1. **Bevezető**

A művészek a történelem során mindig is arra törekedtek, hogy esztétikát csempészszenek 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 öszpontosít, 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 számítástehnikai eljárások összessége, melyek az emberi szemhez hasonlóan képesek jól alkalmazható esztétikai döntéseket hozni.

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 látványosabb képeket lehet készíteni. Azonban az átlagember nem rendelkezik különösebb fotós tapasztalatokkal, 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 hasznos segítséget nyújtani a felhasználónak. Itt jön képbe a számítógépes esztétika használata.

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

1. **Kompozíciós szabályok**

Kompozíció alatt a fényképen található elemek pozícióját, ezek elrendezésének egyensúlyát értjük. E mellett magába foglalja a képek színezetét és textúráját is. 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 a kreativitásban, egyediségben rejlik, éppen ezért nem kell a kompozíciós szabályokat sem köbevésett törvényként kezelni. Éppen ezért sokkal inkább nevezhetőek tanácsoknak, melyeket jó megfogadni, ha egy igazán sikeres és maradandó képet szeretnénk alkotni. Előfordulhat az is, hogy két szabály ellentmond egymásnak, ezért választani kell, melyik teszi előnyösebbé a képet.

A legalapvetőbb és legszéleskörübben ismert a harmadolás szabály. A leggyakoribb hiba az amatőr fotózásban az, hogy minden elem a kép közepére kerül. Azonban 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. Előnyős például egy emberi alakot vagy fát a kép jobb vagy baloldali harmadába helyezni, illetve a horizont vonalat az alsó vagy felső harmadvonal mentére pozicionálni. A legtöbb kamerában van már lehetőség harmadolt nézetben készíteni képeket, segítséget nyújtva a fénykép készítőjének.

Az előző szabállyal ellentmondóan néha igenis előnyös, ha a célpont a kép közepére kerül. A szimmetrikus összeállításokat érdemes a kép közepére helyezni. 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. Ez egyedivé teheti a képet.

Az egyensúly fogalma is központi szerepet játszik a kép-összeállítás harmóniájában. Egy kiegyensúlyozott képen az objektumok egyenló távolságra vannak elosztva a kép középpontja körül. Egy nem megfelelőképpen kiegyensúlyozott kép zavart okozhat, túlzsúfolt hatást kelthet.

A fényképek két dimenzióban örökítik meg a minket körülvevő világot. Bizonyos tehnikákkal azonban három dimenziós 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. Igy mélységérzetet lehet kelteni, illetve a figyelem központjába a közeli célobjektum kerül, nem pedig a színes, túlzsúfolt háttér.

A három dimenziós 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.

Az emberek természetes módon vonzódnak a mintákhoz és különböző textúrákhoz. Vonzóak a szem számára és harmóniát sugallnak. Az ismétlődő minták mindig jó módszert biztosítanak a kellemes képek készítésére.

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 tartalmzó 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 teljes mértékben a célpontra összpontosítson, megfigyelje a részleteket. Ezzel ellentétben az is nagyon látványos lehet, ha sok semleges helyet hagyunk a fő objektum körül. 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.

1. **Beépített szabályok alkalmazása**

Ahhoz, hogy egy képnek számszerűsíteni lehessen az esztétikai értékét fontos az előbb említett szabályok közül néhányat kiválaszatni. Mivel egyes szabályok ellentmondanak egymásnak, olyan szabálycsoportot kell kiemelni 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 válsztottunk ki: a harmadolás szabályt, az átlók dominanciáját illetve a vizuális egyensúlyt. 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 tekintetet, így a néző nem érezheti semmitmondónak a képet. E mellett az emberi lélek igényli az egyensúlyt, így a képekben is hajlamos az egyensúlyt részesíteni előnyben. A rendezetlen, túlzsúfolt, majd lappangó területeket tartalmazó képek zavart okozhatnak, a néző elveszítheti a kép mondandóját. Ezért esett a válsztás, az előbb említett két szabály mellett, a vizuális egyensúly beépítésére is.

 

1. (b)

 

1. (d)
2. **Ábra:** A felhasznált kompozíciós szabályok példákkal

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 helyezhedik 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 bizonyultak. Kisebb változtatásokkal ezen metrikákat implementálva alakult ki az esztétikai értéket számszerűsítő alkalmazás.

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

A célobjektumok azonosításához előbb egy sor képfeldolgozó algoritmust kell alkalmaznunk, mely megfelelően összemossa a képet, ezzel eltüntetve a felesleges éleket, fekete fehérré kell alakítani, majd megfelelő thresholdot alkalmazva 0 és 1 közé redukálnunk a pixelek értékeit. Ezután egy Canny operátort alkalmazva azonosíthatjuk a képen található fő kontúrokat. Ha mindez megtörtént, a kontúrokat magukbafoglaló 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 végett az egymással átfedésben levő téglalapokat egybeolvasztjuk, így kevesebb darab, nagyobb területű célobjektumot kapva eredményül.

A kép túlzott összemosása 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 végzünk, hogy ezek mindenképp bekerüljenek a talált objektumok közé. A felfedezett arcokat nagyobb súllyal ellátva adjuk hozzá a főobjektumok listájához.

A vonalak azonosítása is képfeldolgozó algoritmusok alkalmazásával valósítható meg. Első körben Canny éldetektáló operátort használunk, úgy, hogy előtte nem mossuk különösebben össze a képet. A vonalak keresésekor fontos, hogy minél élesebb képpel dolgozzunk, ugyanis összemosás esetén értékes információt veszíthetünk, ezért előfordulhat, hogy bizonyos vonalak nem lesznek megtalálva. Ezután egy Hough – féle vonaldetektáló algoritmust alkalmazunk. Ha ezt megfelelően paraméterezzük, eredményül meg is kapjuk a képben 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 tartjuk meg. Egy vonal minél hosszabb, annál nagyobb a súlya.

Az általános vonalak mellett számon kell tartanunk az átló közeli vonalakat is. Az átló közeliséget nem a vonal pozíciójára értjük, hanem arra, hogy mennyire áll az átlókhoz hasonló szögben. Az átlók szögéhez viszonyítva 10o-os eltérést engedünk meg. Az átlós vonalak listáját az átlódominancia szabály vizsgálata miatt kell számontartanunk.

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 : , illetve 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.

alatt az aktuális, i-edik objektum súlyát értjük, ami nem más, mint az objektum területe, míg az i-edik vonal súlya, azaz ennek a hossza. az összes átlagos vonalat jelöli, míg az összes átlós vonalat, azaz azon vonalakat melyek legfeljebb 10o-ban térnek el a és által jelölt átlók valamelyikétől.

|  |  |
| --- | --- |
| Szimbólum | Jelentése |
|  | a kép szélessége és hosszúsága |
|  | a kép középpontja |
|  | a négy súlypont |
|  | a négy harmadvonal |
|  | a detektált objektumok |
|  | az objektum középpontja és súlya |
|  | objektum méret – képméret arány |
|  | az átlók |
|  | a detektált vonalak |
|  | nem átlós vonalak halmaza |
|  | átlós vonalak halmaza |
|  | a vonal súlya |
|  | normalizált Manhattan távolság |
|  | két vonal közötti távolság |

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

**A harmadolás szabály (Rule of Thirds, RT)** kiszámítása két fő komponensből tevődik össze : a célobjektumoknak a súlypontoktól mért távolságát méri, míg az a vonalaknak a harmadvonalaktól mért távolságát tükrözi. A képletben szereplő és súlyok. Az összképlet ezen komponensekből a következő képpen épül fel:

A pont értéke az összes célobjektumnak :

ahol az aktuális célobjektumnak a négy, súlyponthoz mért minimális távolságát jelenti, , illetve konstans egyenlő 0.17 .

A vonal értéke az összes megtalált, nem átlós vonalnak :

ahol a minimum vonaltávolság az aktuális vonaltól a négy harmadvonalhoz mérve, illetve a 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 , míg a -nak választottuk meg.

**A vizuális egyensúly (Visual Balance, VB) –**ról akor beszélhetünk, ha a célobjektumok összességének középpontja közel található a képnek a középpontjához. Ennek megfelelően a vizuális egyensúly értékének kiszámítása :

ahol .

**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 :

ahol .

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

ahol , és súlyok, és rendre 1, 0.3, illetve 1 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/ok 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: , és . Az áltlunk használt méret-függvény az ehhez konvergáló arányokat támogatja :

ahol , és konstansok ezen arányok előfordulását reprezentálják.

**Az összetett esztétikai érték (Combined Aesthetic Score)**  a kiszámított és értékekből tevődik össze :

ahol , ugynis a hangsúly az esztétikai értéken van, a célobjektum mérete csak másodlagos, a határozatlan esetek elkerülése végett lett beépítve a rendszerbe.

1. **Az alkalmazás felépítése, megvalósítás lépései**

**4.1. Felhasznált technológiák**

A képelemző 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. Alkalmazásunk sok számítást kell végezzen, ezért szükségeltetett a performanciát szem előtt tartani.

Alkalmazásunk 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 kell mosnunk a képet, ezáltal a felesleges élektől megszabadúlni, míg a fontosakat hangsúlyosabbá tenni. Erre a célra a piramis átlagoló szűrőt alkalmaztuk. Az algoritmus egy *maxlevel+1* szintet tartalmazó Gauss piramist épít, az összemosó algoritmust mindig a kisebb szinteken alkalmazva először, majd az eredményeket átadva a magasabb szinteknek. Itt újból alkalmazva van az összemosás, 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 adaptive threshold-ot 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 bizonyult. Ez egy sor algoritmusnak az alkalmazása, mely magába foglalja a kép összemosását, a zavarok kiszűrését, konvolúciós maszk alkalmazását, a nem éleket képező pixelek kiszűrését, illetve utolsó lépésként, két küszöbérték alkalmazását : ha az alsónál kisebbek 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 összemosást és a Canny – féle algoritmusban is található egy. Azért volt szükséges egy összemosást alkalmazni előre, mivel a Canny algoritmusban nincs megadva a választás lehetősége: Gauss – féle összemosást használ, nekünk azonban a piramis átlgoló a megfelelőbb. Eredményképp a kimeneti kép már csak a kontűröket tartalmazza.

Ezután már csak külön kell választani ezen kontűröket 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. Az így kapott nagyobb területű téglalapok már a célobjektumokat reprezentálják.

A kép színkombinációjából adóthat 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 minden arc bekerült a célobjektumok közé.

**4.2.2. Vonalak azonosítása**

1. Esettanulmány
2. Következtetések
3. Továbbfejlesztési lehetőségek
4. Könyvészet

[1] HOENIG F. : Defining Computational Aesthetics, 2005

[2] <https://petapixel.com/2016/09/14/20-composition-techniques-will-improve-photos/>

[3] <http://www.digital-photo-secrets.com/tip/3372/18-composition-rules-for-photos-that-shine/>

[4] <https://digital-photography-school.com/easy-tips-to-help-beginners-understand-composition/>

[5] <https://digital-photography-school.com/5-elements-of-composition-in-photography/>

[6] RENJIE L. L., WOLF C. L., COHEN-OR D. : Optimizing photo composition, Technical Report, Jan. 2005

[7] http://docs.opencv.org/3.2.0