azért esett a választás, mivel ez az alapvető képfeldolgozási algoritmusokat magába foglalja, nagyon gyors és hatékony megoldásokat biztosítva [13]. Alkalmazásunk sok számítást kell végezzen, ezért szükségeltett a performanciát szem előtt tartani.

Algoritmusunk három fő részre bontható: az első a képen található elemek azonosítása; a második az esztétika számszerűsítése, a szabályoknak megfelelő metrikák implementációja; a harmadik pedig a legesztétikusabb képrészlet megtalálása.

### **4.2. Elemek detektálása**

A képen található elemek azonosítása a célobjektumok, arcok, illetve vonalak detektálását takarja. Ezekre szükségünk van a második lépésnél, az esztétika számszerűsítésénél. Az elemek azonosítása a legidőigényesebb, ezért ezt a lépést az elemzés során célszerű csak egyszer

elvégezni. Az azonosított objektumokat és vonalakat listákban tároljuk, hogy az elemzés során ezt bármikor használni tudjuk.

#### 4.2.1. Célobjektumok azonosítása

A célobjektumok azonosításában első lépésként kissé össze kellett mosnunk a képet, ezáltal a felesleges élektől megszabadulni, míg a fontosakat hangsúlyosabbá tenni. Erre a célra a piramis átlagoló szűrőt [10] alkalmaztuk. Az algoritmus egy  $maxlevel+1$  szintet tartalmazó Gauss piramist épít, az összerosó algoritmust mindig a kisebb szinteken alkalmazva először, majd az eredményeket átadva a magasabb szinteknek. Itt újból alkalmazva van az összerosás, de csak azon pixelekre, melyek adott határnál jobban eltérnek a nekik megfelelő alacsonyabb szintű pixelektől. Ez a színes területek körvonalait élesebbé teszi, a kevésbé hangsúlyos éleket pedig teljesen eltünteti. Az algoritmus paramétereit próbáltuk úgy beállítani, hogy a leginkább helyes eredményt adja: az adott pixel körüli vizsgálandó pixelek számát a kép területének 3%-val tettük egyenlővé, a pixelérték küszöböt 100-ra, míg a piramis szintjeinek számát 2-re állítottuk.

Következő lépésben a képet fekete-fehérré alakítottuk, majd adaptív küszöbértéket (adaptive threshold) [8] alkalmazva 0 és 1 közé szorítottuk a pixelek értékét. Mivel a kép különböző pontjai más-más megvilágításban lehetnek, nem lenne célravezető egy globális küszöbértéket használni. Ezért esett a választás az adaptive threshold alkalmazására, ahol kisebb régióknak saját küszöbértéket számolunk, így különböző helyeken más-más küszöbértékkel dolgozunk, mely a fényviszonyoktól függetlenül jobb eredményeket ad.

Ezután az élek felfedezése következik. Erre a lépésre a legalkalmasabbnak a Canny – féle éldetektálás [9] bizonyult. Ez egy sor algoritmusnak az alkalmazása, mely magába foglalja a kép összerosását, a zavarok kiszűrését, konvolúciós maszk alkalmazását, a nem éleket képező pixelek kiszűrését, illetve utolsó lépésként, két küszöbérték alkalmazását : ha az alsónál kisebb értékűek a pixelek, akkor nem képeznek élt, ha a felsőnél nagyobbak, akkor élt képeznek, illetve ha a kettő között vannak, csak akkor képeznek élt, ha van olyan szomszédos pixelük, mely élhez tartozik. Észrevehető, hogy a kezdetekkor már alkalmaztunk egy összerosást és a Canny-féle algoritmusban is található egy. Azért volt szükséges egy összerosást alkalmazni előre, mivel a Canny algoritmusban nincs megadva a választás lehetősége: Gauss-féle összerosást használ [15], nekünk azonban a piramis átlagoló a megfelelőbb. E lépés eredményeképp a kimeneti kép már csak a kontúrokat tartalmazza.

Ezután már csak külön kellett választani ezen kontúrokat egy pontok halmazát tartalmazó listába, és az adott ponthalmazok egy-egy téglalapba helyezhetőek. Ez még nem jelenti az algoritmus végét, ugyanis az így kapott eredményben egy objektumot több kisebb résztéglalap reprezentálhat. Ezt célszerű volna egy téglalapba tömöríteni. Az egyetlen hátramaradt lépés tehát az egymást metsző vagy átfedő téglalapok egyesítése volt. Az így kapott nagyobb területű téglalapok már a célobjektumokat reprezentálják.

A kép színekombinációjából adódhat olyan probléma, hogy az emberi arc nem lesz célobjektumként azonosítva. Ez probléma, ugyanis az emberi arcok fontos szerepet játszanak a képeken. Annak érdekében, hogy az emberi arcok mindenképp bekerüljenek a célobjektumok közé, az eredeti képen alkalmaztuk az OpenCV által kínált arcfelismerést, így az már nem fordulhat elő, hogy az arcok kimaradnak a célobjektumok listájából.

#### 4.2.2. Vonalak azonosítása

A vonalak azonosítása esetén nincs szükség a kép nagy mértékű összerosására. Itt fontosak az eredeti élek, kockázatos volna eltüntetni egyes éleket, hiszen ezzel akár fontos vonalakat is elveszíthetnénk. A vonalkereső algoritmusunk egy Canny-féle élkereséssel kezdődik, ahol a kis és nagy küszöbérték rendre 100 és 200.

A megtalált élek közül a vonalakat egy valószínűségeen alapuló Hough-féle vonaltranszformációval kapjuk meg. A Hough-féle transzformációban a vonalak polárkoordinátában vannak leírva. Így egy vonal a következőképpen adható meg:  $y = (-\cos \theta / \sin \theta) x + (r / \sin \theta)$ , ahonnan egyszerűen kifejezhető:  $r = x \cos \theta + y \sin \theta$ . Tehát egy  $(x_0, y_0)$  ponton áthaladó egyenesek halmazát a  $(r_\theta, \theta)$  párosok képezik.

Algoritmusunkban az  $r = 1 \text{ pixel}$  és a  $\theta = 1^\circ$ , illetve a vonalhoz szükséges minimális metszéspontok száma, a minimális vonalhossz és a maximális vonal közti szakadás a fénykép méreteitől függ. A minimális metszéspontok számának a szélesség és magasság minimumának a 10% -át vettük, a minimális vonalhosszt az átlónak 25% -a, illetve a maximális vonal közti szakadást az átlónak 4% -a képezi. A Hough-féle vonaldetektálás az előbb említett paramétereket használva egy vonallistát ad eredményül. Ez azonban egyes képeknél igen sok, felesleges vonalat is tartalmaz. Éppen ezért a talált vonalaknak csak kis százalékkal foglalkoztunk. Kiválasztottunk maximálisan 6 darab legértékesebb vonalat. Egy vonal annál értékesebbnek számít, minél nagyobb a hossza.

A átlódominancia szabály vizsgálatához külön számon kellett tartanunk az átlóközei vonalakat is. Ennek érdekében kiszámoltuk az átlók szögét, és a kapott vonallistából kerestük azon vonalakat, melyek ezekhez viszonyítva egy bizonyos küszöbön belül helyezkednek el. Ezt a küszöbértéket  $10^\circ$ -ra állítottuk be. Sorra ellenőrizve a vonalakat, a határon belül esőket hozzáadtuk az átlókat tartalmazó listához.

### 4.3. Az esztétikai érték kiszámítása

Az adott kép esztétikai értékének kiszámításához a fentebb említett metrikákat szükséges implementálni. Az algoritmusunk során nem csupán az eredeti, teljes nagyságú képnek kellett kiszámítani az esztétikai értékét, hanem a javasolt képrészletnek is. Felmerül tehát a kérdés, hogy minden ilyen képrészletre érdemes-e újból megkeresni a célobjektumokat és vonalakat. Performancia szempontjából semmiképp sem. Az objektumok megkeresése időigényes folyamat, ezért nem lenne célravezető minden alkalommal újból es újból azonosítani ezeket. Pontosan ezért volt hasznos az előző lépésekben minden megtalált objektumot listákba elmenteni. Ilyen módon ezeket a későbbiek során bármikor hasznosítani lehet.

Egy képrészlet esetén előfordulhat, hogy bizonyos objektumok vagy vonalak teljesen, vagy csak részben kiesnek a keretből. Ilyen esetben nem dolgozhatunk az eredeti objektumlistával. Az ilyen célobjektumok esetén meg kell keresnünk azt a téglalapot, mely a célobjektum és a keret metszete, ez lesz az aktuális képrészletben használt objektum. Ha egy adott objektum teljesen a kereten kívülre esik, azt nem vesszük számításba az adott értékszámolásban. A kilógó vonalak esetén hasonlóképpen járunk el. Az a vonalrészlet lesz az aktuálisan használt vonal, mely belóg a képbe, vagy adott esetben, ha a vonal teljesen a kereten kívül esik, akkor ezt nem használjuk az értékszámolásban.

### 4.4. A legesztétikusabb képrészlet meghatározása

A dolgozatban tárgyalt algoritmus utolsó lépését a legesztétikusabb képrészlet meghatározása képezi. Ezt a képtimalizációs lépést keresési algoritmussal valósítjuk meg. Egy kép azonban nagyon sok lehetséges részképpel rendelkezhet. Egy ilyen lehetséges képrészletet 4 paraméterrel írhatunk le: a képrészlet bal felső sarkának  $x$  és  $y$  koordinátája, illetve a képrészlet szélessége és hosszúsága,  $(x_i, y_i, w_i, h_i)$ .

Egy kép esetén tehát  $x_i \in [0, w]$ ,  $y_i \in [0, h]$ ,  $w_i \in [0, w]$  és  $h_i \in [0, h]$  intervallumok jöhetnek szóba, ami nagyon nagy számú lehetséges kombinációt adna. Ez egy  $480 \times 320$  – as

kép esetén, ami nem számít nagy felbontásúnak,  $480^2 * 320^2 \approx 32$  *milliárd* volna. Ez nagyon hosszú időt venne igénybe, ugyanis ennyiszor kellene az értékelő algoritmust végrehajtani.

A képtimalizációs algoritmus első lépéseként tehát csökkentenünk kellett a lehetséges esetek számát. Ezt a legegyszerűbben az intervallumok méretének csökkentésével érhetjük el. A kép minőségének megőrzése is fontos szempont, ezért a képet legfeljebb az eredeti területének negyedére csökkentjük. Ez nagyon kis felbontású képek esetén már kellemetlen hatást okozhat, azonban algoritmusunkban feltételezzük a nagy felbontású képek használatát. Az algoritmusban használt intervallumok ennek megfelelően a következőképpen változnak :  $x_i \in [0, w/2]$ ,  $y_i \in [0, h/2]$  illetve  $w_i \in [w/2, w]$ . Szintén segítségünkre válhat az, ha megtartjuk a kép eredeti arányait. Ennek megfelelően elegendő a kép szélességét és a *szélesség/hosszúság* arányt ismernünk, ebből ugyanis egyszerűen kiszámítható a hosszúság.

Az intervallumok csökkentésével a lehetséges kombinációk számát is nagyon lefaragtuk, azonban ez még mindig nem elég. A lehetséges részképek száma még mindig túl nagy ahhoz, hogy ezeknek sorra kiszámoljuk az esztétikai értékét. Szükség van tehát egy olyan, eléggé gyors módszerre, mely megoldja az optimalizációs problémánkat. Erre legalkalmasabbnak a genetikus algoritmus használatát találtuk.

A genetikus algoritmusok olyan keresőalgoritmusok, melyek a genetikára és a természetes szelekcióra épülnek. Általában optimális vagy az optimálishoz közel álló megoldások megtalálására használják, olyan problémák esetén melyeket nagyon hosszú időbe telne megoldani a hagyományos programozási algoritmusokkal. Az optimalizációs problémák esetén az a kihívás, hogy megtaláljuk azon bemeneti értékeket, melyre a legjobb étéket kapjuk [11]. Esetünkben keressük azon  $(x_i, y_i, w_i)$  számhármast, mely a legnagyobb esztétikai értékkel rendelkezik. Ez a számhármast reprezentál tehát egy egyed. Ilyen egyedekből építjük fel a populációt, ezen hajtjuk végre a genetikus algoritmus függvényeit: a kiválasztást, a keresztezést és a mutációt.

A genetikus algoritmusok problémáról problémára változhatnak, de a következő részeket mindig tartalmazniuk kell: inicializáció, szelekció, szaporítás és leállítás. Az inicializáció során a kezdeti populációt generáljuk. Ennek mérete feladattól függően változó. Általában a egyedek a keresési téren belül egyenletesen oszlanak el. Néha hasznosabb a jónak vélt egyedekkel inicializálni a kezdeti populációt, ez viszont a globális helyett lokális optimum megtalálásához vezethet [12].



Minden generációban az egyedek egy része kiválasztásra kerül. Ezek általában a fitness függvény értéke szerint lesznek kiválasztva. Bizonyos genetikus algoritmus változatokban minden egyed kiértékelésre kerül, s a legjobb értékűek lesznek szelektálva. Ez a teljes folyamat azonban túlságosan hosszú lehet, így előnyössé válik a véletlen egyed kiválasztása.

Az újabb egyedek bevezetése kereszteződéssel, illetve mutáció művelettel lehetséges. A kereszteződés esetén két meglévő egyedből hozunk létre egy újat, míg a mutáció esetén egy egyedet módosítva kapunk újabbat.

Végül, a genetikus algoritmusok egy leállási feltétel teljesüléséig futnak. Ez a feltétel lehet egy előre rögzített maximális generációs szám, illetve a maximális fitnessérték elérése.

#### **4.4.1.A genetikus algoritmus paraméterei**

Algoritmusunkban a kezdőpopuláció létszámának 200 – at választottunk, hogy minél nagyobb valószínűséggel kapjuk már első körben egy, az eredetnél nagyobb esztétikai értékkel rendelkező képet, illetve hogy minél változatosabb legyen a kezdeti populációnk. Az algoritmus evolúciójának leállási feltételének 500 generációt állítottunk be. Természetesen az algoritmus akkor is leállási ponthoz érkezne, ha megtalálná a maximálisan esztétikus képet, amelynek értéke 1.0 lenne. Kísérletek elvégzése után, illetve figyelembe véve, hogy nincs olyan kép mely minden szabálynak tökéletesen megfelelne, biztosan elmondható, hogy egy kép esetén sem fogunk maximális értéket kapni. Éppen ezért volt szükséges egy leállási feltételnek megfelelő maximális generációs számot megadnunk.

A genetikus algoritmusban bajnokság (tournament) szelekciót alkalmaztunk. Ez a szülők megválasztási módja. Szülőválasztáskor tehát véletlenszerűen kiválasztunk 20 darab egyedet, ezek közül megkeressük a legnagyobb esztétikai értékkel rendelkezőt, ez lesz az egyik szülő. A másik szülő választásánál is hasonlóképpen járunk el. E két szülőt megfelelően keresztezzük : mindhárom gén, tulajdonság esetén (x, y és szélesség) generálunk egy számot. Ha ez a szám kisebb mint a 0.5 - ös küszöbérték, akkor az első, különben a második egyed tulajdonságát örökli meg a gyermek. Ha a generálás során a kapott keret kilógna a képből, akkor addig választjuk újra a szülőket, majd keresztezzük, amíg a kapott keret teljesen az eredeti képen belül nem lesz.

A mutáció során hasonlóképpen járunk el. Véletlenszerűen választunk egy értéket. Ha ez kisebb mint a mutációs küszöbhatár, akkor a megfelelő gént újrageneráljuk. Ha a kapott mutált egyed nem található a képen belül, akkor addig generáljuk újra, míg olyan keretet nem kapunk,

mely teljesen az eredeti képen belül helyezkedik el. Az általunk használt küszöbhatár 0.3. A mutáció beépítése azért fontos, hogy elkerüljük a lokális optimumok megtalálását, célunk ugyanis az összes lehetőség közül a legjobbat, azaz a globális optimumot megtalálni.

Fontos megemlíteni, hogy a genetikus algoritmusban elitista szelekciót alkalmaztunk tehát előnyben részesítettük az eliteket, azon egyedeket, melyek nagy esztétikai értékkel rendelkeznek. Minden generációban kiválasztottuk tehát a legsztétikusabb részképet, majd első lépésként ezt hozzá is adtuk a következő generáció egyedeihez.

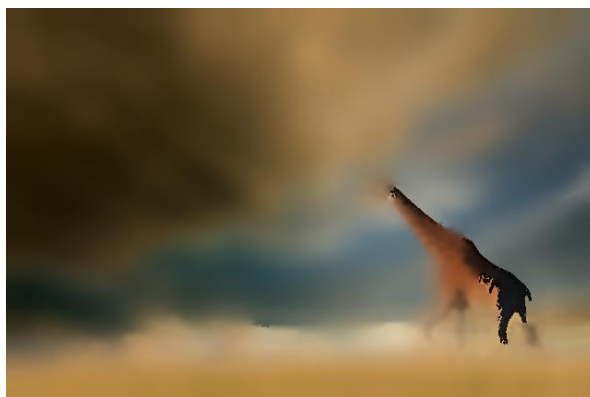
Ezen módszerrel elérve a maximális generációs számot és kiválasztva a legsztétikusabb képrészletet, ha nem is a legoptimálisabb, de jelentősen jó megoldást kaptunk viszonylagosan rövid idő alatt.

## 5. Esettanulmányok, észrevételek

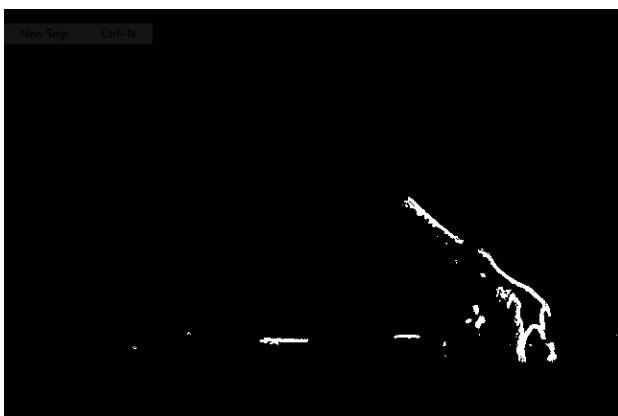
A következőkben néhány konkrét eseten keresztül mutatjuk be az algoritmusunk működését, felhívva a figyelmet az algoritmus tulajdonságaira. Példáinkon keresztül bemutatjuk az objektumok azonosításának lépéseit, a szabályokat képviselő függvények értékeit, illetve a bemeneti és eredményül kapott képet is. Az első esettanulmány lépései:



**3a). Ábra:** A bemeneti kép.



**3b).Ábra:** Az összerosott kép.



**3c). Ábra:** A kép threshold alkalmazás után.



**3d). Ábra:** A kép a detektált célobjektumokkal



**3e). Ábra:** Simán thresholdolt kép



**3f). Ábra:** A detektált vonalak

(Forrás: <https://www.outdoorphotographer.com>)

A fenti ábra az algoritmus különböző fázisaiban mutatja be a bemeneti a képet. A 3b). ábra a kezdeti összemérés utáni állapotot örökíti meg, látszik, hogy a felesleges élek teljesen eltűntek, csak a zsiráf mentén maradtak meg láthatóak. A 3c). ábra a fennmaradt éleket mutatja, ez alapján lettek azonosítva a 3d). ábrán bekeretezett objektumok. A következő ábra az összemérés nélkül alkalmazott thresholdolást mutatja, itt látható, hogy sokkal több él maradt, amire szükségünk is van a vonalak azonosításához. Végül az utolsó, 3f). ábrán a megtalált vonalak vannak feltüntetve. Bár az előző, 3e). kép azt a látszatot kelti, hogy sokkal több vonal fedezhető fel, mégis sok vonal azonosítva. Ez annak köszönhető, hogy a többi vonaljelölt vagy túlságosan rövidnek bizonyult, vagy túl nagy szakadások darabolták fel őket.

A kezdeti kép esztétikai értéke 0.666046 volt, ami azt jelenti, hogy a kép megközelítőleg 66% - ban tartotta be a kompozíciós szabályokat. Egész pontosan a harmadolás szabály értéke 0.770615, a vizuális egyensúlyé 0.421820, illetve az átlódominanciáé 0.947065 volt.

Ez a kapott érték nem nevezhető jónak. Szükséges tehát az optimalizációs algoritmussal megkeresni azon képrészletet, mely maximális eredményt szolgáltat. Az algoritmus eredményeképpen egy olyan képrészletet kaptunk, melynek bal felső sarkát a (191, 123) pont képezi illetve szélessége 405, míg hosszúsága 270 pixel. Ennek a képnek az esztétikai értéke jelentősen jobb, 0.954201, azaz megközelítőleg 95% - os. Észrevehető, hogy mindhárom szabály értéke nőtt: a harmadolás szabály értéke 0.956310, a vizuális egyensúlyé 0.955622, illetve az átlódominanciáé 0.957842 lett. Végül a 3g). ábra mutatja a kapott eredményt.



**3g). Ábra:** Az eredményül kapott, optimalizált kép.

Az algoritmus többszöri végrehajtása során észrevettük, hogy a túlszűfolt, vagy túl közeli képek esetén a kép területének nagyon nagy százaléka a célobjektumot képezi, amint ez a 4. ábra képein is látható. Ez annak köszönhető, hogy az algoritmus nagyon sok, összefüggő élet talál. Az ilyen képek nem lehetnek igazán esztétikusak.



**4. Ábra:** Túlzsúfolt vagy túl közeli képek, a kék keret a célobjektumokat jelöli

(Források: <http://robbygotay.weebly.com/3/post/2013/04/group-photos.html>,

<http://www.freeiconspng.com/free-images/tree-png-714>,

<https://media-cdn.tripadvisor.com/media/photo-s/03/f9/eb/22/peppermill-resort-spa.jpg>)

A 4. ábrán mutatott képek alapján könnyű levonni a következtetést: nem előnyös ilyen túlzsúfolt képekre alkalmazni az algoritmust. A közeli képeket is érdemes elkerülni, ugyanis az algoritmus részlet kivágásával, azaz beközelítéssel próbálja feljavítani a képet, egy nagyon közeli kép esetén azonban már ezen nem lehet ilyen módszerrel optimalizálni.

Egy másik fontos észrevétel, hogy a jó felépítésű, 0.8 feletti esztétikai értékkel rendelkező képek esetén az algoritmus hajlamos ezt az értéket optimálisnak elfogadni. Nem minden esetben találja meg az ennél optimálisabb értékeket. Látványos javítást 0.6-0.7 - nél kisebb esztétikai értékkel rendelkező képek esetén érhetünk el.

Az algoritmus megírásában a legfőbb kihívást a megfelelő képfeldolgozási paraméterek megtalálása jelentette. Jelen paraméterek mellett is előfordulnak olyan képek, melyeken az algoritmus nem talál meg minden kiemelkedő területet. Ez természetesen befolyásolja az esztétikai érték kiszámolását is. Azonban nem lehet általánosan mindenre jól működő paramétereket találni. Minden kép egyedi a maga módján, így ami az egyik kép esetében nagyon jó paraméterhalmaz, az egy másik kép esetén nagyon rossznak bizonyulhat. Célunk tehát egy általánosan elég jól működő paraméteregyüttes megtalálása volt.



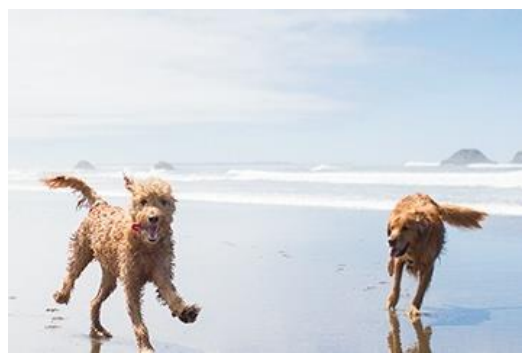
(a)  $E = 0.899226$



$E = 0.94805786$



(b)  $E = 0.511398$



$E = 0.861275$



(c)  $E = 0.919799$



$E = 0.930881$



(d)  $E = 0.899165$



$E = 0.921564$

(Források: (a) [http://lh6.ggpht.com/\\_5N\\_f6noTS2M/SNbhkUr6NQI/AAAAAAAAABnY/Z6EGbAiYAyw/Acacia%2BTree%2BOver%2Bthe%2BSavannah\\_Kenya.jpg](http://lh6.ggpht.com/_5N_f6noTS2M/SNbhkUr6NQI/AAAAAAAAABnY/Z6EGbAiYAyw/Acacia%2BTree%2BOver%2Bthe%2BSavannah_Kenya.jpg),

(b) [http://cdn2-www.dogtime.com/assets/uploads/gallery/goldendoodle-dog-breed-pictures/running-1\\_0.jpg](http://cdn2-www.dogtime.com/assets/uploads/gallery/goldendoodle-dog-breed-pictures/running-1_0.jpg),

(c) <http://hdwallpaperbackgrounds.net/3d-wallpapers-hd-nature/>,

(d) <http://www.motherpedia.com.au/article/holiday-at-home-phillip-island> )

A fenti ábrán a baloldali képek az eredetiek, míg a jobboldali képek az általunk ajánlott optimalizált változataik. Mindegyik kép alatt fel van tüntetve az általunk kiszámított esztétikai érték is. Látható, hogy egyes képek esetén jelentős optimalizációt sikerült elérnünk.

Az általunk használt algoritmus csak a szabályok betartását ellenőrzi, nincs beépítve semmilyen szintű kreativitás. Ezért mindenki eldöntheti, hogy a kapott képek valóban esztétikusabbak-e. Az azonban biztos, hogy a vizsgált szabályokat jobban betartják.

## 6. Továbbfejlesztési lehetőségek

Dolgozatunk jelen állapotában egy olyan algoritmust ír le, mely képes egy bemeneti képen azonosítani a kiemelkedő területeket és vonalakat, ezek felhasználásával esztétikai értéket számolni, illetve az adott képnek egy olyan részképét eredményül adni, mely optimális értékkel rendelkezik. Segítséget nyújt tehát az amatőr fényképészeknek a sikeresebb, esztétikusabb képek készítésében. Az eredményesebb felhasználás érdekében érdemes néhány továbbfejlesztési lehetőséget figyelembe venni.

Annak ellenére, hogy az algoritmus jelen állapotában sem nevezhető lassúnak, érdemes volna a számítások gyorsítására törekedni. A dolgozat elsődleges továbbfejlesztési lehetőségei közé sorolható tehát a performancia javítása. Mivel az algoritmus elég időigényes képfeldolgozási számításokat igényel, illetve az esztétikai értéket is több ezerszer számítjuk ki a genetikus algoritmusban, ezeket jó volna a lehető legrövidebb idő alatt elvégezni. Az objektumok, illetve vonalak azonosítását külön szálon futtatva, illetve az esztétikai értékszámolásban is párhuzamosítást alkalmazva csökkenthető volna az algoritmus futási ideje.

Szintén az algoritmus gyorsítását szolgálná a genetikus algoritmus továbbfejlesztése. Érdeemes volna tehát különböző típusú genetikus algoritmusokat kipróbálni, közben figyelni arra, hogy milyen típusú genetikus operátorok esetén kapunk gyorsabb, vagy adott esetben jobb eredményt.

Egy másik továbbfejlesztési lehetőségként megemlíthető egyéb kompozíciós szabályok bevezetése. Minél több szabályt sikerül beépíteni, annál pontosabban lehetne esztétikai értéket számolni. Ehhez azonban olyan szabályokat kell találnunk, melyek jól együttműködnek a már felhasználtakkal. A megfelelő szabályok megtalálása kihívást jelenthet. Érdeemes volna tehát olyan szabályok és mérési metrikáik után nézni, melyek növelnék a kép esztétikai megítélésének helyességét.



## 6. Könyvészet

[1] Florian Hoenig. Defining computational aesthetics. *Comp. Aesthetics in Graphics, Vis. and Imaging* (2005): 13-18. (<https://pdfs.semanticscholar.org/53ff/fdf60dc8d2a70d8e731c21a5ddb042b959ee.pdf>).

Megtekintve: 2017.03.18.

[2] Barry O'Carroll. 20 Composition Techniques That Will Improve Your Photos. (<https://petapixel.com/2016/09/14/20-composition-techniques-will-improve-photos/>)  
.Megtekintve: 2017.03.10.

[3] David Peterson. 18 Composition Rules For Photos That Shine. (<http://www.digital-photo-secrets.com/tip/3372/18-composition-rules-for-photos-that-shine/>).

Megtekintve: 2017.03.10.

[4] Darren Rowse. 5 Elements of Composition in Photography. (<https://digital-photography-school.com/5-elements-of-composition-in-photography/>).

Megtekintve: 2017.03.11.

[5] Dékán István. Mélységélesség – Képrajzolás és gyújtótávolság, a perspektíva hatás alakítása. (<http://www.fotovilag.hu/foto/objek-1/melysg.htm>).

Megtekintve: 2017.04.22.

[6] Ligang Liu, et al. Optimizing photo composition. *Computer Graphics Forum*. Vol. 29. No. 2. Blackwell Publishing Ltd, 2010 (<https://contrastly.com/wp-content/uploads/Composition-TR-low.pdf>).

Megtekintve: 2017.03.10.

[7] OpenCV Hivatalos Weboldal. <http://docs.opencv.org/3.2.0>.

Megtekintve: 2017.04.18.

[8] Image Thresholding. [http://docs.opencv.org/trunk/d7/d4d/tutorial\\_py\\_thresholding.html](http://docs.opencv.org/trunk/d7/d4d/tutorial_py_thresholding.html).  
Megtekintve: 2017.03.24.

[9] Canny Edge Detection. [http://docs.opencv.org/3.1.0/da/d22/tutorial\\_py\\_canny.html](http://docs.opencv.org/3.1.0/da/d22/tutorial_py_canny.html).  
Megtekintve: 2017.03.25.

[10] Image Filtering. <http://docs.opencv.org/3.0-beta/modules/imgproc/doc/filtering.html>.  
Megtekintve: 2017.03.25.

[11] Wei Luo, Xiaogang Wang, Xiaoou Tang. Content-based photo quality assessment. 2011 IEEE International Conference on Computer Vision (ICCV), 2011 (<https://pdfs.semanticscholar.org/627a/b487cbf0bf0062b318da37b479d571957ff7.pdf>).

Megtekintve: 2017.04.12

[12] Zbigniew Michalewicz. Genetic algorithms + data structures = evolution programs. Springer, 1992.

Megtekintve: 2017.04.13.

[13] Gary Bradski, Adrian Kaehler. Learning OpenCV. O'Reilly, 2008.

Megtekintve: 2017.03.28.

[14] Genetic algorithm, [https://en.wikipedia.org/wiki/Genetic\\_algorithm](https://en.wikipedia.org/wiki/Genetic_algorithm),

Megtekintve: 2017.04.20.

[15] John Canny. A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence 6 (1986): 679-698. (<https://pdfs.semanticscholar.org/55e6/6333402df1a75664260501522800cf3d26b9.pdf>).

Megtekintve : 2017.04.01.