2023.09.10. - A 3D-scanner őse, s egyben a szépség-matematikájának mérnöki alapja: <https://en.wikipedia.org/wiki/Beauty_micrometer>

A szépség evolúciós szempontból a leginkább „átlagos”/normaszerű (arc) állapot, mert az agynak bármely pontja töredékei alapján is egyszerűen (energiatakarékosan) vélelmezhető. A sok egyed eltérő méreteinek leggyakoribb/átlagos/… értéke jól közelítik vélhetően az aranymetszés szabályát?!

Robot-esztéta, avagy a szépség matematikája

(Robot-aesthetic-expert or the mathematics of the beauty)

Pitlik László, Pitlik Marcell, Pitlik Mátyás (MY-X team)

Kivonat: Az aranymetszés a természet formavilágának egy karakteres arányára (phi=1.618) mutat rá. A szépség matematikáját firtató kutatás célja ezzel párhuzamosan a robotszemek számára egy olyan mesterséges szépség-skála alkotása (itt és most egy nonfiguratív tesztkép és egy random pontokból álló kép összehasonlítására támaszkodva), mely skála alapján egy robot különbséget tud tenni context-free módon (vagyis a mindenkori kép potenciális tartalmától függően) szép és csúnya, szebb és még szebb között. Vagyis a szépség matematikájától annyit várunk, hogy átmenjen a Turing-teszten és mindennemű speciális esztétikai képzés nélkül, tanulást támogató kép-adatbázis nélkül, korábbi emberi értékítéletek ismerete nélkül, vagyis már egyetlen egy képről is vélelmezni tudja, hogy (koncepciózus) műalkotás-e, vagy inkább a véletlen (koncepciótlan) terméke. Ezen skála megteremtése matematikailag egyszerű, mert a skála eredete a kockázat context-free fogalmára vezethető vissza, ahol minden adatpont kockázatos, mely a többi adatpont (az adatpontok egymáshoz képesti hasonlóságai) alapján nem becsülhető tetszőleges pontossággal. A skála melléktermékeként is tekinthetünk a robot-esztétára, aki nem más, mint egy-egy kép kapcsán a leginkább kritikus pixelek, foltok kijelölésére alkalmas, vagyis a kompozíciót racionálisan (s ismét csak nem véletlenszerűen) kritizálni képes mesterséges intelligencia-alapú eljárás. A robot-esztéta operatív működését úgy lehet elképzelni, mint a kevésbé jól éneklő ember hangfelvételét az ideális hangmagasságokhoz korrigálni képes eljárás, ahol az ideál a képpontok közötti matematikai harmónia maga.

Kulcsszavak: művészet, automatizálás, hasonlóságelemzés-láncok,

Abstract: The Vitruvian man of Leonardo Da Vinci is a kind of evidence of the specific ratio (divine proportion, golden ratio) being available in the nature (c.f. phi=1.1618). Parallel, the mathematics of the beauty has to derive a scale for robot-eyes. This scale should be capable of estimating/describing the difference between two pictures (here and now between a randomized pixel-set and a real nonfigurative painting) in a context-free way. The scale has to support the ranking of the beauty of pictures. It is therefore possible to derive beauty and its inverse, or even a good-better-best distinction between pictures. The kind of artificial intelligence-based beauty-definition has to lead to a positive Turing-test – without any training, without human evaluation of pictures in advance. The robot-aesthetic-expert (the robot-eye) is capable of sensing/estimating the ratio of the visual conception and the coincidence. This beauty-scale could be derived from the context-free risk-definition of the similarity analysis where risk-free is a data pattern where each data position can be derived from the rest. The same logic can be used for the definition of the harmony/beauty. A kind of side-effect is a visual risk-map. This map gives red alert signals if the interpretability of a part of the picture is less than the interpretability of the neighbourhood. The risk-map can be used by artists to improve the compositions – like sounds of a less trained singers can be optimized based on automated adjustments processes.

Keywords: arts, automation, chained similarity analysis

# Bevezetés

A Photoshop kapcsán már régóta ismert (vö. <https://helpx.adobe.com/hu/photoshop/using/content-aware-fill.html>), hogy a rendszer képes egy adott fotórészlet (pl. emberalak előtérből való) kitörlése után a potenciális hátteret a környező képpontok alapján rekonstruálni (mely rekonstrukció sikeressége objektíven visszamérhető, ha a kitörölt képrészt fizikailag is eltávolítva új fotót készítünk – vö. tartalomérzékeny kitöltés).

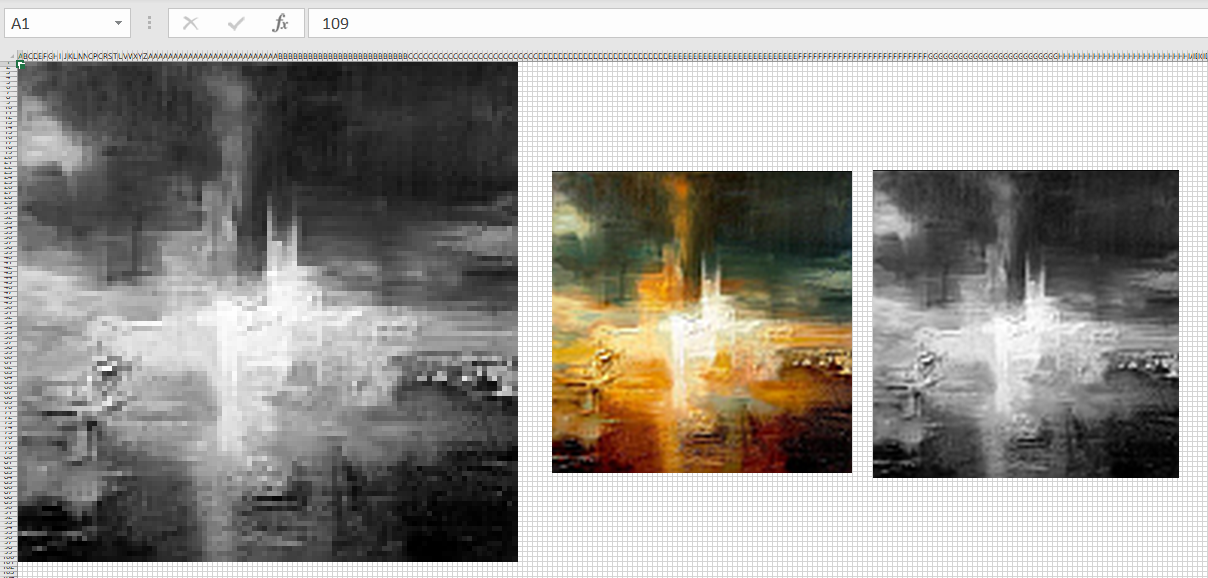
A Photoshop szoftver tehát arra világos példát jelent, hogy a képpontok között összefüggések vélelmezhetők – méghozzá context-free összefüggések, hiszen a Photoshop sem tudja, hogy adott pixel egy falevél vagy egy felhő része-e, s nem azért alakítja olyanná a hátteret, amilyenné, mert érti, mit ábrázol a kép, hanem azért, mert valamiféle mintázatot mer vélelmezni a létező képpontok alapján, mely mintázat kiterjeszthető a következő hiányzó képpontokra is lépésről lépésre. Sőt, már ez esetben is végtelen az iterációk száma, ha a már valahogyan becsült képpont RGB/szürkeárnyalat-kódja(i) felhasználásra kerülnek a következő hiányzó pixel becsléséhez – hiszen a sok-sok hiányzó képpont egymásutáni bejárásának lehetősége önmagában végesen nagy, de bármikor bármilyen sorrend szerint újrakezdhető lenne (vö. végtelen – vö. <https://miau.my-x.hu/miau/262/excel_plus.xls>).

További előzmények:

* Robotszem: <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=eye>
* Gyanúgenerálás: <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=suspicion>
* …

# A kiindulási helyzet

## A valós kép



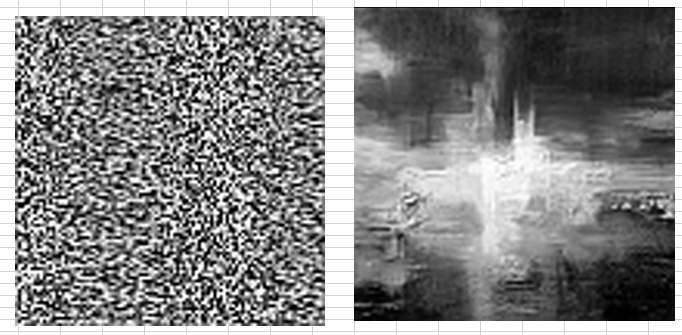
1. ábra: Az eredeti színes és szürkeárnyalatos kép (középen és jobbra) feltételes formázással rekonstruált szám-alapú kontroll-nézete (balra) – forrás: saját ábrázolás

További részletek: <https://miau.my-x.hu/miau/270/rgb_grey.xlsm>

A kísérlethez két képre van szükségünk. Válasszunk egy valódi, nonfiguratív alkotást: <https://www.google.com/imgres?imgurl=https://i.pinimg.com/originals/4d/ff/71/4dff7170bc7f18084e0d27f77601ef57.jpg&imgrefurl=https://www.pinterest.com/hasnarozana/painting/&h=480&w=368&tbnid=_t7Ir0fc95mHNM&tbnh=256&tbnw=197&usg=AI4_-kTg3V3EHbqzEe40PP07NBXIb-1Qtg&vet=1&docid=K042IN3GOu37RM&itg=1&hl=en-HU> (vö. 1. ábra középen, ill. szürke-árnyalatos verzióban jobb).

## A véletlen benchmark

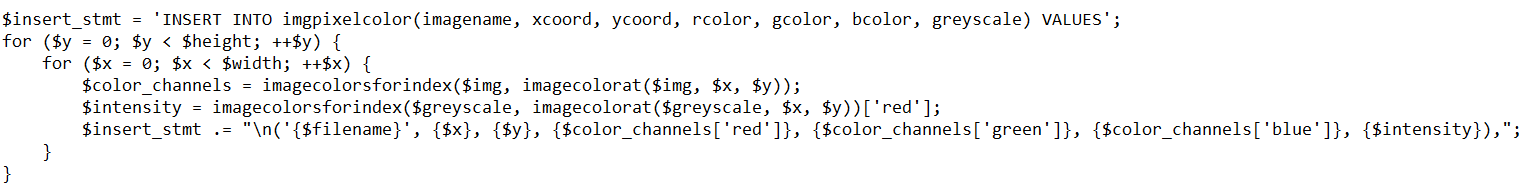
A másik kép legyen eleve szürkeárnyalatos és keletkezzen 100\*100 pixel felbontásban eleve, ahol minden képpont helyére a véletlenszám-generátor 0-255 között számot generál, s ezen mátrix az Excel-ben feltételes formázás keretében fekete és fehér közötti átmenetet = szürkeárnyalatot kap: vö.



2. Ábra: RND (balra) és szándék (jobbra) – forrás: saját ábrázolás

## Képből adat

Képből adatot (RGB-paramétereket) például a PHP beépített függvényei segítéségével lehet kinyerni (vö. kódrészlet alább):



Vagyis az 1. ábra baloldali ábrája nem más, mint a fenti konvertálás nyomán előállt RGB-kód, ill. az ebből nyert szürkeárnyalatos verzió maga (vö. <https://www.tutorialspoint.com/dip/grayscale_to_rgb_conversion.htm>).

# A robot-esztéta működése

A robot-esztéta lényege, hogy minden egyes képpontot megkísérel levezetni a rendelkezésre álló egyéb képpontok alapján. Operatívan ez nem pontonként, hanem oszloponként, ill. soronként történik. A pont-alapú becslés is lehetséges (vö. fentebb excel\_plus.xls, de itt és most a blokkosított eljárás lesz bemutatva).

## Alapbeállítások

A 100\*100 képpont alapján 100-100 modell kialakítása szükséges, vagyis elsődlegesen az 1-99 oszlopok alapján kerül becslésre a 100. képoszlop szürkeárnyalat-kódja

## Transzponált állapotok

A transzponált nézet nem más, mint a sorok és oszlopok megcserélése utáni új 100 modell, mely érdekében (hasonlóan az alapbeállításokhoz) ismét csak 9 csoportba kerültek besorolásra az eredeti (0-255) szürke-kódok.

## A modell

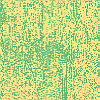
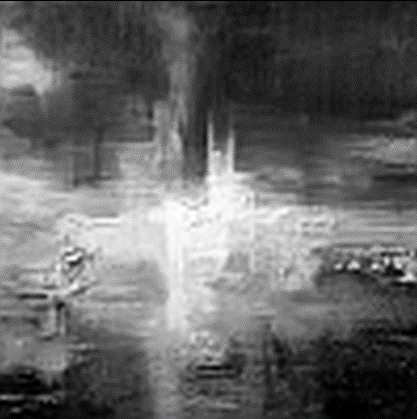
A mindenkori 100\*99 inputadat (melyek 1 és 9 közötti értékeket tartalmaztak) a COCO MCM (<https://miau.my-x.hu/myx-free/>) keretében ad vissza „fékezett habzású” polinomokat, vagyis nem minden nyers adat kaphat csereértéket az irányítatlan lépcsős függvényben, hanem csak az 1—9 csoportokra létezik csereérték, s ezek összege a 99 attribútum mentén adja ki a keresett képpont becsült értékét.

# Eredmények

Összesen készült tehát 100 oszlop-becslő MCM-modell a valós képre + 1 RND-modell a benchmarkképre, ill. 1+1 transzponált modell valós képpont-blokkra és rnd-képpontblokkra. A keresett skála kapcsán feltárásra kerültek az alábbi tények:

* Az RND-kép direkt és transzponált maximális, minimális hibája 3 nagyságrenddel nagyobb volt, mint a valós kép esetén.
* Az RND-kép esetén az 1—9 kódtartományban tehát lehetséges volt kódtévesztés, míg ez a jelenség a valós kép esetén egyetlen egyszer sem lépett fel (vö. 100\*100 pont esetén egyszer sem).
* A valós kép tény-becslés korrelációi 1.00 értéket adtak, míg az RND-kép esetén ez a korrelációs érték 0.98 volt direkt és transzponált állapotban is.
* Az RND-jelleg és a valós-jelleg tehát egymástól „szignifikánsan” elválik, ami alapját adja annak a keresett skálának, melyben a 0 hiba jelzi, ha egy „valós” kép koncepciózus, konzisztens, tehát SZÉP, míg a hiba végtelensége jelzi a véletlenséget, vagyis a szépség ellentettjét.

A robot-esztéta képpont-gyanút megfogalmazó képessége kapcsán az alábbi vizuális élmények álltak elő:

3. ábra: A pixeles és a foltos gyanú-nézet (jobbra, balra) az eredeti képhez képest (forrás: saját ábrázolás)

A 3. ábra alapján látható, hol leginkább piros a riasztási (gyanú) nézet (balra), ill. hol a legkevésbé zöld a foltnézet (jobbra). Különösen a foltnézet alapján kijelenthető, hogy a kép közepe a leginkább konzisztens (vagyis az üzenettér a leginkább tervezett), míg a kép konstrukciós gyengeségére több halványpiros folt utal ennek környezetében. Ezen gyanús foltok (pixelek) művészi újragondolása mindenkor az alkotó feladata lesz majd a jövőben, ha egy-egy művész akar ilyen típusú támogatást igénybe venni.

A 4. ábra(pár) felső része az egyes (pixel-széles) képoszlopok magyarázhatóságának inverzét (vagyis a képoszlopok hibáinak négyzetösszegét) mutatja, ahol a piros vonal a tízpontos mozgóátlag. A kép/festmény konstrukciók erősségét mutatja, hogy a magyarázhatóság zavarai a kép közepén a legkisebbek, vagy a centrális fókuszt a festő rel. jól garantálta.

A 4. ábra(pár) alsó része az egyes (pixel-széles) képsorok magyarázhatóságának inverzét (vagyis a képsorok hibáinak négyzetösszegét) mutatja, ahol a piros vonal ismét csak a tízpontos mozgóátlag. A kép/festmény konstrukciók erősségét mutatja, hogy a magyarázhatóság zavarai a kép közepén a legkisebbek, vagy a centrális fókuszt a festő rel. jól garantálta – jobban, mint oszlopirányból vizsgálva.

4. ábra: Magyarázhatósági gyanú-hisztogramok (forrás: saját ábrázolás – 0-100 sorok, ill. oszlopok azonosítója, 0-1000 hibanégyzetösszegek dimenzió nélküli számok)

5. ábra: Az átlagos magyarázó erők (oszlopok) hisztogramjai (forrás: saját ábrázolás – Y-tengely átlagos %-os magyarázóerők, melyek összege 0-100 attribútumra/oszlopra = 100%)

Az 5. ábra a 100 modell átlagában mutatja meg az egyes képoszlopok hozzájárulását a többi képoszlop szürkeárnyalatának numerikus levezethetőségéhez. Az egyes magyarázóerők alapvetően kiegyensúlyozottak (a százalékos magyarázóerők szórása rel. kicsi), vagyis a kép/festmény konstrukciója rel. stabil. A kiugró/maximum-érték pedig a kép közepéhez közel helyezkedik el, ami a 4. ábrapár centrális fókuszra vonatkozó vélelmét erősíti.

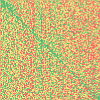
# A robotszem specialitásai

Ha az RGB kódok alapján várjuk el a robotszemtől, hogy megértse az szürkeárnyalat-kód (lineáris) számítását, ez természetesen lehetséges (vö. 6. ábra):

6. ábra: Az RGB-kódok konverziós szabályának megsejtése – forrás: saját ábrázolás

A 6. ábra a 100\*100 pontból véletlenszerűen kiválasztott 500 RGB-kód (X1-2-3) és Y-ként a szürkeárnyalat kódját tartalmazó OAM COCO-STD-alapú megértésének hátterét mutatja be, vagyis a lineáris transzformáció felismerésének részletei láthatók. A lineáris transzformáció kapcsán a 3 komponens lineáris hatásmechanizmusa nem teljesen tiszta, de jól felismerhető az 500 input-triplet és az Y értékek alapján.

# A szomszédosság elve

7. ábra: A magyarázó erők oszloponkénti megoszlása – forrás: saját ábrázolás

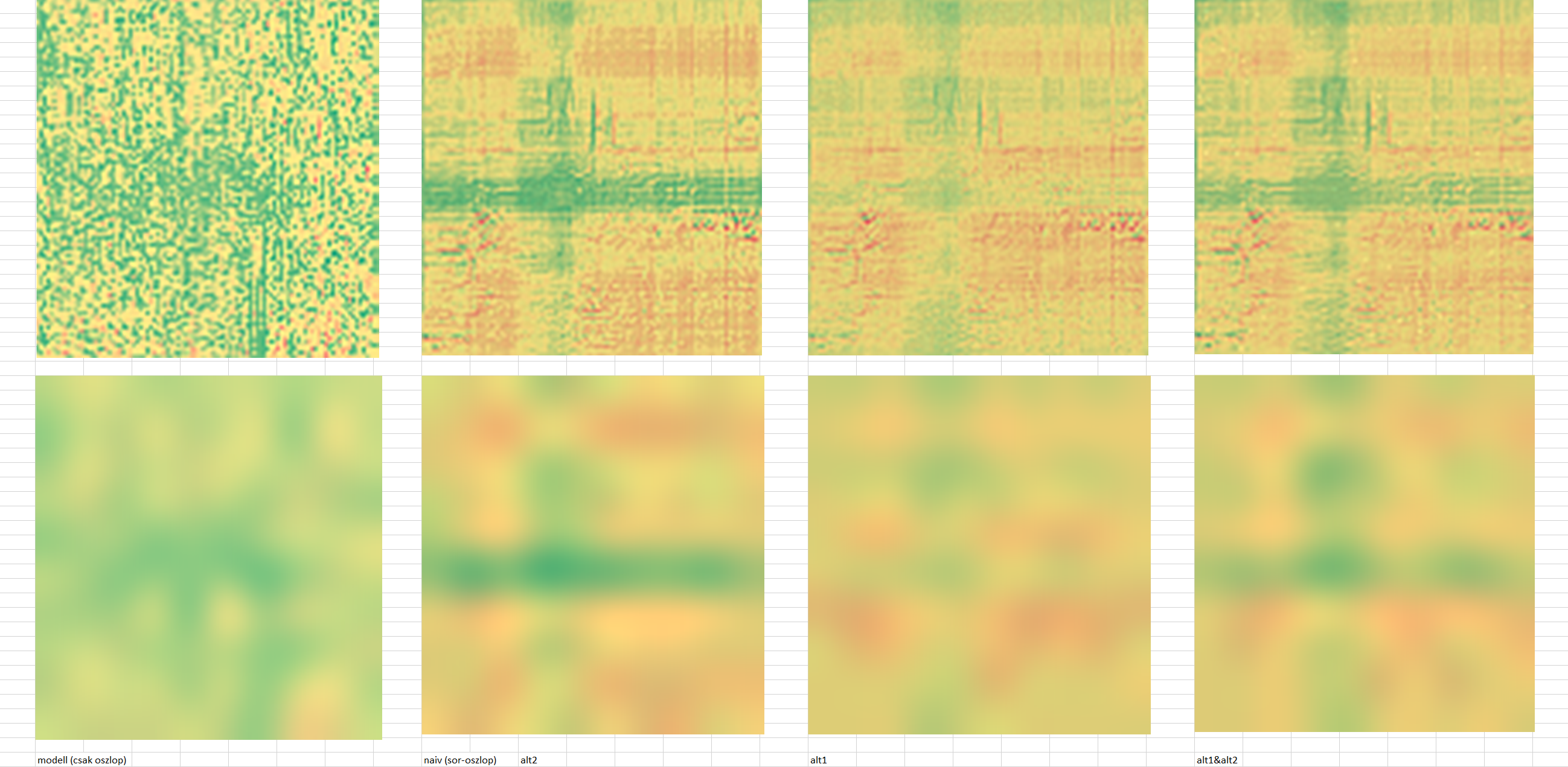
A 7. ábra óvatosan jelzi, hogy a fenti oszlopirányú modellezés kapcsán alapvetően elvárható, hogy bármely oszlop értéke a szomszédos (előtte, utána) lévő oszlopok által jobban magyarázható legyen, mint a távolabbi képpontokkal – hasonlóan, mint ahogy a pixelesség előáll egy alacsonyabb felbontású videóban.

Az átló maga a tény=becslés állapot, de a 7. ábra jobboldali képe az átló vékony vonalánál egy erősebb zöld átlós sáv létét engedi felismerni: vagyis a szomszédosság elve felismerhető. (Ahogy felismerhető kell, hogy legyen majd a transzponált esetekben is.) Ez a sáv is jelzi a contex-free modellezés logikus voltát!

Ha az átlódomináns képet alapképnek = festménynek tekintjük, akkor ennek minél kevésbé random jellege irányába való elmozdulás a képalkotás koncepciózusságának mértékeként értelmezhető.

Részletek: <https://miau.my-x.hu/miau/270/rgb_grey_naiv.xlsm>

A naiv modellek



8. ábra: Az optimalizált és a naiv modellek riasztási nézetei - forrás: saját ábrázolás

A 8. ábra baloldali képpárja jelzi a 100\*100-as és a 10\*10 felbontású optimalizált riasztási képet. A további három állapotpár a naiv riasztásokat mutatja be (alternatíva1 és alternatíva2, ill. ezek eredője formájában). A naiv megoldások a mindenkor balról szomszédos, tehát azonos sorú egyetlen kép szürke-kódját módosítja a max, min és átlag értékek eltérő verzióival. Az optimalizált és a naiv megoldások közös jegye a centrális fókusz (zöld dominancia). A naiv megoldások számításigénye töredéke a modellezés számításigényének.

# Konklúziók

* Az elvárt skála megalkotható!
* A gyanúgenerálás = inspiráció-generátor megalkotható!
* A képközép elvárt idealitás-többlete levezethető!
* A naiv és optimalizált megoldások jelentősen eltérnek, de a centrális fókuszt mindkettő kimutatja.
* A context-free optimalizálás esetén igaz a szomszédság-dominancia elvárása exploratív modellezés esetén is (vö. COCO MCM, mely semmilyen restrikciót nem használ a modell-paraméterek kapcsán, csak a hibaminimalizálást és a 100 pontról 9 pontcsoportra szűkítést.
* A jövőkép többszörösen meghaladja a fentebb bemutatott potenciált…

# Jövőkép

A fentebb bemutatott oszlop-irányú képszeletelés 100 modellel elvégezhető sor-irányú (transzponált modellsorozattal is, mely eredménye a 3-4-5. ábrák alternatíváiként értelmezendő. A két alternatív gyanúkép eredője egy harmadik 100\*100-as kép lesz. Minden gyanúkép önmagában is lehet a művész számára inspiráló.

Amennyiben az RGB adatok is feldolgozásra kerülnek, akkor külön az R, külön a G és külön a B rétegeknek lesznek 3-4-5. ábrái és ennek transzponált nézetei, sőt ezek eredőnézetei.

A szürke+piros+zöld+kék nézetek 3-3 alternatívája immár 12 gyanúnézetet generál, ill. egy 13. nézet ezek eredője (az eredők eredője).

A keresett és megtalált skála mentén elsődlegesen tehát minden képről statisztikai adatok gyárthatók, melyek alapján n darab önálló festmény (kép) közül a minden kép másként egyformán szép-e kérdés is megválaszolható az ismert antidiszkriminatív úton. A szóba jöhető attribútumok itt is olyanok, melyek kapcsán az irány fogalma racionálisan értelmezhető: pl.

* Minél kisebb bármely nézet hibája (100\*100) pontra, annál ideálisabb/szebb a kép.
* Minél kisebb a mindenkori (nézetenkénti) attribútumok hibáinak szórása, annál ideálisabb/szebb a kép.
* Minél centrálisabb (karakteresebb) a képközepi hibaminimum, annál ideálisabb/szebben komponált a kép.
* …

Az itt bemutatott demo egy 2D-s, azaz síkvetületi potenciálként értelmezendő, de semmi akadálya nincs annak, hogy 3D-ben (vö. szobrászat), ill. 4D-ben (vö. mozgókép/valóság) is hasonló harmóniakereső értelmezések operacionalizálásra kerüljenek, hiszen a 2D-nek nincs érdemi specialitása az elemzési folyamat szempontjából…

A riasztási képek nem feltétlenül, de az átló-hatást mutató képek tekinthetők önálló festménynek, s ezek minél erőteljesebb szabályszerűsége (szépsége) új szimulációs lehetőségeket nyit meg – végtelenített folyamatként… (vö. HPC-projektek)…

# Referenciák

…szövegközben…

# Mellékletek

A szürkeárnyalatos kép képszerkesztő szerinti és Excel-riport-alapú hisztogramjai:

