



EÖTVÖS LORÁND TUDOMÁNYEGYETEM
INFORMATIKAI KAR
NUMERIKUS ANALÍZIS TANSZÉK

Képmínőséget mérő metrikák

Témavezető:

Dr. Fridli Sándor

egyetemi docens

Készítette:

Bognár Gergő

programtervező informatikus MSc

Budapest, 2014.

Tartalomjegyzék

1. Bevezetés	2
2. A képminőség mérése	4
2.1. Célkitűzés	5
2.2. Szubjektív mérés, adatbázisok	5
2.3. A metrikák csoportosítása	6
2.4. Alapvető módszerek	7
3. Torzulások modellezése	10
3.1. Zaj	10
3.2. Elmosódás, elmozdulás	12
3.3. Luminancia, kontraszt változás	12
3.4. Blokkosodás	14
4. Referenciaképet használó metrikák	15
4.1. A hiba mérésén alapuló metrikák	15
4.2. Strukturális hasonlóság	17
4.3. Természetes képek modellezése	21
5. Referenciakép nélküli metrikák	25
5.1. Típusok	26
5.2. Q metrika	27
6. Összehasonlító elemzés	30
7. Implementáció és a szemléltető alkalmazás	34
7.1. A szemléltető alkalmazás dokumentációja	34
7.2. Metrika függvények dokumentációja	38
Hivatkozások	41
Függelék: Rövidítések	43

1. Bevezetés

Az objektív képminőség mérés célkitűzése olyan automatikus mérési módszerek kialakítása, melyek a vizuális tartalom minőségét az emberi érzékelésnek megfelelő módon határozzák meg. A mérés eszközei a képminőséget mérő metrikák, amelyek a vizsgált kép minőségét egy objektív pontszámon keresztül hivatottak jellemezni.

A képminőség mérésének kiemelt szerepe van a digitális kép- és videófeldolgozásban. A minőségvizsgálat felhasználása széleskörű, fontos részét képezi a kép- és videófeldolgozó algoritmusoknak és rendszereknek, megjelenik azok implementációja, optimalizációja és tesztelése folyamán. A felhasználási területei közé tartoznak a képrögzítés, átvitel, tárolás, tömörítés, képjavítás, rekonstrukció, képernyőn és nyomtatásban történő megjelenítés algoritmusai és módszerei. Például a képrögzítés során a rögzítőeszköz beállításainak optimalizációja, a képtömörítés esetén az eredeti és a tömörített kép közötti eltérés vizsgálatán keresztül a tömöríthetőség mértékének becslése, az iteratív rekonstrukciós és képjavító algoritmusok célfüggvényének meghatározása egyaránt képminőséget mérő metrikák segítségével történhet.

A képminőség mérése jelentős szakirodalommal rendelkező, napjainkban is aktívan kutatott terület. Kialakulásához az emberi látás és a digitális képképzés közötti különbségekből származó problémák vezettek. Az emberi látórendszer sajátosságaiól fakadóan egy képen nem minden változtatás érzékelhető, nem minden régió kap ugyanakkora figyelmet, egy kép ugyanolyan mértékű mérhető megváltozásainak nincs ugyanolyan érzékelhető hatása. Az objektív mérés nehézségét az emberi látórendszer összetettsége, részletes ismeretének hiánya adja. A gyakorlatban alkalmazott metrikák az emberi érzékelés pontos modellezése helyett sokszor a kép bizonyos minőségbeli jellemzőit vizsgálják, mely jellemzők a vizsgálni kívánt képek típusának figyelembe vételével kerülnek meghatározásra. Természetes képek esetén például gyakori az élesség, kontrasztosság, zajosság mértékének vizsgálata.

A munkám során megvizsgáltam a képminőség mérésének alapvető feladatait és problémáit, a minőségmérésre szolgáló metrikákkal szemben támasztott alapvető elvárásokat. A metrikák általános jellemzése, csoportosítása és meghatározó módszereinek bemutatása mellett elemeztem a gyakorlatban leginkább elterjedt metrikákat, kitérve azok felhasználási területére, előnyeikre és hátrányaikra. Megvizsgáltam a képminőség romlásához vezető alapvető torzulás típusokat és azok modellezhetősé-

gét. A munkám eredménye egy elemző összehasonlítás, több meghatározó metrika implementálása, valamint egy MATLAB környezetben futó grafikus szemléltető alkalmazás, melynek segítségével különböző torzulások hatását lehet szimulálni, és a torzulások metrikákkal történő mérhetőségét vizsgálni.

2. A képminőség mérése

Az emberi látás alapvető funkciói közé tartozik a különbségtétel egy kép „jó” és „rossz” minőségű változatai között. Az 1. ábrán közepen a szürkeárnyalatos Lena tesztkép, mellette kétoldalt annak módosított változatai láthatók. Megfigyelhető, hogy a baloldali, elmosódott változatot jellemzően lényegesen más (rosszabb) minőségűnek érzékeljük, mint az eredeti képet. Ezzel szemben a jobboldali, kontrasztosított változat és az eredeti kép között nem érzékelünk lényeges minőségbeli különbséget.



1. ábra. Elmosódott, eredeti és kontrasztos Lena

A számítógépes feldolgozás szempontjából a képek közötti különbség másként jelenik meg. Két digitális kép közötti távolságot a legkézenfekvőbb a pixelenkénti eltérések négyzetének átlagaként definiálni (MSE – *mean-squared error*). Az ábrán szereplő két módosított változatot az eredeti képpel összehasonlítva az MSE ugyanazt az értéket adja. A példa tehát azt demonstrálja, hogy ez a metrika nem jellemzi jól a képek minőségbeli különbségét az emberi észlelés szempontjából, ez képezi az alapját a további vizsgálatoknak.

Az emberi látórendszernek ugyanakkor általában szüksége sincs referenciaképre egy kép minőségének meghatározásához. Az ábrán baloldalt látható, elmosódott változatot az eredeti kép ismerete nélkül is rossz minőségűnek, a jobboldali, kontrasztos változatot jó minőségűnek érzékeljük. A számítógépes feldolgozás során is felléphet olyan szituáció, amikor nem áll rendelkezésre referenciaként a vizsgált kép egy jó minőségű változata.

2.1. Célkitűzés

A képminőség mérésének általános célkitűzése tehát – a bevezetőben leírtaknak megfelelően – olyan mérési módszerek, képminőséget mérő *metrikák* kialakítása, melyek a képek minőségét az emberi érzékelésnek megfelelő módon határozzák meg.

Az általános célkitűzés mellett a metrika megkonstruálása során figyelembe kell venni a metrika által vizsgálni kívánt képek típusát, jellemzőit, a minőségbeli romlásukhoz vezető okokat. A képek gyakorlatban előforduló torzulásai néhány alapvető torzulás típusra vezethetők vissza: a képrögzítés során a képre rakódó zaj, elmosódás és elmozdulás, luminancia és kontraszt változás, és a képtárolás során megjelenő blokkosodás, élelmosódás. A metrikákkal szemben elvárás, hogy a célterület képeinek jellemző torzulásaira az emberi érzékeléssel összhangban reagáljanak.

2.2. Szubjektív mérés, adatbázisok

A legtermészetesebb módszer a képminőség mérésére az emberi látórendszer közvetlen bevonásával, tesztalanyok segítségével végzett szubjektív mérés. A mérés eredménye egy szubjektív pontszám (MOS – *mean opinion score*), ami egy átlagos szemlélő várható véleményének feleltethető meg.

A szubjektív mérés során a tesztalanyok többször ismételve megtekintik, és adott skálán értékelik a vizsgált képeket. A mérés eredményeképp egy adott képhez rendelt MOS pontszám az egyes szemlélők egyes pontszámainak az átlaga, ezáltal reprezentálva egy átlagos szemlélő véleményét. A módszer előnye, hogy a lehető legjobban határozza meg az érzékelhető képminőséget, mivel közvetlenül az emberi látórendszert használja a pontszámok meghatározásához. A hátrányai közé tartozik, hogy időigényes, körülményes és drága, és egy objektív metrikával szemben nem megismételhető az eredménye. A résztvevők és a mérési módszer kialakítása döntő fontosságú, ugyanis a mérési eredmény nagyban függ a résztvevők fizikai kondíciójától, érzelmi állapotától, személyes tapasztalaitól és a megelőző vizuális környezettől.

Több adatbázis is publikusan elérhető, ahol a tesztképek minőségét tesztalanyok segítségével határozták meg: LIVE [14], CSIQ [6], IVC [2], Toyama [5], A57 [1], TID [9] és WIQ [3]. Az adatbázisok jellemzően néhány (jó minőségű) referenciaképet, és azoknak többféle módszerrel torzított változatait tárolják. Az MOS pontszám helyett egyes adatbázisokban a tesztkép és a referenciakép pontszámának különbségeként

kapott DMOS (*differential mean opinion score*) pontszám található meg.

Az objektív képminőséget mérő metrikákkal szembeni elvárás az, hogy az általuk meghatározott pontszám segítségével minél hatékonyabban lehessen becsülni a várható szubjektív MOS pontszámot, ezért ezek az adatbázisok jó alapot nyújtanak egy metrika teszteléséhez és validálásához. A validálás során az adatbázis MOS illetve DMOS értékei kerülnek összehasonlításra a metrika által meghatározott értékekkel. Az összehasonlítás történhet egyszerűen az objektív és a szubjektív pontszámok közötti átlagos négyzetes eltérés négyzetgyöke (RMSE – *root-mean-squared error*) segítségével, vagy statisztikai vizsgálat, rangkorreláció használatával is. A gyakorlatban a Pearson-féle és a Spearman-féle korrelációs együtthatók használata a legelterjedtebb. Mindkét együttható egy 0 és 1 közötti számmal jellemzi az objektív és szubjektív pontszámok közötti kapcsolatot, ahol az 1-es érték jelenti a tökéletes egyezést. [7]

2.3. A metrikák csoportosítása

2.3.1. Modellezésen alapuló csoportosítás

A felhasznált módszerek szerint a metrikák két csoportba oszthatók: az emberi látórendszer (továbbiakban HVS – *human vision system*) modellezésén és a jelváltozás modellezésén alapuló metrikákra.

A HVS modellezésén alapuló eljárások az emberi látásról alkotott pszichológiai, fiziológiai és humánbiológiai ismeretek alapján, többek között dekompozíciós módszerek, kontrasztérzékeny szűrés és fényadaptáció segítségével. Ezen módszerek hátránya összetettségük és magas számításigényük.

A jelváltozás modellezésén alapuló metrikák bár figyelembe veszik a HVS tulajdonságait, de annak pontos modellezése helyett kép torzulásának mértékére próbálnak következtetni a kép minőségbeli jellemzőit vizsgálva. Módszerei többek között a strukturális hasonlóság vizsgálata, statisztikai modellezés, jellemző és torzulás detektálás. Ezen módszerek a gyakorlatban szélesebb körben alkalmazottak könnyebb számíthatóságuk miatt, hátrányuk azonban, hogy eredményük adott esetekben eltérhet az ember érzékelt minőségtől.

2.3.2. Referenciakép szerinti csoportosítás

A referenciaképet használó (FR – *full-reference*, és RR – *reduced-reference*) metrikák akkor alkalmazhatók, ha a kiértékelni kívánt kép mellett rendelkezésre áll annak jó minőségű, eredeti, torzításoktól mentes változata is. Ezen metrikák a vizsgált és a referenciakép közötti eltéréseken keresztül határozzák meg a vizsgált kép minőségét. Az FR és RR metrikák közötti különbség az, hogy míg az FR metrikák a teljes referenciaképet felhasználják, az RR metrikák pedig annak csak egy részét: a referenciakép dekompozíciójaként kapott adatokat vagy a referenciakép bizonyos jellemzőit. A referenciaképet használó metrikák népszerűek például egyes képfeldolgozó algoritmusok esetén, a bemeneti és kimeneti kép közötti eltérés jellemzésére.

A referenciakép nélküli (NR – *no-reference*, vagy *blind*) metrikák csak a vizsgált képet használják fel annak kiértékeléséhez. Ez a megközelítés közelebb áll az emberi érzékeléshez, de a kiértékelésükhöz általában valamilyen *a priori* ismeret szükséges a képről. A referenciakép nélküli metrikák használata népszerű például iteratív képjavító eljárások optimalizációs célfüggvényének meghatározásánál.

2.4. Alapvető módszerek

A metrikák jelentős részének alapját a következő módszerek, technikai fogás alkotják: dekompozíció, jellemző és torzulás detektálás, az éppen érzékelhető megváltozás és a vizuális figyelem modellezése. [7]

2.4.1. Dekompozíció

A dekompozíciós technikák használatának alapja az, hogy a HVS is különböző módon dolgozza fel az akromatikus és kromatikus jeleket, a gyors és a lassú mozgást, valamint a feldolgozás során különbséget tesz a jel frekvenciája és orientációja szerint. A dekompozíció szín vagy luminancia, frekvencia, és orientáció meghatározását foglalhatja magába. A dekompozíció szerinti csatornákra bontás segítségével lehetőség nyílik az egyes csatornákra vonatkozó különböző HVS reakció modellezésére. A szín szerinti dekompozíció színtértranszformációval történhet, gyakran alkalmazott színtérek az YC_bC_r és a $CIELAB$ színtér, amelyek az RGB színtérnél relevánsabb módon írják le a színeket az emberi érzékelés szempontjából. A minőségvizsgálat

során azonban a luminancia vizsgálata az elterjedtebb, mivel ez fontosabb szerepet játszik az emberi észlelés során, mint a kromatikus komponensek. A frekvencia szerinti dekompozíció különböző matematika transzformációk segítségével valósítható meg, a leggyakoribbak a DFT (diszkrét Fourier-transzformált), DCT (diszkrét koszinusz-transzformált) és a DWT (diszkrét wavelet-transzformált).

2.4.2. Jellemző és torzulás detektálás

A digitális képfeldolgozás során a jellemző (*feature*) a képből kinyert információkat jelöli. A jellemző detektálás gyakori eleme a képfeldolgozó algoritmusoknak, feladata a vizsgált kép pontjaiban vagy régió meghatározni a jellemző értékét, jellemzően a pont vagy régió egy lokális környezetét vizsgálva. Gyakran keresett jellemzők például az élek illetve a sarkok. A képminőség mérése során a leggyakrabban vizsgált jellemző a kontraszt, mivel a HVS jobban érzékeli egy jel kontrasztosságát, annak intenzitásával szemben.

A képrögzítés és feldolgozás során a kép különböző típusú torzulásokat szenvedhet, ezek detektálása, mérése fontos részét képezi a metrikák működésének, a bevezetőben és a célkitűzés részben leírtaknak megfelelően.

2.4.3. JND modellezés

A HVS működéséből fakadóan egy képen nem minden változás érzékelhető. Az éppen észrevehető megváltozás (JND – *just-noticeable distortion, just-noticeable difference*) az a küszöb, ami alatt a szemlélők nagy része nem érzékeli a különbséget.

Egy általános, gyakran alkalmazott JND modell a *Weber-törvény*. A törvény a HVS fiziológiai vizsgálatára alapozva kimondja, hogy a HVS egy jel intenzitásának relatív megváltozására érzékeny: az éppen érzékelhető ΔI intenzitásváltozás közelítőleg arányos az eredeti I intenzitással, vagyis $\Delta I / I \approx konst.$

Az összetett JND modellek lehetnek pixel- vagy dekompozíció alapúak is. Széles körben vizsgált a DCT alapú JND modelleket, mivel a DCT képezi az alapját több kép- és videótömörítő algoritmusnak.

2.4.4. VA modellezés

A HVS működésének másik jellemzője, hogy egy képen nem terület kap ugyanakkora figyelmet. Ez a jelenség mozgóképek vizsgálata esetén még relevánsabb, mint állóképek esetén. A vizuális figyelem (VA – *visual attention*) modell feladata a HVS ezen szelektivitásának modellezése. A VA modellek jellemzően összetett működésűek, dekompozíciós, jellemző detektáláson és JND modellezésen alapuló technikákat használva.

3. Torzulások modellezése

A fejezetben a képek minőségbeli romlását kiváltó leggyakoribb torzulás típusok és azok modellezése kerül bemutatásra. A képminőséget mérő metrikák általános feladata, hogy ezen torzulások hatását az emberi látórendszerhez hasonlóan érzékeljék. Munkám során a szürkeárnyaltos képek torzulásával foglalkoztam, illetve színes képek esetén a luminanciát érő torzulással.

3.1. Zaj

Egy jel torzulásának egyik leggyakoribb oka a ráakódó zaj, amely általában valamilyen véletlen hatásnak tekinthető és valószínűségi változóként modellezhető. Többféle zajtípus különböztethető meg, a jelfrögzítés, tárolás és átvitel módjától függ, hogy egy adott jel esetén milyen típus vagy típusok jelentkeznek.

A hagyományos digitális képalkotás során fellépő zaj forrásaként két fő összetevőt lehet megjelölni: az elektromos eszközök háttérzaját és a fény természetéből fakadó kvantumzajt. Különböző képalkotási technikák során más forrású és jellegű zaj is felléphet, ezért először egy rövid összefoglalás következik a zaj általános jellemzőiről.

3.1.1. A zaj általános jellemzői

Jelöljék az $x = (x_i)_{i=1}^N$ és $y = (y_i)_{i=1}^N$ sorozatok az eredeti és a torzult jel N számú diszkrét pontban mért értékeit. A zajt modellező valószínűségi változót jelölje a $\xi = (\xi_i)_{i=1}^N$ sorozat. A torzult jel adott értéke ekkor $y_i = F(x_i, \xi_i(x))$ alakban írható, ahol F a zaj megjelenési módját reprezentáló függvény.

A zaj megjelenésétől függően beszélhetünk *additív*, *multiplikatív*, és *só-bors* zajról. Az additív zaj olyan ξ valószínűségi változóként modellezhető, amely hozzáadódik az eredeti jelhez: $y_i = x_i + \xi_i$. A multiplikatív zaj pedig szorozódik: $y_i = \xi_i \cdot x_i$. Hagyományos képek rögzítése során a zaj általában additív formában jelentkezik, a multiplikatív zaj például radarképek esetén jellemző. A só-bors zaj, avagy impulzus zaj jellemzően jelfrögzítési és adatátviteli hibákból származik, melyek hatására a sérült érték helyére a maximális vagy a minimális mérhető érték kerül. Képek esetén ez fekete és fehér pontok formájában jelentkezik.

Az eredeti jel és a megjelenő zaj közötti kapcsolat alapján megkülönböztethető

jelfüggetlen és jelfüggő zaj. Jelfüggetlen zaj esetén a megjelenő zaj és az eredeti jel között nincs kapcsolat. A képrögzítés során az elektromos eszközök háttérzaja képfüggetlen zaj formájában jelentkezik, a fotonok természetéből adódó kvantumzaj pedig képfüggő zaj formájában.

Ha a zajt modellező ξ valószínűségi változó sorozat korrelálatlan, akkor *fehér* zajról, ellenkező esetben különböző *színes* zajokról beszélhetünk. Az elektromos és kvantumzajt általában fehér zajnak szokás tekinteni.

A fellépő zaj annak típusától függően különböző *eloszlású* valószínűségi változókkal modellezhető. Az elektromos eszközök zaja normális eloszlású (Gauss-zaj), a kvantumzaj Poisson-, a diszkretizálás hatására fellépő kvantálási zaj pedig egyenletes eloszlási valószínűségi változóval írható le.

3.1.2. Digitális képek zaja

A hagyományos digitális képalkotás során fellépő domináns zajok közül a digitális eszközök háttérzaja képfüggetlen, 0 várható értékű, additív fehér Gauss-zajként modellezhető, a fény természetéből fakadó kvantumzaj pedig képfüggő Poisson-zajként. A következőkben ezt a két hatást együttesen tartalmazó zajmodell kerül ismertetésre. [4] A modell elektromágneses szenzorokból származó felvételek leírására szolgál, itt a hagyományos digitális képalkotásra való alkalmazása következik.

A fenti jelöléseket használva a modell a két hatást két egymástól független, additív komponensként írja le:

$$y_i = x_i + \eta_p(x_i) + \eta_{g,i} \quad (1)$$

ahol η_p a képfüggő Poisson-zaj, melynek eloszlására $\chi > 0$ paraméter esetén:

$$\chi(x_i + \eta_p(x_i)) \sim \mathcal{P}(\chi x_i) \quad (2)$$

η_g pedig a képfüggetlen Gauss-zaj, 0 várható értékkel és $b > 0$ (konstans) szórásnégyzettel: $\eta_{g,i} \sim \mathcal{N}(0, b)$. A Poisson-eloszlás statisztikai tulajdonságai alapján a Poisson-zaj is 0 várható értékű, de képfüggő szórással:

$$E(\eta_p(x_i)) = 0, \quad D^2(\eta_p(x_i)) = \frac{x_i}{\chi} \quad (3)$$

Ennek következtében a torzult jel várható értéke az eredeti jel, szórása pedig $a = \chi^{-1}$ jelöléssel:

$$E(y_i) = x_i, \quad D^2(y_i) = ax_i + b \quad (4)$$

A centrális határeloszlás tétele alapján a Poisson-eloszlás „elég nagy” paraméter esetén normális eloszlással közelíthető: $\mathcal{P}(\lambda) \approx \mathcal{N}(\lambda, \lambda)$, ezért a két hatás együttesen is modellezhető egy 0 várható értékű, képfüggő szórású $\mathcal{N}(0, ax_i + b)$ eloszlású zajjal.

A modellben szereplő χ paraméter jellemzi a rögzítőeszköz érzékenységét, és így a kvantumzaj mértékét, a b paraméter pedig a Gauss-zaj szórásaként az elektromos háttérzaj mértékét.



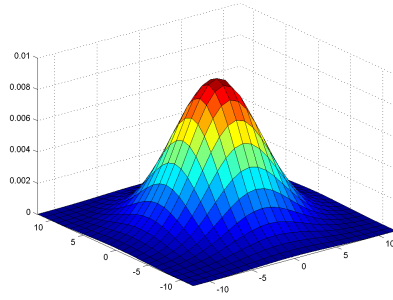
2. ábra. Zaj szimuláció a Barbara tesztképen: $a = 0.01$, $b = 0.04$

3.2. Elmosódás, elmozdulás

A digitális képek a képrögzítés és a képtárolás során is elmosódhatnak. A képrögzítés során az elmosódás forrása lehet a kamera vagy egy céltárgy elmozdulása és a nem megfelelő fókuszbéállítások. A képtárolás során az egyes tömörítő algoritmusok okozhatnak élességromlást és az élek torzulását. Az elmosódás mindkét esetben modellezhető egy átlagoló vagy egy Gauss-szűrő segítségével. Az 3. ábrán 25×25 méretű, 4 szórású Gauss-szűrő látható térbeli felületként ábrázolva. A szűrőnek a Lena tesztképre alkalmazott hatása az 1. ábrán látható.

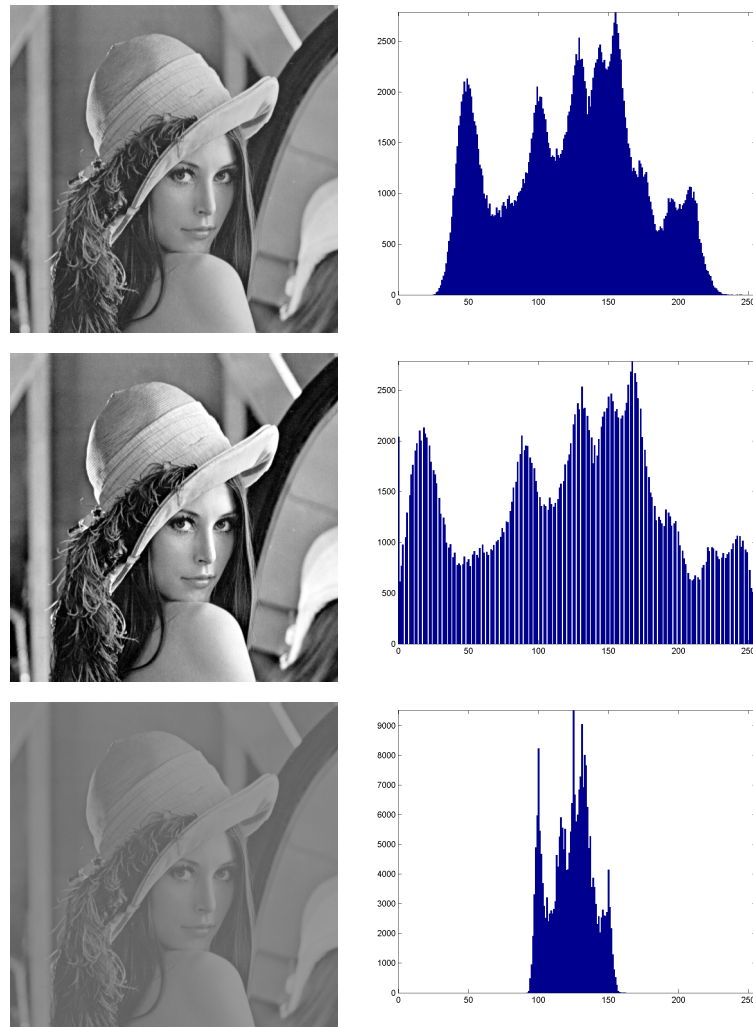
3.3. Luminancia, kontraszt változás

A digitális képek rögzítése során a kép minősége függ a környezetben jelenlévő megvilágítás erősségétől, és a rögzítőeszköz érzékelőinek beállításaitól. Ezen hatások a luminancia és a kontraszt megváltozásával írhatók le: a pixelek intenzitásértékének



3. ábra. 25×25 méretű, 4 szórású Gauss-szűrő

eltolódásával, illetve az intenzitások hisztogramjának nyújtásával és összenyomásával.



4. ábra. Kontraszt javított és rontott Lena tesztkép hisztogramokkal

Megjegyzendő, hogy a luminancia és a kontraszt változtatása nem minden eset-

ben okoz érzékelhető képminőség romlást. Esetenként pont ellenkezőleg: a HVS egy kép kontrasztosabb változatát jobb minőségűnek érzékelheti, mint az eredeti változatot. Ez a jelenség az alapja annak, hogy egyes egyszerű képjavító eljárások pont a luminancia és a kép hisztogramjának manipulálásán alapulnak. A luminancia vagy a kontraszt nagymértékű megváltozása azonban információvesztést és így torzulást okoz: a változtatás hatására a kép dinamikus tartományán kívül eső intenzitásértékek levágásra kerülnek.

A képminőséget mérő metrikák különbözőképpen értékelik luminancia és kontraszt megváltozását, a VIF metrika például az emberi érzékeléshez hasonlóan kifejezetten jobb minőségűnek értékelheti a kép kontrasztosított változatát az eredetivel szemben, az 6. ábrán látható módon.

3.4. Blokkosodás

A blokkosodás a képtárolás során létrejövő mesterséges torzulás, a képminőség mérése során mégis sokat vizsgált jelenség, mivel napjaink egyik legelterjedtebb képtároló módszere, a JPEG tömörítés is blokkosodást vált ki. A JPEG tömörítés a képet 8×8 méretű, nem átfedő blokkokra bontva tömöríti. Mivel az egyes blokkok feldolgozása egymástól független, így a blokkokban lokális hatások dominálnak, az 5. ábrán látható módon.



5. ábra. JPEG tömörített Lena

4. Referenciaképet használó metrikák

A fejezetben a referenciaképet használó FR metrikák főbb típusai, azok jellemzése és a gyakorlatban használt legfontosabb metrikák leírása található. Az FR metrikák 3 fő csoportba oszthatók a típusuk szerint: a hiba mérésén, strukturális hasonlóságon, és a természetes képek modellezésén alapuló metrikákra.

A munkám során a szürkeárnyaltos állóképeket vizsgáló metrikákra koncentráltam. Ezekben a képeken a szürkeárnyalat intenzitása a luminancia értéknek felel meg, amely fontosabb szerepet játszik az emberi észlelésben, mint a színek. A metrikák nagyobb része ezért szürkeárnyaltos képekkel foglalkozik. Természetesen azonban színes képek is vizsgálhatók ezekkel a módszerekkel, a színeket luminancia értékre transzformálva.

4.1. A hiba mérésén alapuló metrikák

A kiértékelendő kép tekinthető úgy, mint a referenciakép és a hiba összege. Az FR metrikák alapvető megközelítése ennek a hibának a mérésén, számszerűsítésén alapul. Az egyik legegyszerűbb módszerek az átlagos négyzetes eltérés (MSE – *mean-squared error*), és az ezen alapuló jel-zaj arány (SNR – *signal-to-noise ratio*, és PSNR – *peak signal-to-noise ratio*).

4.1.1. MSE, SNR, PSNR

Jelöljék az $x = (x_i)_{i=1}^N$ és $y = (y_i)_{i=1}^N$ sorozatok a referencia- és a kiértékelendő kép pixeleit, N a pixelek számát. Ekkor a hiba a két kép különbsége: $(y_i - x_i)_{i=1}^N$

Az MSE közvetlenül a hiba nagyságát méri négyzetes értelemben, átlagolva:

$$\text{MSE}(x, y) = \frac{1}{N} \sum_{i=1}^N (y_i - x_i)^2 \quad (5)$$

Az SNR, a jel-zaj arány a referenciakép nagyságához viszonyítja a hibát, általában logaritmikus skálát alkalmazva:

$$\text{SNR}(x, y) = 10 \cdot \log_{10} \left(\frac{\frac{1}{N} \sum_{i=1}^N x_i^2}{\text{MSE}(x, y)} \right) \quad (6)$$

A PSNR, a jel-zaj arány másik változata esetén a referenciakép nagyságát annak maximuma jelenti:

$$\text{PSNR}(x, y) = 10 \cdot \log_{10} \left(\frac{\max(x)^2}{\text{MSE}(x, y)} \right) \quad (7)$$

Az egyszerűség érdekében a tényleges maximumot legtöbbször a képek dinamikus tartománya (L) helyettesíti, amely szürkeárnyalatos képek esetén általában 255:

$$\text{PSNR}(x, y) = 10 \cdot \log_{10} \left(\frac{L^2}{\text{MSE}(x, y)} \right) \quad (8)$$

4.1.2. Tulajdonságok, előnyök és hátrányok, alkalmazhatóság

Az MSE nemnegatív valós érték, a vizsgált és a referenciakép megegyezése esetén 0, az eltérés növekedésével nő. Az SNR és PSNR tetszőleges valós értéket felvehet, a vizsgált és a referenciakép megegyezése esetén értékük $+\infty$, az eltérés növekedésével csökkennek.

A bemutatott három metrika a referenciakép minden megváltozását érzékeli, így reagálnak a leggyakoribb torzítás típusokra is. A legfőbb hátrányuk azonban az, hogy ez a reakció sok esetben nincs összhangban az ember számára érzékelhető megváltozással. Az 1. ábrához kapcsolódó példa is szemlélteti, hogy ugyanazon referenciakép két különböző módon torzított, de megegyező MSE (és így SNR és PSNR) értékkel rendelkező változatát lényegesen eltérő minőségűnek érzékeljük. Egyszerűségük miatt azonban gyakran alkalmazzák őket, ha csak a képminőség durva becslése szükséges.

4.1.3. Hibamérés általánosan

Általánosan a hiba mérésén alapuló rendszerek a következő fázisokra bonthatók: előfeldolgozás, kontrasztérzékeny szűrés, csatornákra bontás, csatornák hibájának mérése és normalizálása, és a hibák összegzése. [17]

Az *előfeldolgozás* feladata egyszerű átalakítások végrehajtása, néhány ismert torzulás kiküszöbölésére: színtérkonverzió a HVS működéséhez közelebb álló színtérre vagy luminanciára, a luminancia értékek transzformációja, szűrés és fényadaptáció. Az esetleges *kontrasztérzékeny szűrő* a HVS frekvenciaérzékenységének megfelelően súlyozza és szűri a képet. A *csatornákra bontás* történhet összetett módszerek segítségével, amelyek a látókéreg neuronválaszainak modellezésén alapulnak, vagy egysze-

rőbb módszerek, például diszkrét koszinusz-transzformált és wavelet-transzformált alapján. A *csatornák hibájának mérése és normalizálása* után az egyes hibaegyütthetők összegzése adja a metrika eredményét. Az összegzés leggyakrabban a Minkowski norma segítségével történik:

$$E(\{e_{l,k}\}) = \left(\sum_l \sum_k e_{l,k}^\beta \right)^{1/\beta} \quad (9)$$

ahol $e_{l,k}$ az l . csatorna k . együtthetője, β egy jellemzően 1 és 4 közé eső hatványkitevő.

4.2. Strukturális hasonlóság

A strukturális hasonlóságon alapuló metrikák, az SSIM (*structural similarity*) [17] és az SSIM-et megelőző, annak speciális esetét képező UQI (*universal quality index*) [18], a leginkább elterjedt, legszélesebb körben alkalmazott módszerek. A strukturális hasonlóság vizsgálatának alapgondolata az az észrevétel, hogy a természetes képek strukturáltak, és hogy az emberi látórendszer is strukturális információk kinyerésén keresztül értelmezi a képet, ezért a strukturális eltérések mérése jól becsülheti az érzékelhető képminőséget.

A strukturális hasonlóság mérése három összehasonlításra különíthető el: luminancia, kontraszt és struktúra. A luminancia összehasonlítása az átlagos intenzitáson, a kontraszt a szóráson, a struktúra pedig a korreláción keresztül történik. A három összehasonlítás eredményének kombinációja adja a metrika értékét.

4.2.1. UQI

Jelölje továbbra is $x = (x_i)_{i=1}^N$ és $y = (y_i)_{i=1}^N$ a referencia- és a vizsgált képet, N a pixelek számát.

Az átlagos intenzitásértékek a luminancia összehasonlításhoz:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i, \quad \mu_y = \frac{1}{N} \sum_{i=1}^N y_i \quad (10)$$

Az $l(x, y)$ luminancia összehasonlító függvény a μ_x és μ_y intenzitások függvénye:

$$l(x, y) = \frac{2\mu_x\mu_y}{\mu_x^2 + \mu_y^2} \quad (11)$$

Az l függvény összhangban van a HVS fény adaptációjával, a Weber-törvénynek megfelelően. Ha R jelöli a az emberi érzékelés szempontjából releváns háttérhez képest vett relatív luminancia változást a torzult képen, vagyis $\mu_y = (1 + R) \cdot \mu_x$, akkor az l függvény kifejezhető csak R -től függő alakban:

$$\frac{2(1 + R)}{1 + (1 + R)^2} \quad (12)$$

A szórás értékek a kontraszt összehasonlításához:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2}, \quad \sigma_y = \left(\frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_y)^2 \right)^{1/2} \quad (13)$$

A $c(x, y)$ kontraszt összehasonlító függvény a σ_x és σ_y szórások függvénye:

$$c(x, y) = \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2} \quad (14)$$

A c függvény konzisztens a HVS kontraszt-feldolgozásával: ugyanolyan mértékű $\Delta\sigma = \sigma_y - \sigma_x$ kontrasztváltozásra kevésbé érzékenyen reagál magas kontraszt, mint alacsony kontraszt mellett, σ_x -hez viszonyítva.

Kovariancia a struktúra összehasonlításához:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y) \quad (15)$$

Az $s(x, y)$ struktúra összehasonlító függvény a két jel közötti korreláció¹:

$$s(x, y) = \frac{\sigma_{xy}}{\sigma_x\sigma_y} \quad (16)$$

A metrika a három komponens kombinációjaként áll elő:

$$Q(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y) = \frac{2\mu_x\mu_y}{\mu_x^2 + \mu_y^2} \cdot \frac{2\sigma_{xy}}{\sigma_x^2 + \sigma_y^2} \quad (17)$$

¹Megjegyzendő, hogy a $c(x, y)$ kontraszt összehasonlító függvény független a luminanciától, az $s(x, y)$ struktúra összehasonlító függvény pedig a luminanciától és a kontraszttól is. A függvényre ugyanis tekinthetünk úgy, mint a normalizált (0 átlagú és 1 szórású) $\tilde{x} = \frac{x - \mu_x}{\sigma_x}$ és $\tilde{y} = \frac{y - \mu_y}{\sigma_y}$ jelek közötti korrelációs együtthatóra, mivel:

$$s(x, y) = \frac{1}{N-1} \sum_{i=1}^N \frac{x - \mu_x}{\sigma_x} \cdot \frac{y - \mu_y}{\sigma_y} = \frac{1}{N-1} \sum_{i=1}^N \tilde{x} \cdot \tilde{y} = \frac{1}{N-1} \sum_{i=1}^N \frac{\tilde{x} - 0}{1} \cdot \frac{\tilde{y} - 0}{1}$$

A luminancia, a kontraszt és a képet érő torzulások a képtérben különböző mértékek lehetnek, ezért a kiértékelés lokálisan $B \times B$ méretű csúszóablakok segítségével történik. A képhez rendelt végső érték az egyes ablakok értékeinek átlaga lesz. M ablak esetén, az egyes ablakokban számított értékeket $(Q_j)_{j=1}^M$ -vel jelölve:

$$\text{UQI} = \frac{1}{M} \sum_{j=1}^M Q_j \quad (18)$$

(Az eredeti implementáció $B = 8$ méretű ablakot használ.)

4.2.2. UQI tulajdonságai

A metrika szimmetrikus, dinamikus tartománya $[-1, 1]$, a vizsgált és a referencia-kép megegyezése esetén értéke 1. Az alkalmazott képletek miatt az UQI numerikus instabil, ha egyes ablakok átlagintenzitása vagy szórása 0 közeli érték. Az alkalmazott négyzetes csúszóablak hatására a lokálisan számított Q_j értékek gyakran blokkosodási torzulást szenvednek. Az UQI továbbfejlesztésének tekinthető SSIM metrika figyelembe veszi ezeket a problémákat.

4.2.3. SSIM

A blokkosodás kiküszöbölésére az SSIM $w = (w_i)_{i=1}^N$, $\sum_{i=1}^N w_i = 1$ súlyokat használ a vizsgált ablakban az értékek kiszámolására:

$$\mu_x = \sum_{i=1}^N w_i x_i, \quad \mu_y = \sum_{i=1}^N w_i y_i \quad (19)$$

$$\sigma_x = \left(\sum_{i=1}^N w_i (x_i - \mu_x)^2 \right)^{1/2}, \quad \sigma_y = \left(\sum_{i=1}^N w_i (y_i - \mu_y)^2 \right)^{1/2} \quad (20)$$

$$\sigma_{xy} = \sum_{i=1}^N w_i (x_i - \mu_x) (y_i - \mu_y) \quad (21)$$

(Az eredeti implementáció 11×11 -es méretű ablakokban 1.5 szórású Gauss-szűrőnek megfelelő súlyokat használ.)

A numerikus instabilitás kiküszöbölésére pedig bevezeti az elég kicsi C_1 , C_2 és C_3 értékeket, és az összehasonlító függvények ezek segítségével módosulnak:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (22)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (23)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (24)$$

A három érték meghatározása rögzített $K_1, K_2 \ll 1$ konstansok segítségével, a vizsgált képek L dinamikus tartományának figyelembe vételével történik:

$$C_1 = (K_1L)^2, \quad C_2 = (K_2L)^2, \quad C_3 = \frac{C_2}{2} \quad (25)$$

(Az eredeti implementáció $K_1 = 0.01$, $K_2 = 0.03$ értékeket használ.)

Az SSIM a fenti két probléma kezelése mellett lehetővé teszi az egyes komponensek relatív fontosságának figyelembe vételét a komponensek kombinációjakor:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (26)$$

ahol a $\alpha, \beta, \gamma > 0$ paraméterek segítségével állítható az adott komponens fontossága, eredetileg $\alpha = \beta = \gamma = 1$.

A teljes képre történő kiértékelés itt is az egyes ablakokban felvett $(SSIM_j)_{j=1}^M$ értékek átlaga:

$$MSSIM = \frac{1}{M} \sum_{j=1}^M SSIM_j \quad (27)$$

4.2.4. Tulajdonságok, előnyök és hátrányok, alkalmazhatóság

Az SSIM örökli az UQI tulajdonságait, az UQI pedig megközelítőleg az SSIM speciális esetének tekinthető a $w_i = \frac{1}{N}$ ($i = 1, 2, \dots, N$) és $K_1 = K_2 = 0$ paraméterekkel.

Az SSIM, bár konzisztens a HVS fény- és kontrasztérzékelésével, nem egy összetett HVS modellre épül. A gyakorlati, többek között a szubjektív adatbázison alapuló tesztek azonban azt mutatják, hogy természetes képek esetén az SSIM jól közelíti az emberi érzékelésnek megfelelő értékelést, a gyakorlatban előforduló torzítás típusokra a HVS-sel összhangban reagál. Hatékonysága és matematikai egyszerűsége miatt ez az egyik legelterjedtebb metrika. Felhasználása túlmutat a hagyományos képalkotási eljárásokkal készült képek vizsgálatán, például CT képek rekonstrukciója során is alkalmazzák. [8]

4.3. Természetes képek modellezése

Természetes képnek (*natural scene*) nevezzük a látható környezetünk látható fénytartományban rögzített, jó minőségű, torzítatlan felvételeit. A természetes képek modellezése (továbbiakban NSS – *natural scene statistics*) a természetes képek statisztikai jellemzőivel foglalkozik. A természetes képek az összes lehetséges kép egy nagyon szűk hányadát képezik, megkülönböztetendők többek között az írásos tartalomtól, a számítógépes grafikától, rajzoktól és festményektől, valamint a nem látható tartományban rögzített – radar, röntgen, ultrahang – felvételektől. A természetes képek elemzése kiemelt fontosságú a képminőség mérés algoritmusai között. Általános feltevés, hogy a természetes környezetből érkező vizuális hatások feldolgozása vezetett a HVS evolúciójához, így az NSS modellezés a HVS modellezés duális feladatának tekinthető. [12]

Az NSS modellezést használó metrikák aktívan kutatottak és széles körben alkalmazottak, különböző képfeldolgozó algoritmusok részeként. Általános minőségmérési feladatok esetén elterjedt a VIF (*visual information fidelity*) [13] és az annak előzményét képező IFC (*information fidelity criteria*) [12]. A következőkben a VIF működési elvének ismertetése következik, példaként az NSS modellezés módszereire.

4.3.1. VIF

A VIF (*visual information fidelity*) [13] működése során csatornákra bontja a torzult és a referenciaképet wavelet-transzformáció segítségével. Az egyes csatornákra alkalmazott NSS modell alapján információelméleti módszerek segítségével megméri a képekből az emberi agy számára kinyerhető információmennyiséget. A metrika eredménye a torzult és a referenciaképből kinyerhető információmennyiség hányadosa.

Forrás modell. A wavelet-transzformált egyes csatornáinak modellezése GSM (*Gaussian scale mixture*) modell segítségével történik, mivel a GSM jól modellezi a természetes képek wavelet-transzformáltjának statisztikai jellemzőit [16]. A GSM egy olyan véletlen mező, amely két független véletlen mező szorzataként írható le. Egy I térbeli indexekkel rendelkező $\mathcal{C} = \{\vec{C}_i : i \in I\}$ véletlen mező a következő

alakban írható fel:

$$\mathcal{C} = \mathcal{S} \cdot \mathcal{U} = \{S_i \cdot \vec{U}_i : i \in I\} \quad (28)$$

ahol $\mathcal{S} = \{S_i : i \in I\}$ egy véletlen pozitív skalármező és $\mathcal{U} = \{\vec{U}_i : i \in I\}$ egy véletlen normális eloszlású vektormező, 0 várható értékkel és \mathbf{C}_U varianciával.

A VIF a wavelet-transzformált egyes csatornáit $M \times M$ méretű blokkokra bontja, amelyek a feltevés szerint függetlenek egymástól, és a blokkokból képzett oszlopvektorok alkotják az egyes \vec{C}_i értékeket. A GSM modellezés csatornánként külön zajlik.

Torzítás modell. A VIF az egyes csatornák torzulását intenzitásváltozásként és additív zajként modellezi:

$$\mathcal{D} = \mathcal{G} \cdot \mathcal{C} + \mathcal{V} = \{g_i \cdot \vec{C}_i + \vec{V}_i : i \in I\} \quad (29)$$

ahol \mathcal{C} a referenciakép csatornáját, $\mathcal{D} = \{\vec{D}_i : i \in I\}$ a vizsgált kép csatornáját modellező véletlen mezők, $\mathcal{G} = \{g_i : i \in I\}$ az jel változását leíró skalármező, $\mathcal{V} = \{\vec{V}_i : i \in I\}$ pedig a stacionárius, 0 várható értékű, additív, fehér Gauss-zajt leíró vektormező, $\mathbf{C}_V = \sigma_v^2 \mathbf{I}$ varianciával.

A fenti, wavelet-transzformáltakra alkalmazott egyszerű zajmodell segítségével modellezhetők a kép természetes torzulásai: additív fehér zaj, elmosódás, és a megvilágítás megváltozásából adódó luminancia és kontraszt eltolódás.

Kiegészítő HVS modell. Bár az NSS modellezés a HVS modellezés duális feladatának tekinthető, az emberi látásnak az NSS során modellezésre nem került aspektusai esetenként külön figyelmet érdemelnek. A VIF az alkalmazott NSS modell mellett külön modellezi a belső neurális zajt. A neurális zaj modellezése az $\mathcal{N} = \{\vec{N}_i : i \in I\}$ vektormező segítségével történik, ami egy 0 várható értékű, additív, korrelálatlan Gauss-zajt ír le. Az agyban ténylegesen a vizsgált és a referenciakép neurális zajjal torzított eredménye kerül feldolgozásra:

$$\mathcal{E} = \mathcal{C} + \mathcal{N} \quad (\text{referenciakép}) \quad (30)$$

$$\mathcal{F} = \mathcal{D} + \mathcal{N} \quad (\text{vizsgált kép}) \quad (31)$$

A zaj varianciája $\mathbf{C}_N = \sigma_n^2 \mathbf{I}$ alakban modellezhető, ahol σ_n^2 a HVS modell paramétere.

A VIF működése. A VIF, mint metrika az NSS mellett információelméleti eszközöket használ: a fenti modellek felhasználásával a csatornákból kinyerhető információmennyiség mérésével dolgozik. Az információelméleti modellhez legyen $\vec{C}^N = \{\vec{C}_1, \vec{C}_2, \dots, \vec{C}_N\}$ a \mathcal{C} mező N eleme, S^N az \mathcal{S} , \vec{D}^N a \mathcal{D} , \vec{E}^N az \mathcal{E} , \vec{F}^N pedig az \mathcal{F} mező elemei, hasonlóan. Az információmennyiség mérése \mathcal{C} és \mathcal{E} illetve \mathcal{C} és \mathcal{F} közötti *feltételes kölcsönös információ* vizsgálatával történik, adott \mathcal{S} mellett. A feltételes kölcsönös információk az agy számára a referencia- és a tesztkép egyes csatornáiból elméletileg kinyerhető információmennyiséget reprezentálják. A vizsgálat során a modelleket paraméterező \mathcal{G} , σ_v^2 és σ_n^2 ismertnek tekintendő.

Az S^N egy realizációját s^N -nel jelölve a feltételes kölcsönös információ a következő alakra hozható:

$$I(\vec{C}^N; \vec{E}^N | S^N = s^N) = \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^{M^2} \log_2 \left(1 + \frac{s_i^2 \lambda_k}{\sigma_n^2} \right) \quad (32)$$

$$I(\vec{C}^N; \vec{F}^N | S^N = s^N) = \frac{1}{2} \sum_{i=1}^N \sum_{k=1}^{M^2} \log_2 \left(1 + \frac{g_i^2 s_i^2 \lambda_k}{\sigma_v^2 + \sigma_n^2} \right) \quad (33)$$

ahol λ_k a (diagonalizálható) \mathbf{C}_U kovarianciamátrix sajátértékeit jelöli.

A fenti modell a referencia- és a vizsgált kép egy csatornáját írta le, a teljes kép információmennyisége az egyes vizsgált csatornákból kinyerhető információmennyiségek összegével modellezhető. A VIF metrika a vizsgált képből és a referenciaképből kinyerhető információk hányadosa:

$$VIF = \frac{\sum_{j \in \text{csatorna}} I(\vec{C}^N; \vec{F}^N | S^N = s^N)}{\sum_{j \in \text{csatorna}} I(\vec{C}^N; \vec{E}^N | S^N = s^N)} \quad (34)$$

A tényleges implementáció során a feltételes kölcsönös információ a véletlen mezők ergodicitását feltételezve, azok realizációiból kerül kiszámításra, paraméterbecslések segítségével. Az NSS és HVS modellekben szereplő paraméterek becslése a következőképpen történhet:

$$\hat{s}_i^2 = \frac{\vec{C}_i^T \mathbf{C}_U^{-1} \vec{C}_i}{M} \quad (35)$$

$$\hat{\mathbf{C}}_U = \frac{1}{N} \sum_{i=1}^N \vec{C}_i^T \vec{C}_i \quad (36)$$

$$\hat{g}_i = \widehat{\text{Cov}}(C, D) \widehat{\text{Cov}}(C, C)^{-1} \quad (37)$$

$$\hat{\sigma}_{v,i}^2 = \widehat{\text{Cov}}(D, D) - \hat{g}_i \widehat{\text{Cov}}(C, D) \quad (38)$$

(Az eredeti implementáció *steerable pyramid* [15] wavelet-dekompozíciót, 3×3 és 18×18 méretű ablakokat használ, valamint $\sigma_n^2 = 0.1$, kézi optimalizáció alapján kiválasztva.)

Tulajdonságok, előnyök és hátrányok, alkalmazhatóság. A VIF pozitív metrika, a referencia- és a vizsgált kép megegyezése esetén értéke 1, a minőség romlásával nő. A torzítás típusok nagy része esetén értéke 0 és 1 közötti. A VIF ugyanakkor nem szimmetrikus metrika, 1-nél nagyobb értéket is felvehet, ha a vizsgált képet jobb minőségűnek értékeli, mint a referenciaképet. Ilyen szituáció például kontrasztosítás esetén fordulhat elő, az 6. ábrán látható módon.



6. ábra. A Lena tesztkép eredeti, és kontrasztosított változata (VIF = 1.21)

Az alkalmazott NSS modell, valamint az NSS és a HVS modellezés közötti kapcsolatból adódóan a VIF jól modellezi a természetes képek minőségét az ember érzékelés szempontjából, és jól reagál a gyakorlatban előforduló torzítás típusokra. Hatékonyságát a szubjektív adatbázisok segítségével történő tesztek is alátámasztják, más metrikákkal (például MSE, PSNR, SSIM) szemben is. Hátránya algoritmikus bonyolultsága, számításigénye, és hogy csak természetes képeken, a szóbjáhető képek kis hányadán működik megbízhatóan.

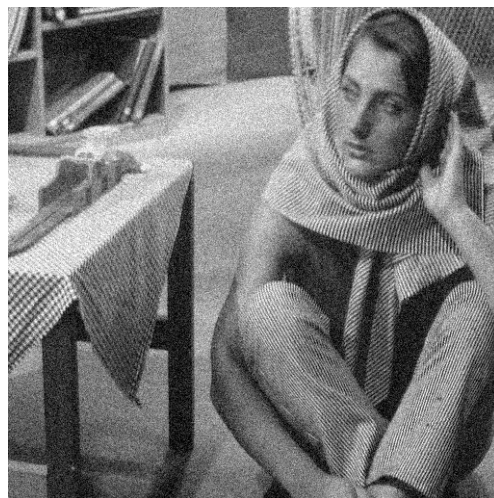
5. Referenciakép nélküli metrikák

A fejezetben a referenciakép nélküli NR metrikák főbb típusainak jellemzése és egy példaként bemutatott NR metrika található. Az NR metrikák általános előnye – miszerint működésükhöz nem szükséges a vizsgált kép jó minőségű referenciája, amely megközelítés közelebb áll a HVS működéséhez —, egyben a hátrányuk is: a referenciakép nélküli megbízható minőségméréshez *a priori* ismeretekre van szükség a vizsgált képről, így ezek a metrikák sok esetben csak a képek egy szűk körén működnek hatékonyan.



SSIM: 1

Q: 57.6



SSIM: 0.488

Q: 26.74



SSIM: 1

Q: 26.16



SSIM: 0.356

Q: 6.63

7. ábra. Az SSIM és a Q metrika összehasonlítása 0.006 szórású Gauss-zaj mellett

A referenciakép nélküli metrikák sokszor nem önálló metrikaként, hanem egy iteratív képjavító algoritmus részeként, annak optimalizációs célfüggvényeként jelennek meg a gyakorlatban. Önálló metrikaként való megjelenésüket sok esetben gátolja, hogy míg az FR metrikák többnyire a minőséget abszolút értelemben jellemző értéket rendelnek a képhez, az NR metrikák esetén ez sokszor egy relatív érték². Ha tekintünk két különböző referenciaképet, és azoknak többféle torzult változatát, akkor egy „abszolút” FR metrika esetén az első kép és annak egy torzított változatát összehasonlítva illetve a második kép és annak egy torzított változatát összehasonlítva a metrikából két kapott érték egymással is összemérhető. „Relatív” NR metrikák esetén a metrikából kapott értékek egy adott referenciakép torzult változatai között összemérhetők, de a másik referenciaképpel és annak torzult változatainak értékeivel nem feltétlenül. Az 7. ábra ezt a jelenséget szemlélteti, a SSIM (FR) és Q metrika (NR) között, két tesztképet és azok ugyanolyan módon torzított változatát megjelenítve.

5.1. Típusok

Az NR metrikák 3 fő csoportba oszthatók a típusuk szerint: torzítás elemzésen, jellemző alapú tanuláson és a természetes képek modellezésen alapuló metrikákra. [11] A felosztás jellemzően nem kizárólagos, egy metrika többféle tulajdonsággal is rendelkezhet.

5.1.1. Torzítás elemzés

A torzítás elemzésen alapuló módszerek egy vagy több adott torzítás hatását mérik, és az alapján értékelik a vizsgált képet. Alkalmazhatóságuk nagyban függ az alkalmazott torzítás modelltől, amely korlátozza a vizsgálható képek körét. A vizsgált torzítások főként a gyakorlatban megjelenő főbb torzítás típusok közül kerülnek ki.

²Megjegyzendő, hogy az MSE, mint FR metrika is csak (a fenti értelemben) relatív módon jellemzi a képeket, illetve egyes torzulásokat mérő NR metrikák pedig pont (a fenti értelemben) abszolút módon jellemzik azokat.

5.1.2. Jellemző alapú tanulás

A jellemző alapú tanuló módszerek több, a képekből kiemelt jellemzőt vizsgálnak. Az algoritmusok működése két fázisra osztható: tanulási és kiértékelési fázisra. Az első fázis egy tanuló algoritmus betanítása a később vizsgálandó képekhez hasonló jellegű, de ismert MOS vagy DMOS értékű, változatos minőségű képhalmaz segítségével. A tanulás a kiemelt jellemzők vizsgálatán keresztül történik, például SVM (*support vector machine*) vagy *neurális háló* modell segítségével. A vizsgálandó képek kiértékelése a betanított algoritmus segítségével, szintén a megfelelő jellemzők vizsgálatával történik.

A módszer előnye, hogy használatához nem szükséges a képeken jelenlévő torzítások pontos modellezése, megfelelő jellemző kiválasztás és tanítás mellett a módszer a tanításhoz használt képhalmaz torzítás típusaira lesz érzékeny. Hátránya azonban, hogy hatékonysága nagyban függ a vizsgált jellemzőhalmaztól, a tanításhoz használt képhalmaztól és az alkalmazott tanuló algoritmustól.

5.1.3. Természetes képek modellezése

Az NSS alapú NR metrikák az NSS alapú FR metrikákhoz hasonló elveken működnek, az ott leírtaknak megfelelően. NR metrikák esetén, referenciakép hiányában a minőség például a vizsgált kép és a természetes képek halmaza közötti távolságként közelíthető meg.

5.2. Q metrika

A példaként bemutatott NR metrika, a *Q metrika* [19] egy élességet mérő metrika, egy iteratív képjavító eljárás részeként. A metrika a szinguláris érték felbontást (SVD – *singular value decomposition*) használja a domináns gradiens irányok és azok energiájának meghatározásához. A metrika értéke a szinguláris értékek mellett a képre rakódó zaj modellezésével kerül kiszámításra.

5.2.1. Működés

Jelölje $p = (p_i)_{i=1}^{N^2}$ a vizsgált kép egy $N \times N$ méretű blokkját.

A metrika alapja a gradiens mátrix elkészítése. A kép egy adott p_k pontjában a gra-

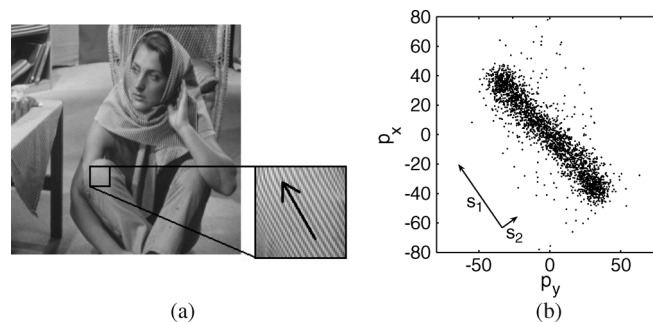
diens vektor többféle szűrő segítségével meghatározható, például Prewitt- és Sobel-operátor, Canny-szűrés vagy akár Hough-transzformáció használatával is. A gradiens mátrix a következő alakban írható:

$$G = \begin{bmatrix} \vdots & \vdots \\ p_x(k) & p_y(k) \\ \vdots & \vdots \end{bmatrix} \quad (39)$$

A domináns orientáció meghatározása SVD segítségével történik:

$$G = USV^T = U \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T \quad (40)$$

ahol U és V ortonormált mátrixok, $s_1 \geq s_2 \geq 0$ a szinguláris értékek. A v_1 vektor jelöli a domináns irányt, a v_2 vektor a rá merőleges domináns él irányt, a s_1 és s_2 szinguláris értékek a két irány energiáját reprezentálják.



8. ábra. Gradiens irányok, és a domináns gradiens irány és energia [19]

Az élesség mérésére a Q metrika a következő értéket javasolja:

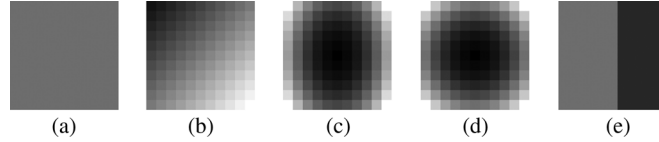
$$Q = s_1 \frac{s_1 - s_2}{s_1 + s_2} \quad (41)$$

A képlet megalkotása az s_1 és s_2 értékek egyszerű blokkokon: egyszínű, lineáris színátmenetes, kvadratikusan (izotróp és anizotróp) színátmenetes és élt tartalmazó foltokon történő vizsgálata alapján, additív, fehér Gauss-zajt feltételezve történt.

A képlet fontos része az R koherencia, amely a zaj szórásnégyzetének reciprokát becsüli:

$$R = \frac{s_1 - s_2}{s_1 + s_2} \approx \frac{1}{\sigma^2} \quad (42)$$

A foltok vizsgálata alapján a metrika nem a kívánt viselkedést mutatja izotróp blokkokon, ezért a metrika végső mérése előtt szűrésre kerülnek a vizsgált kép blokkjai



9. ábra. A vizsgálathoz használt foltok [19]

anizotrópia szempontjából. A szűrés a koherencia statisztikai jellemzői alapján hipotézisvizsgálat segítségével történik: a nullhipotézis szerint az adott blokk izotróp additív, fehér Gauss-zajt feltételezve. A nullhipotézis $0 < \delta < 1$ szignifikancia szint mellett elvetésre kerül – vagyis a blokk anizotróp –, ha $R \geq \tau$, ahol az R statisztikai jellemzői alapján:

$$\delta = \left(\frac{1 - \tau^2}{1 + \tau^2} \right)^{N^2 - 1} \quad (43)$$

A teljes képre történő kiértékelés $N \times N$ méretű nem átfedő ablakok segítségével, az ablakokat izotrópia szerint szűrve, csak az anizotrópokot vizsgálva történik. A képhez rendelt végső érték az egyes ablakok értékeinek átlaga, ahol egy izotróp ablak értéke 0. M ablak esetén, az egyes ablakokban számított értékeket $(Q_j)_{j=1}^M$ -vel jelölve:

$$Q = \frac{1}{M} \sum_{j=1}^M Q_j \quad (44)$$

(Az eredeti implementáció $N = 8$ méretű ablakot használ, és $\delta = 0.001$ szignifikancia szintet.)

5.2.2. Tulajdonságok, előnyök és hátrányok, használhatóság

A metrika célzottan képminőséget javító algoritmusokhoz készült, azok iteratív minőségjavításának optimalizálásához. Értéke nemnegatív, a kép minőségének romlásával csökken. A szubjektív adatbázisokon történő tesztek azt mutatják, hogy változása jól követi a minőség romlását elmosódás és Gauss-zaj jelenléte esetén. Hátránya, hogy csak relatív értelemben méri a kép minőségét: ugyanazon kép különböző változatai között észleli a minőségbeli különbséget, de két különböző kép esetén nem összemérhetők a kapott értékek, a fejezet elején látható 7. ábrán látható módon.

6. Összehasonlító elemzés



Referencia

SSIM: 1

VIF: 1

Q: 26.16

Intenzitáseltolás

SSIM: 0.982

VIF: 1

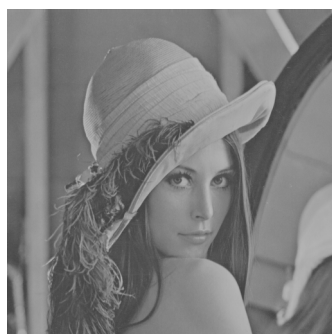
Q: 26.16

Kontrasztosítás

SSIM: 0.91

VIF: 1.212

Q: 37.096

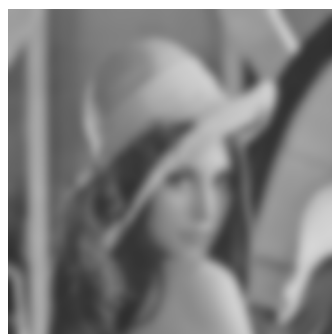


Kontraszt gyengítés

SSIM: 0.922

VIF: 0.688

Q: 15.19



Elmosódás

SSIM: 0.647

VIF: 0.0536

Q: 10.43



Só-bors zaj

SSIM: 0.583

VIF: 0.292

Q: 16.03



Gauss-zaj

SSIM: 0.356

VIF: 0.288

Q: 6.63



Kvantumzaj

SSIM: 0.378

VIF: 0.307

Q: 8.156

10. ábra. A vizsgált metrikák értékei a Lena tesztkép néhány torzítása esetén

A metrikák működésének bemutatása és jellemzésük után a fejezetben a metrikák egymással kerülnek összehasonlításra. Az összehasonlítás során az MSE, SNR, PSNR, SSIM [17], VIF [13] és a Q metrikát [19] vizsgáltam, numerikus tulajdonságaik, alkalmazhatóságuk köre, a különböző torzításokra történő reakciójuk és a szubjektív adatbázison mért hatékonyságuk alapján.

Az 10. ábrán a metrikák szemléletes összehasonlítása látható, a metrikák által meghatározott pontszám néhány jellemző torzítás használata mellett. A referenciakép kivételével az összes változat esetén a hibamérésen alapuló metrikák azonos eredményt mutatnak (MSE: 400, SNR: 16.46, PSNR: 22.11).

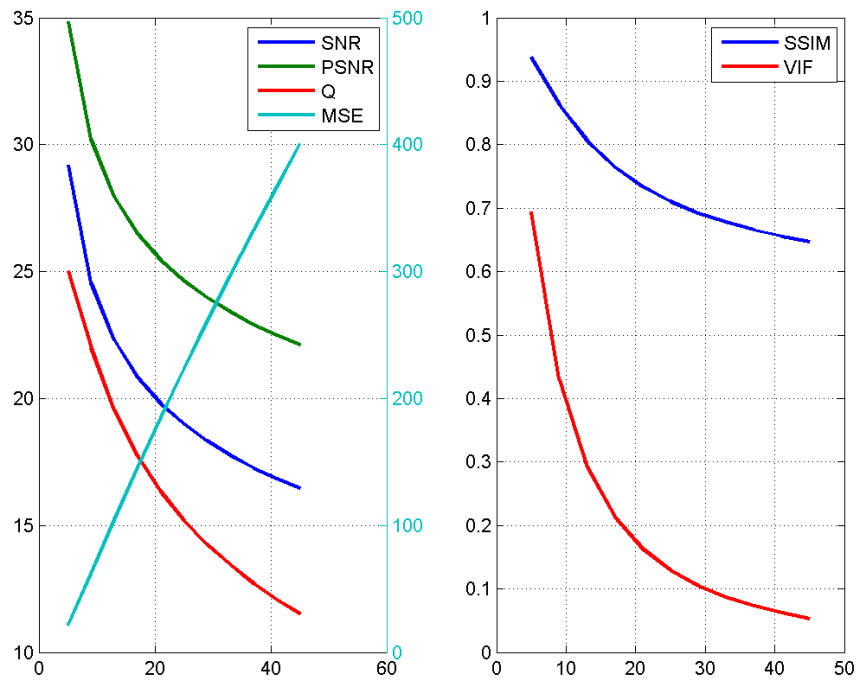
Metrika	Intervallum	Optimális érték	Monotonitás
MSE	$[0, \infty]$	0	növekvő
SNR	$[0, \infty]$	∞	csökkenő
PSNR	$[0, \infty]$	∞	csökkenő
SSIM	$[-1, 1]$	1	csökkenő
VIF	$[0, \infty], ([0, 1])$	1	csökkenő
Q metrika	$[0, \infty]$	–	csökkenő

1. táblázat. Metrikák numerikus összehasonlítása

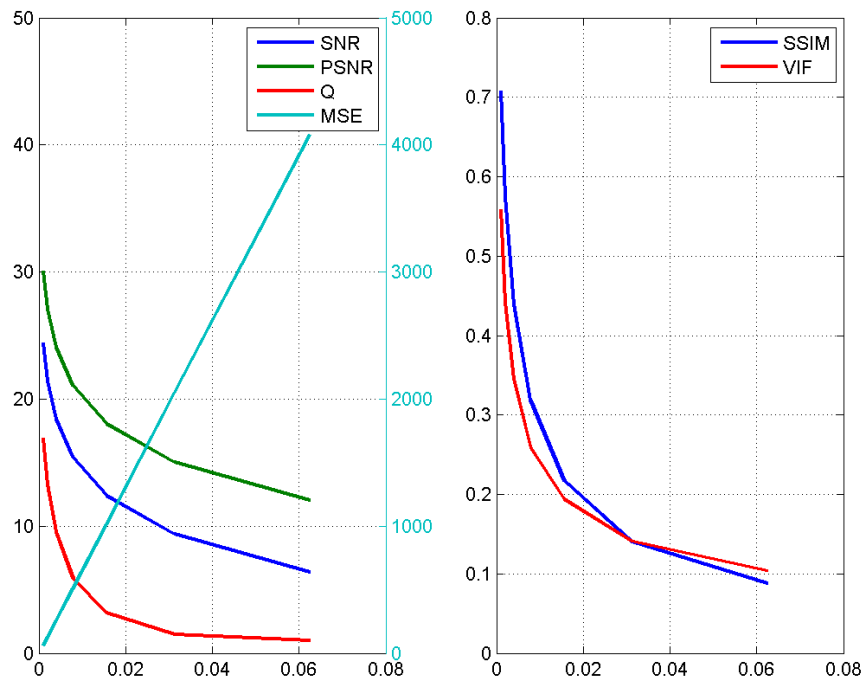
Az 1. táblázat a metrikák numerikus összehasonlítását tartalmazza. Az intervallum oszlop a metrika lehetséges értékei, az optimális érték referenciaképet használó metrikák esetén a tökéletes minőségbeli egyezést jelző érték, a monotonitás a metrika viselkedése a képminőség romlásának hatására. A VIF metrika a leggyakoribb torzítások esetén a $[0, 1]$ intervallumba esik, 1-nél nagyobb felvett érték esetén a vizsgált képet jobb minőségűnek értékeli, mint a referenciaképet, az 6. ábrán látható módon. A Q metrika intervalluma függ a vizsgált kép dinamikus tartományától és a gradiens kiszámításának módjától.

Az 11. illetve 12. ábrák grafikonjai a metrikák numerikus viselkedését szemléltetik a képekhez képfüggetlen, 0 várható értékű, additív, fehér Gauss-zajt adva, illetve a kép elmosódását Gauss-szűrő segítségével szimulálva, a zaj szórásának, illetve a Gauss-szűrő méretének függvényében. A szimuláció során alkalmazott $N \times N$ méretű Gauss-ablakok szórása $N/6$.

Az 2. táblázat a metrikák alkalmazhatóságának körét és hatékonyságukat mu-



11. ábra. Az egyes metrikák viselkedése elmosódás hatására



12. ábra. Az egyes metrikák viselkedése zaj hozzáadása esetén

Metrika	Referencia	Korlátozás	Hatékonyság	SROCC
MSE	igen	–	közepes	0.5531
SNR	igen	–	közepes	0.5235
PSNR	igen	–	közepes	0.5531
SSIM	igen	–	jó	0.6233
VIF	igen	természetes képek	legjobb	0.6233
Q metrika	nem	anizotróp képek elmosódás és zaj	gyenge	0.3646

2. táblázat. Metrikák összehasonlítása alkalmazhatóság szempontjából

tatja. A referencia oszlop a metrika típusára utal: használ-e referenciaképet működése során. A korlátozás a felhasználható képek körére illetve a torzulás típusokra vonatkozó megkötések, „–” esetén a metrika általánosan használható, és reagál a gyakorlatban előforduló torzítás típusokra. A hatékonyság a szubjektív eredményekkel való összehasonlítás szöveges értékelése, az SROCC oszlop pedig a mért rangkorrelációs együtthatók. Az összehasonlítást a TID [9] adatbázis képei és szubjektív pontszámai alapján végeztem el, Spearman rang-korreláció segítségével, a szubjektív mérés, adatbázisok részben írtaknak megfelelően. Az adatbázis 25 referenciakép többféleképpen és különböző mértékben torzított változatát tartalmazza, azok 0 és 9 közötti MOS értékeivel. A mérést az implementációról szóló fejezetben bemutatott eszközökkel és megvalósításokkal végeztem, a Q metrika pontszámát az adott referenciakép pontszámának felhasználásával a $[0, 1]$ intervallumra normáltam. A mérés azt mutatja, hogy a Q metrika csak elmosódás és additív zaj esetén írja le jól a kép érzékelhető minőségét, általános esetben gyengén teljesít. Az FR metrikák közül az emberi észlelés szempontjából leggyengébben az MSE és a PSNR teljesít, az SSIM felülmúlja őket, a legmegbízhatóbb pedig a VIF metrika eredménye.

7. Implementáció és a szemléltető alkalmazás

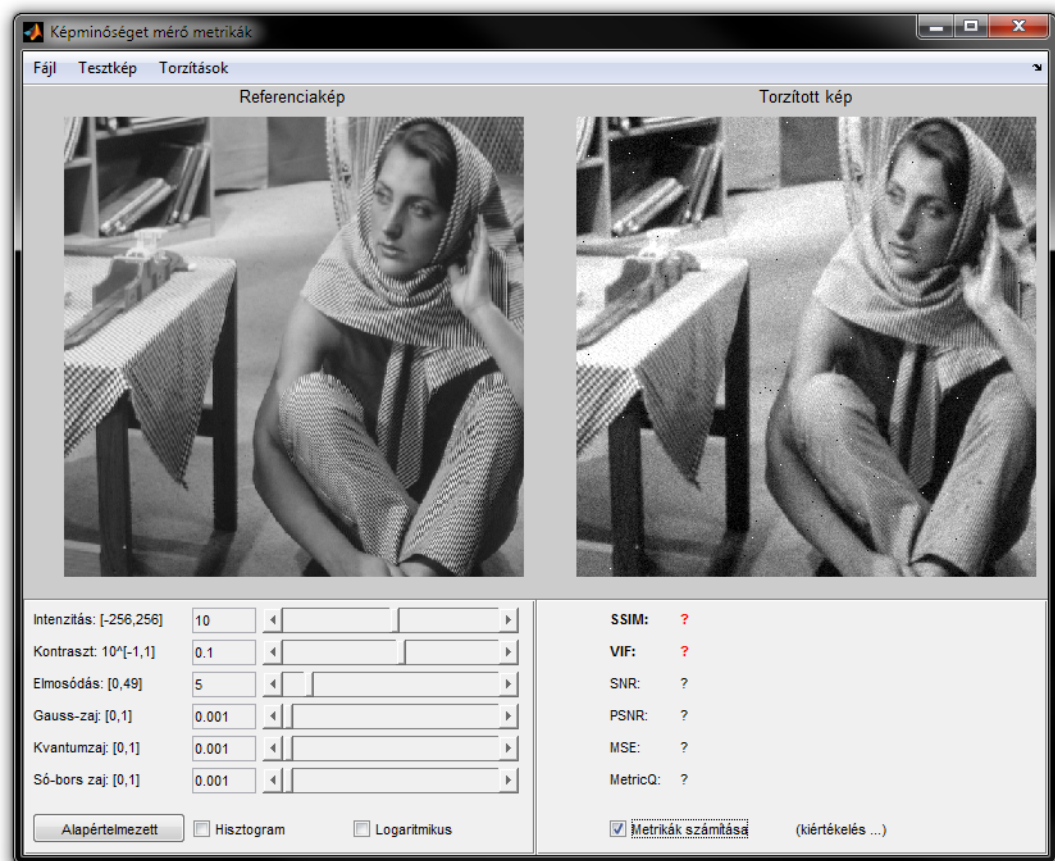
Munkám során MATLAB környezetben implementáltam az FR metrikák közül az alapvető hiba mérésen alapuló MSE, SNR és PSNR metrikákat, és a strukturális hasonlóságon alapuló SSIM [17] metrikát, valamint az NR metrikák közül a Q metrikát [19]. Munkám eredménye az előző fejezetbeli összehasonlító tesztek mellett egy grafikus felületű szemléltető alkalmazás, melynek segítségével többféle torzulást lehet szimulálni, és megvizsgálni ezen torzulások metrikákkal történő mérhetőségét. A program az implementált metrikákon kívül kiszámítja a VIF [13] metrika értékét is, a szerzők honlapján közzétett implementáció és a *steerable pyramid* [15] csomag segítségével. A fejezet a metrikákat megvalósító MATLAB függvények és a szemléltető alkalmazás dokumentációját tartalmazza.

7.1. A szemléltető alkalmazás dokumentációja

Az alkalmazás egy referenciakép betöltésére és torzítására ad lehetőséget. A program szürkeárnyalatos, 8 bites, 0 és 255 közötti dinamikus intenzitástartományú képekkel dolgozik. Lehetőség van színes képek betöltésére is, ezek luminanciája kerül felhasználásra.

7.1.1. A program indítása és felülete

A program MATLAB parancssorból, a `MetricGUI` paranccsal indítható. A program indításakor opcionálisan megadható paraméterként a használandó referenciakép. Ha nincs kép megadva, akkor a Lena tesztkép illetve annak hiányában egyszínű szürke kép kerül betöltésre, de a program lehetőséget nyújt referenciakép fájlból történő betöltésére, és néhány standard tesztkép használatára. A program indítása után annak grafikus felülete jelenik meg, az 13 ábrán látható. A program ablaka tetszőleges méretezhető és mozgatható, indításkor a képernyő közepén, 800×600 méretű felülettel jelenik meg. A felület egy menüsorból, a referenciakép és a torzított kép illetve hisztogramjaik megjelenítéséből, a torzításokat szabályozó és a metrikák pontszámát mutató panelekből áll.



13. ábra. A szemléltető alkalmazás felülete

7.1.2. Menüpontok

A menüpontokon keresztül végezhető a referenciakép betöltése, a torzított kép mentése, és néhány előre definiált torzítás kombináció kiprobálása. A menük egyes menüpontjainak funkciói:

A **Fájl** menün keresztül végezhető a referenciakép betöltése és a torzított kép mentése, menüpontjai a következők:

- *Referenciakép beolvasása*: referenciakép fájlból történő beolvasása, a leggyakoribb fájlformátumok (JPEG, PNG, TIFF, GIF) valamelyikének használatával. A referenciakép tetszőleges felbontású, színes vagy szürkeárnyaltos lehet, színes kép esetén az `rgb2gray` paranccsal szürkeárnyaltossá konvertálva kerül feldolgozásra.
- *Torzított kép mentése*: a torzított kép fájlba történő mentése. A kép szürkeárnyaltos formában kerül mentésre, a fájlnev kiterjesztésében meghatározott formátumban.

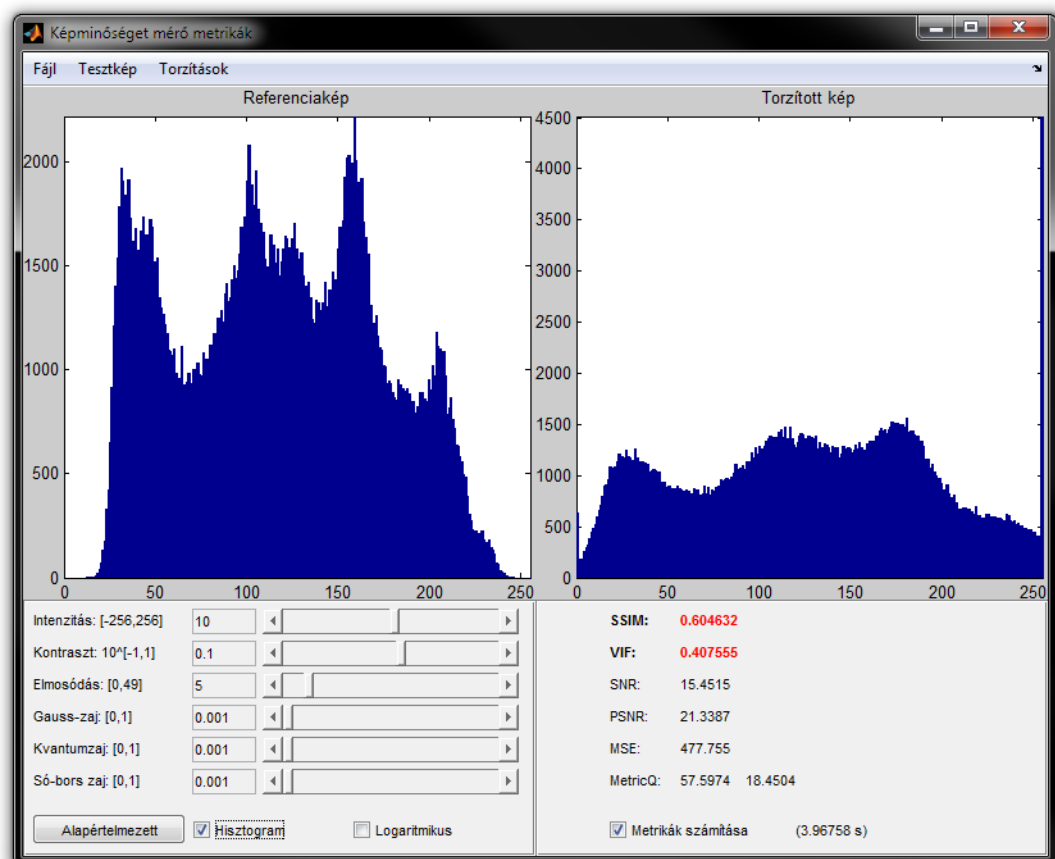
- *Kilépés*: a program bezárása.

A **Tesztkép** menü keresztül egy egyszínű szürke kép, valamint a standard *Lena*, *Barbara*, *Man*, *Cameraman*, *Goldhill*, *Boat* és *Peppers* tesztképek tölthetők be.

A **Torzítások** menü néhány előre definiált torzítás kombináció kipróbálására nyújt lehetőséget. A torzítások a 10. és a ?? ábrán bemutatott beállításokat tartalmazzák.

7.1.3. A képek és a hisztogramok megjelenítése

A grafikus felület központi része a referenciakép és a torzított kép megjelenítését szolgálja. A képek azonos méretben, az ablak méretéhez igazodó nagyításban jelennek meg. A *Hisztogram* jelölőnégyzet bekapcsolásával a képek helyett azok hisztogramjait lehet megjeleníteni, egyszerűen vagy a *Logaritmikus* jelölőnégyzetet bekapcsolva Y tengely mentén logaritmikus skála mellett.



14. ábra. Hisztogramok megjelenítése a képekhez

7.1.4. Torzítás szimuláció

Az alkalmazás segítségével a képrögzítés során jelentkező főbb torzítás típusok hatása szimulálható. A torzítás mértéke minden esetben egy szövegbeviteli mező és egy csúszka segítségével adható meg, amelyek a referenciakép alatt elhelyezkedő panelen találhatók. Az egyes típusoknál megadott tartományok csak a csúszka korlátai, a program ezen kívül eső értékeket is elfogadhat, de ekkor a csúszka nem használható a paraméter megadására. A paraméterek az *Alapértelmezett* gomb segítségével visszaállíthatók kezdeti, torzításmentességet reprezentáló értékeikre. A szimulált torzítások a következők:

- *Intenzitás*: -256 és 256 közötti paraméterrel megadott intenzitás-eltolódás.
- *Kontraszt*: 0.1 és 10 közötti szorzófaktorral 10 alapú logaritmikus skálán megadott kontraszt nyújtás. A hisztogram a kép átlagintenzitásához képest kerül nyújtásra.
- *Elmosódás*: Gauss-szűrővel modellezett elmosódás, a Gauss-szűrő ablakmérete 0 és 49 között megadható egész paraméter. A program N méretű ablak esetén $N/6$ méretű szűrőt használ.
- *Gauss-zaj*: 0 várható értékű, képfüggetlen, fehér, additív Gauss-zaj, 0 és 1 között megadható szórással.
- *Kvantumzaj*: a szimulált kvantumzaj megadható paramétere annak χ intenzitása, 0 és 1 közötti értéként.
- *Só-bors zaj*: az egyenletes térbeli eloszlású só-bors zaj 0 és 1 közötti paramétere a sérült pixelek arányát jelenti a képen. A sérült pixelek fele feketére, fele fehérre módosul.

A program az egyes torzításokat a fenti sorrendben hajtja végre, azaz az alkalmazott kombinált torzítás modell a megvilágításból származó intenzitás és kontrasztváltozást, a képrögzítés során az esetleges bemozdulás és az eszköz beállításai miatt bekövetkező elmosódást, az elektromos eszközök háttérzaját, a kvantumzajt, a rögzítési hibákból adódó só-bors zajt foglalja magába. Az egyes hatások alkalmazása után eredményként kapott torzított kép 0 és 255 közötti dinamikus tartományra és 8 bitre korlátozva kerül további feldolgozásra.

7.1.5. Metrikák pontszámai

Az egyes metrikák által meghatározott pontszámok a torzított kép alatt elhelyezkedő panelen láthatók. A metrikák csak a *Metrikák számítása* jelölőnégyzet bekapcsolása mellett kerülnek kiértékelésre, mivel az egyes metrikák, elsősorban a VIF metrika kiszámítása a számítógéptől függően több időt vehet igénybe, így esetenként előnyösebb csak a torzítások paramétereinek behangolása után kiszámítani a metrikák eredményét. A jelölőnégyzet melletti szövegmező a *(kiértékelés ...)* feliratot tartalmazza a metrikák pontjainak számítása során, a kiértékelés végeztével az ahhoz szükséges időt jeleníti meg, másodpercben kifejezve. A Q metrika esetén két érték szerepel a felületen: a referenciakép és a torzított kép pontszáma.

7.2. Metrika függvények dokumentációja

Az MSE, SNR, PSNR, SSIM és Q metrikák MATLAB függvények formájában kerültek implementálásra, melyeket önállóan is lehet használni képek minőségének kiértékelésére. A függvények paraméterei a szürkeárnyalatos referencia- és vizsgált kép, opcionálisan további beállítások, kimenete a meghatározott pontszám, opcionálisan további információk a számításról. Az egyes függvények hibát jeleznek, ha a bemeneti paraméterek nem megfelelőek: nem egyezik meg a referencia- és a vizsgált kép mérete, érvénytelen vagy hiányzó paraméterértékek kerültek megadásra.

Általános dokumentációként a függvényt **X**-szel jelölve:

- `[x, ...] = X(reference, test, ...)`

Bemeneti paraméterek:

- **reference**: referenciakép (szürkeárnyalatos)
- **test**: vizsgált kép (szürkeárnyalatos)

Kimeneti paraméter:

- **x**: az X metrika értéke

Hibaüzenet:

- *A referenciakép és a vizsgált kép különböző méretű*

7.2.1. MSE, SNR, PSNR

- `e = mse(reference, test)`
- `r = snr(reference, test)`
- `p = psnr(reference, test, L)`

Opcionális bemeneti paraméter:

- `L`: a dinamikus tartomány felső határa, alapértelmezett értéke 255.

7.2.2. SSIM

- `[mssim, map] = ssim(reference, test, K, window, L)`

Opcionális bemeneti paraméterek:

- `K`: az SSIM numerikus stabilitását biztosító `[K1 K1]` konstansok, alapértelmezett értékük `[0.01 0.03]`.
- `window`: a számításhoz használt lokális ablak, alapértelmezetten egy 11×11 méretű, 1.5 szórású Gauss-szűrő ablaka.
- `L`: a dinamikus tartomány felső határa, alapértelmezett értéke 255.

Opcionális kimeneti paraméter:

- `map`: a lokálisan kiszámított SSIM értékek, a bemeneti képekkel megegyező méretű mátrix formájában.

Lehetséges hibaüzenetek:

- *Az ablakméret nagyobb, mint a kép*
- *Hiányzó/túl sok `[K1 K2]` konstans*
- *Negatív `[K1 K2]` konstans*

7.2.3. Q metrika

- `[Q,map,anisomap] = metricq(test, N, delta)`

Opcionális bemeneti paraméterek:

- `N`: a blokkok mérete, alapértelmezett értéke 8.
- `delta`: az anizotróp blokkok szűréséhez megadható szignifikancia szint, alapértelmezetten 0.001.

Opcionális kimeneti paraméter:

- `map`: a lokálisan kiszámított SSIM értékek.

Lehetséges hibaüzenetek:

- *A blokkméret 2-nél kisebb*
- *Hiányzó/túl sok [K1 K2] konstans*
- *A szignifikancia szint negatív, 0, vagy 1-nél nagyobb*

Hivatkozások

- [1] A57 dataset. <http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>
- [2] P. Le Callet, F. Autrusseau: Subjective quality assessment irccyn/ivc database. <http://www2.irccyn.ec-nantes.fr/ivcdb/>
- [3] U. Engelke, H.-J. Zepernick, M. Kusuma: Wireless imaging quality database. <http://www.bth.se/tek/rcg.nsf/pages/wiq-db>
- [4] A. Foi, M. Trimeche, V. Katkovnik, K. Egiazarian: Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data, IEEE Transactions on Image Processing, 2008
- [5] Y. Horita, Y. Kawayoke, Z.M. Parvez Sazzad: Image quality evaluation database. http://160.26.142.130/toyama_database.zip
- [6] E.C. Larson, D.M. Chandler: Most apparent distortion: full-reference image quality assessment and the role of strategy, Journal of Electronic Imaging, 2010
- [7] W. Lin., C.-C.J. Kuo: Perceptual visual quality metrics: A survey, J. Vis. Commun. Image R., 2011
- [8] L. Pfister, Y. Bresler: Tomographic reconstruction with adaptive sparsifying transforms, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2014
- [9] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti: TID2008 – a database for evaluation of full-reference visual quality assessment metrics, Advances of Modern Radioelectronics, 2009. <http://www.ponomarenko.info/tid2008.htm>
- [10] M.A. Saad, A.C. Bovik, C. Charrier: A DCT statistics-based blind image quality index, IEEE Signal Processing Letters, 2010
- [11] M.A. Saad, A.C. Bovik, C. Charrier: Blind image quality assessment: a natural scene statistics approach in the DCT domain, IEEE Transactions on Image Processing, 2012

- [12] H.R. Sheikh, A.C. Bovik, G. de Veciana: An information fidelity criterion for image quality assessment using natural scene statistics, *IEEE Transactions on Image Processing*, 2005
- [13] H.R. Sheikh, A.C. Bovik: Image Information and visual quality, *IEEE Transactions on Image Processing*, 2006
- [14] H.R. Sheikh, Z. Wang, A.C. Bovik, L.K. Cormack: Image and video quality assessment research at live. <http://live.ece.utexas.edu/research/quality/>
- [15] E.P. Simoncelli, W.T. Freeman: The steerable pyramid: A flexible architecture for multi-scale derivative computation, *IEEE Second Int'l Conf on Image Processing*, 1995
- [16] M.J. Wainwright, E.P. Simoncelli, A.S. Wilsky: Random cascades on wavelet trees and their use in analyzing and modeling natural images, *Applied and Computational Harmonic Analysis*, 2001
- [17] Z. Wang, A.C. Bovik, H. R. Sheikh, E. P. Simoncelli: Image quality assessment: From error visibility to structural similarity, *IEEE Transactions on Image Processing*, 2004
- [18] Z. Wang, A.C. Bovik: A universal image quality index, *IEEE Signal Processing Letters*, 2002
- [19] X. Zhu, P. Lianfar: Automatic parameter selection for denoising algorithms using a no-reference measure of image content, *IEEE Transactions on Image Processing*, 2011

Függelék: Rövidítések

Rövidítés	Kifejtés	Magyarázat
DCT	<i>discrete cosine transform</i>	diszkrét koszinusz-transzformált
DFT	<i>discrete Fourier transform</i>	diszkrét Fourier-transzformált
DMOS	<i>differential mean opinion score</i>	a vizsgált és a referenciakép szubjektív MOS pontszámainak különbsége
DWT	<i>discrete wavelet transform</i>	diszkrét wavelet-transzformált
FR	<i>full-reference</i>	referenciaképet használó metrika (a teljes referenciaképet)
GSM	<i>Gaussian scale mixture</i>	NSS modellezési technika [16]
HVS	<i>human vision system</i>	emberi látórendszer
IFC	<i>information fidelity criteria</i>	NSS modellezésen alapuló metrika, a VIF elődje [12]
MOS	<i>mean opinion score</i>	emberi szemlélők segítségével meghatározott szubjektív pontszám
MSE	<i>mean-squared error</i>	az átlagos négyzetes eltérés
NSS	<i>natural scene statistics</i>	természetes képek statisztikai modellezésén alapuló metrika típus
PSNR	<i>peak signal-to-noise ratio</i>	jel-zaj arány
RMSE	<i>root-mean-squared error</i>	az átlagos négyzetes eltérés négyzetgyöke
RR	<i>reduced-reference</i>	referenciaképet használó metrika (a referenciakép egy részét)
SNR	<i>signal-to-noise ratio</i>	jel-zaj arány
SVD	<i>singular value decomposition</i>	szinguláris érték felbontás
SSIM	<i>structural similarity</i>	strukturális hasonlóságon alapuló metrika [17]
UQI	<i>universal quality index</i>	strukturális hasonlóságon alapuló metrika, az SSIM „elődje” és speciális esete [18]
VIF	<i>visual information fidelity</i>	NSS modellezésen alapuló metrika [13]

3. táblázat. Rövidítések