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?!

Robotesztéta II.

(Robot-Aesthete II.)

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

Kivonat: A robotesztéta elméleti bemutatását kínáló I. rész után, ebben a cikkben egy valós Turing-teszt-folyamat kerül dokumentálásra. A véletlenszerűen választott 15 nonfiguratív (modern) képet véletlenszerűen választott 36 tesztelő személy értékelte 1<10-es szépségskálán színes és szürke verzióban külön-külön. Az elemzés első kérdése az volt, vajon képesek vagyunk-e hasonlóságelemzési alapon skálát építeni a véletlenszerű szépség-értékeléstől való képenkénti különbségtételre. Vagyis képesek vagyunk-e a szignifikancia fogalma nélkül kijelenteni, vajon melyik képről alkotott emberi vélemény mennyire különbözik a véletlentől? Melyik különbözik a legkevésbé? Melyek különböznek norma alatti módon, ill. melyik különbözik leginkább? Ezt követően levezetésre került egy fajta statisztikai aggregált szépség-index ismét csak hasonlóságelemzési, ill. naiv (nem optimalizált) alapon a kapott emberi szépség-pontok leíró statisztikái alapján. Mindezzel párhuzamosan létrehozásra került a legegyszerűbb robotszem, mely naiv módon abból indult ki, hogy egy adott pixel szürkeárnyalatos értékét az alapkép szóban forgó pixelt körülvevő pixeleinek szürkeárnyalatos adatainak átlaga racionálisan képes leírni. A becsült kép és a valós kép eltéréseiből számos statisztikai leíró mutatót számolva, hasonlóságelemzéssel létre jött az aggregált szépség a robotszem által érzékelve, ahol szép az, ami minél inkább mintázatszerű – jelen esetben naiv mintázatokat követő. Az eredmények alapján egyrészt kijelenthető, hogy 15-ből csak 3 képről illik humán szépség-ítéletet formálni az erősen szóró, vagyis véletlenszerűséghez közel álló nyers pontok és a bármilyen okból invaliddá váló modell-részeredmények kapcsán. Az elméleti részhez képest egyelőre még csak naiv mintázatokkal dolgozó robotszem és a humán értékítéletek egyetlen egy ponton közösek, vagyis egy képről nagy biztonsággal állítható, hogy rel. csúnya a többihez képest. A humán szépség-fogalom véletlenszerűsége kapcsán a Turing-teszt csak annyiban tekinthető sikeresnek, hogy az egy közös horgonypont létezését sikerült feltárni racionális/szabályelvű keretrendszeren belül.

Kulcsszavak: anti-diszkriminatív modellezés, optimalizálás, automatizálás, mesterséges látás, mesterséges intelligencia

Abstract: The I. part about the robot-aesthete presented the theoretical backgrounds concerning the mathematics of the beauty as such. This part (II) demonstrates a Turing-test-process, where 15 random-selected, but nonfigurative pictures got evaluated by a random set of human beings (by 36 people) on an artificial beauty-scale (1<10) in a separated evaluation process for coloured and grey-scaled versions. The analysis of the human evaluation should deliver a statement for the question: Is it possible to create a scale for measuring the randomness of the human answers concerning each picture? This quasi classic (statistic-oriented) challenge should be solved based on similarity analyses – based on the same AI-engine what will be used for further analyses concerning the beauty as such. This randomness-scale will let derive: which picture could be evaluated in the most robust way (it means: less random-like) and which picture generated the most random-like human evaluations. Therefore, we will also know, which pictures are under-norm, over-norm or even norm-like concerning the randomness of the human evaluation. All these will be derived without using the logic of significant differences. Based on the raw human evaluation, a kind of aggregated (similarity-oriented) beauty-index will also be calculated. Parallel, a naïve robot-eye will also be constructed being just capable of estimating grey-scale-values for each pixel based on the neighbourhood (it means: on 3\*3 pixels). The estimated pixels and the real pixels lead to differences, where the more beautiful is a picture, the less is the volume of the differences (describe through a lot of statistical attributes). The results: 3 of the analysed pictures could be evaluated as valid objects because of random-like human evaluations and/or modelling anomalies. There is one single picture where the naïve robot-eye and the human beauty-conceptions lead to the same conclusion: this single picture is an under-norm-object concerning the pattern-oriented beauty-definition. Therefore, it is not a trivial successful Turing-test – but the success is more depending on the human beings as the naïve robot-solution.

Keywords: anti-discriminative modelling, optimizing, automation, artificial seeing, artificial intelligence

# Bevezetés

Előzmények:

* Robotszem-cikkek: <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=eye>
* Robot-esztéta I: <https://miau.my-x.hu/miau/270/roboteszteta.docx>

Az emberi szépség-fogalom megteremtése, különösen nonfiguratív, azaz tartalmilag kontextus-mentesítésre törekvő képek esetén olyan kihívás elé állítja az emberi agyat, mely kihívás kapcsán azonnal felmerül annak a kérdésnek a kényszerű vizsgálata, vajon beszélhetünk-e adott képhalmaz esetén egyáltalán emberi (közös) mintázatokról, vagy a képek sok emberi szem/agy általi szépként való értelmezése inkább felel meg a véletlenszerű találgatás jelenségét jellemző leíró statisztikáknak.

A véletlenszerűségtől való különbözni tudás az alapfeltétele annak, hogy egy potenciális Turing-teszt esetén létezzen olyan benchmark, mely az emberit, mint olyat képes számszerűen jellemezni.

A véletlentől való különbözés látszólag a statisztikai próbák világába tartozó jelenségkör, de itt és most arra láthatunk példát, hogy az a robotszem/robotagy (MI), ami eleve a szépség fogalmát kell, hogy operacionalizálja, képes arra is, hogy skálát alkosson az egyes képek emberi értelmezésének véletlentől való különbözését illetően. Ha a szépség matematikája értelmében szép az, ami bármilyen komplex módon, de mintázat/szabály-szerű, akkor véletlenszerű és egyben csúnya mindaz, ami ennek ellentettje.

# Alapadatok

Ahhoz, hogy hasonlóságelemzés-alapú aggregált skálák, mesterséges fogalmak kerülhessenek kialakítása, elkerülhetetlen objektumok és attribútumok definiálása, melyek kapcsolatai, hasonlóságai optimalizált értelmezés keretében vezetnek az emberi elemző szubjektivitás kontextus-független kizárásához.

Párhuzamosan a mindenkori OAM-ok alapján naiv, vagyis az attribútumonként sorszámozott objektum-rangsorok összege, ill. átlaga szolgál benchmarkként a nem optimalizált adatelemzés egyik képviselőjeként.

További, a következő fejezetek reprodukálhatóságát támogató részletek az alábbi Excel-állományban érthetők el: <https://miau.my-x.hu/miau/270/roboteszteta_naiv.xlsx> / <https://miau.my-x.hu/miau/270/roboteszteta_human.xlsx>

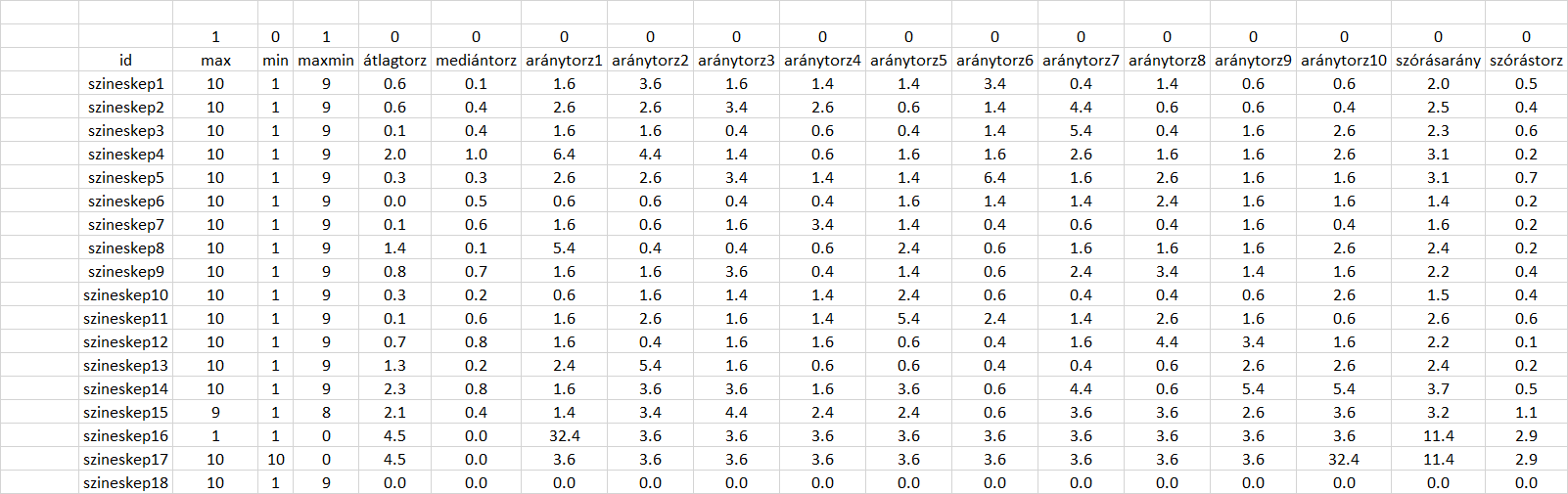
## Az véletlenszerűségi skála OAM-ja

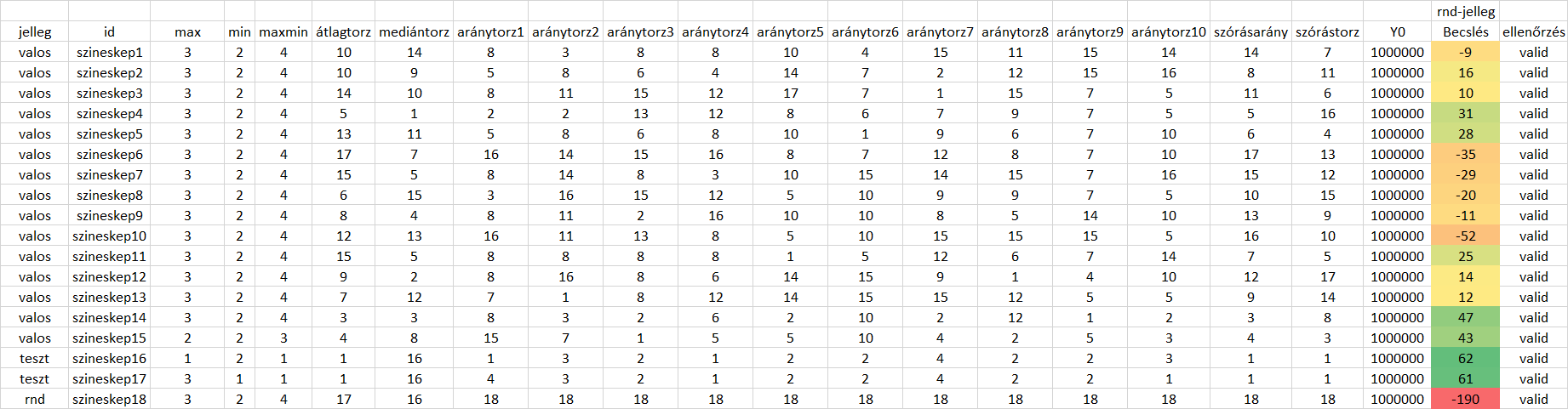
A véletlenszerűen választott 15 nonfiguratív (modern) képet ismét csak véletlenszerűen választott 36 tesztelő személy értékelte 1<10-es szépségskálán színes és szürke verzióban külön-külön. Mivel szemmel látható volt, hogy ugyanazon kép kapott 1-es és 10-es értéket is, azonnal felmerült a kérdés: tekinthető-e az emberi véleményeket leíró (15+15)\*36 cellányi nyersadat egy potenciális véletlenszám-generátor termékének? Illetve: melyik kép esetén mennyire véletlenszerű a kapott emberi értékítéletek halmaza?

Objektumok tehát mindenkor a képek (15 db).

Az attribútumokat (17 darab) az alábbi lista tartalmazza:

* Max: minél kisebb a legnagyobb humán szépség pont, annál jobban eltér a véletlentől a válaszsor
* Min: minél nagyobb a legkisebb humán szépség pont, annál jobban eltér a véletlentől a válaszsor
* Maxmin: minél kisebb a legnagyobb és a legkisebb humán szépség pont közötti távolság, annál jobban eltér a véletlentől a válaszsor
* Átlagtorz: Minél nagyobb az abszolút értékben kifejezett eltérése a válaszok átlagának 5.5-től, annál jobban eltér a véletlentől a válaszsor
* Mediántorz: Minél jobban eltér a medián az átlagtól abszolút értékben, annál jobban eltér a véletlentől a válaszsor
* aránytorz1-2-3-4-5-6-7-8-9-10: minél nagyobb az 1-2-3-4-5-6-7-8-9-10-es értékítélet abszolút eltérése a 36/10-es szinttől, annál jobban eltér a véletlentől a válaszsor
* szórásarány: minél nagyobb az arány1-2-3-4-5-6-7-8-9-10 egymástól való eltérése (szórása), annál jobban eltér a véletlentől a válaszsor
* szórástorz: minél nagyobb az abszolút eltérés a 2.87-es benchmarktól az összes képenkénti humán válasz esetében, annál jobban eltér a véletlentől a válaszsor

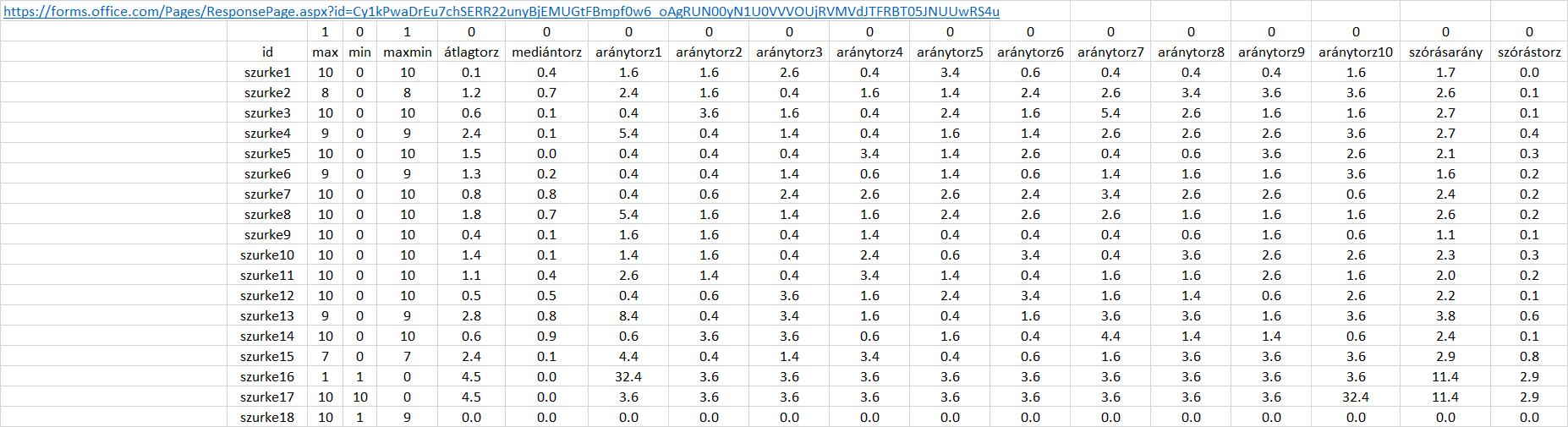


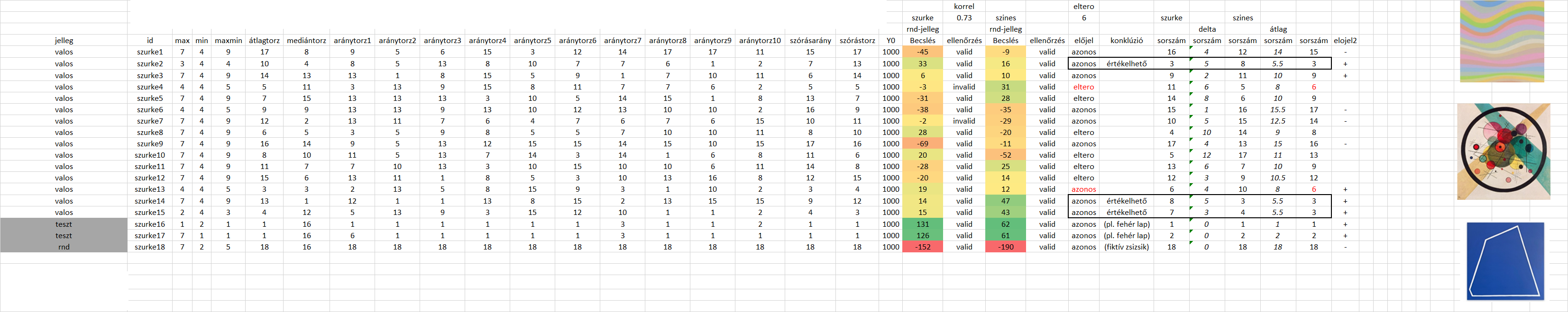


1. ábra: A színes képek véletlentől való eltérésének nyers és sorszámozott OAM-ja (forrás: saját ábrázolás)

A színesképek és a szürkeképek kapcsán a 16. és a 17. objektum a monoton 10 és a monoton 1 fiktív értékelés lenyomata. A 18. objektum a teljesen/elméletileg véletlen pontokat reprezentáló objektum – benchmark.

Az elvárásoknak megfelelően a színes és a szürke képhalmaz esetén is a 18. objektum véletlentől való eltérése a legkisebb, ill. a monoton 1-es és 10-es értékelések adják a véletlentől való eltérés maximumát.





1. ábra: A szürkeárnyalatos képek véletlentől való eltérésének nyers és sorszámozott OAM-ja (forrás: saját ábrázolás)

A szürkeárnyalatos OAM 2 invalid objektumot produkált, míg a színes objektumok mindegyike valid (szimmetrikus) lépcsősfüggvényre vezetett.

A színes és a szürkeárnyalatos képek norma-alattisága vagy norma-felettisége (negatív/pozitív színes hátterű becslések) lehet egyidejű (előjel=azonos), vagy aszinkron (előjel=eltérő). 6 darab kép esetén (egyik sem szürke-invalid) nem volt az értékítélet azonos irányultságú. A színek információértékét a 6/15 arány fejezi ki.

A 15 objektumból az invaliditás és az előjel-eltérések kapcsán (részleges átfedéssel) 7 objektum esik ki. 3 további objektum véletlentől való eltérése volt valid/norma-alatti, így ezek is kizárásra kerültek az értelmezésből. Maradt tehát 5 objektum, amiből 3 volt azonos értékűnek tűnően véletlentől különböző: vagyis a 2., ill. a 14. és 15. objektumok voltak kellően karakterisztikusan értékelve a humán pontozások hatalmas eltérései alapján.

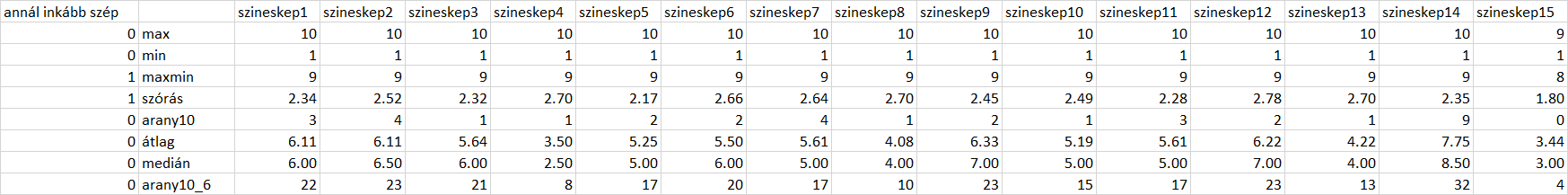
## Az emberi szépség-skála OAM-ja

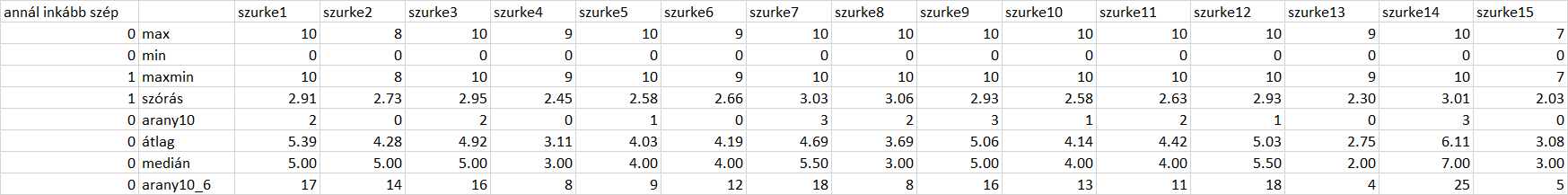
A korábbi 15 valódi kép (objektum) a szépség-számítás sor-fejléce. Az attribútumok () az alábbiak voltak:

* max (0): minél nagyobb a kapott humán értékelés maximuma, annál szebb a kép
* min (0): minél nagyobb a kapott humán értékelés minimuma, annál szebb a kép
* maxmin (1): minél kisebb a max-min távolság, annál szebb a kép
* szórás (1): minél kisebb a szórás, annál szebb a kép
* arany10 (0): minél nagyobb a maximális pontszám részaránya, annál szebb a kép
* átlag (0): minél nagyobb a pontszám-átlag, annál szebb a kép
* medián (0): minél nagyobb a medián, annál szebb a kép
* arany10\_6 (0): az arány1-2-3-4-5-6 rétegek összege legyen minél nagyobb annál szebb a kép

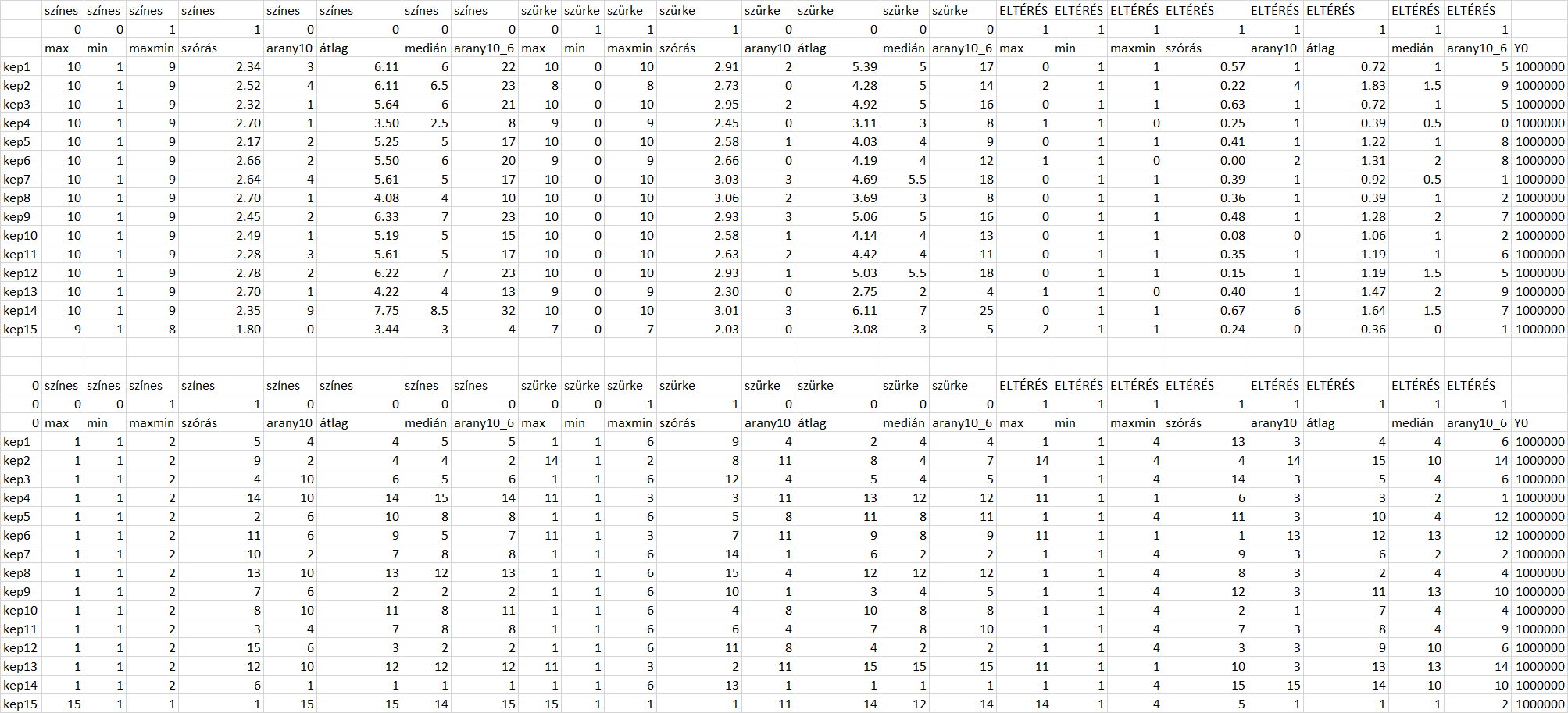
A minimális érték monotonná vált mind a színes, mind a szürkeárnyalatos rétegben.

A származtatott attribútumok a fenti színes ÉS szürkeárnyalatos attribútumok mellett ezek abszolút eltéréseként kerültek értelmezésre – minél kisebb bármilyen eltérés, annál szebb a kép.

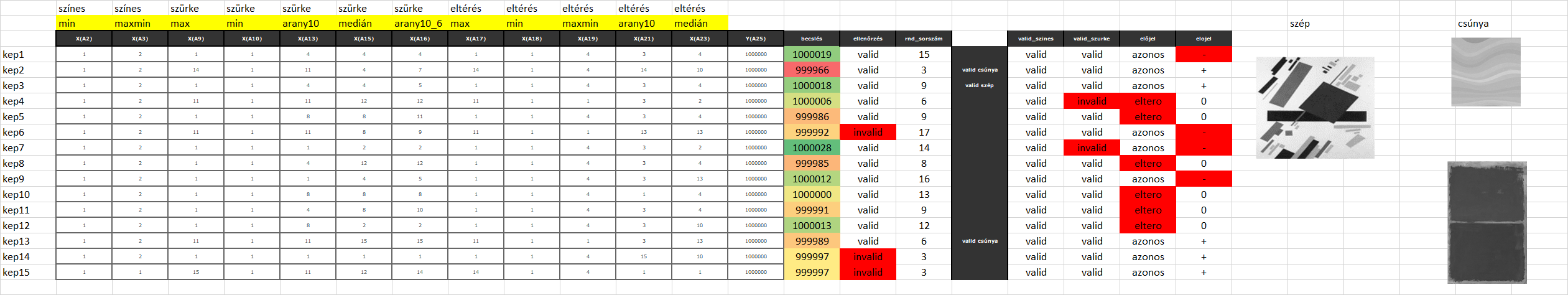




1. ábra: A színes és szürkeárnyalatos objektumok szépségének nyers leíró adatai (forrás: saját ábrázolás)



1. ábra: A színes és szürkeárnyalatos attribútumok önálló és összevont nézete (forrás: saját ábrázolás)



1. ábra: A szépség-számítás második modell-rétege (forrás: saját ábrázolás)

Az 5. ábra alapján látható, hogy az első modellrétegben rendelkezésre álló 3\*8 attribútumból 12 maradt a második körre, hiszen a többi attribútum együttállása a minden kép másként egyforma elvet kielégítő eredményre vezetett, de a 12 fel nem használt attribútum információértékét kötelező ilyenkor kiaknázni.

A szépség-becslés 3 invalid objektumra mutat rá. A véletlenszerűség kizáró elvei és a szépség-invaliditás eredője, hogy 3 objektum marad értékelhető: 2 rel. csúnya (norma-alatti – valid csúnya) és 1 rel. szép (norma-feletti – valid szép).

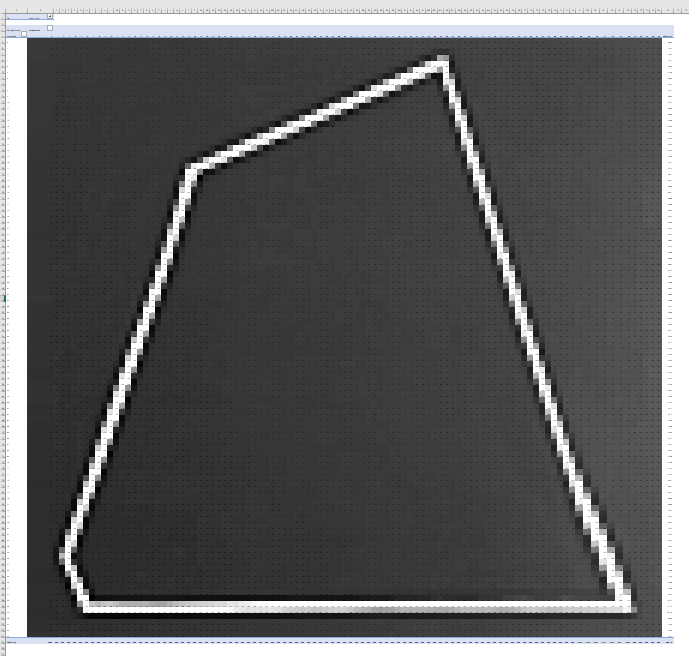
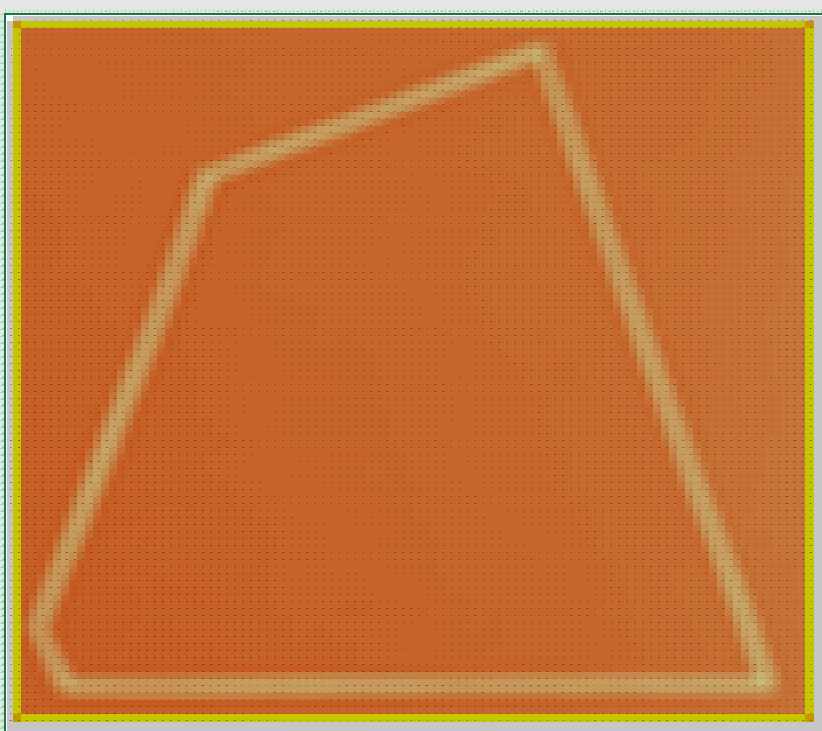
A különböző invaliditások egyidejű kezelése a konzisztencia-képzés egy speciális esete: konzisztens az, ami nem mutat fel semmilyen ellentmondást.

## A robotszem általi szépség-skála OAM-ja

A robotszem célja, hogy úm. megértse a valós kép pontjai közötti összefüggéseket, s ez a megértés nagyon rugalmas matematikai apparátus alapján illik, hogy leképezésre kerüljön. A robotesztéta I. bemutatta a rugalmasság maximumát közelítő COCO MCM eljárást és a naiv megoldások közül is 2 eltérőt annak érdekében, hogy láthatóvá váljon az, hogy a naiv megoldások rel. kevés képpont alapján (pl. sorok, oszlopok, ill. ezek eredője) vezetnek el oda, hogy egy képpont szürke árnyalatos értéke becslésre kerülhessen.

Az itt használt naiv megoldás a mindenkori képpont körüli 3\*3-es minimátrix átlagát vette becslésként – szembe a robotesztéta sor/oszlop/együttes logikájával. Ezek a naiv beállítások nem igényelnek indoklást. Nincs miért, mert a válasz mindenkor az, hogy csak azért, mert éppen így sikerült kialakítani adott pillanatban egy megoldást.

A véletlenszerűen naiv becslés azt demonstrálja, ahogy a zeneileg, vizuális képzetlen átlagember érzékeli a modern zenét és/vagy a modern festészetet. Ha néhány egyszerű séma nem vezet értelmezhető eredményre (függőleges=oszlop-alapú, vízszintes=sor-alapú, merőleges=sor&oszlop, középpontos = 3\*3-as, stb.) – akkor az már frusztráló, s ami frusztráló, az csúnya, mert energia-igényes a megértési folyamata.

1. ábra: A nyers (szimulált), a 3\*3-becslést alapján előálló kép és ezek különbözetei - forrás: saját ábrázolás

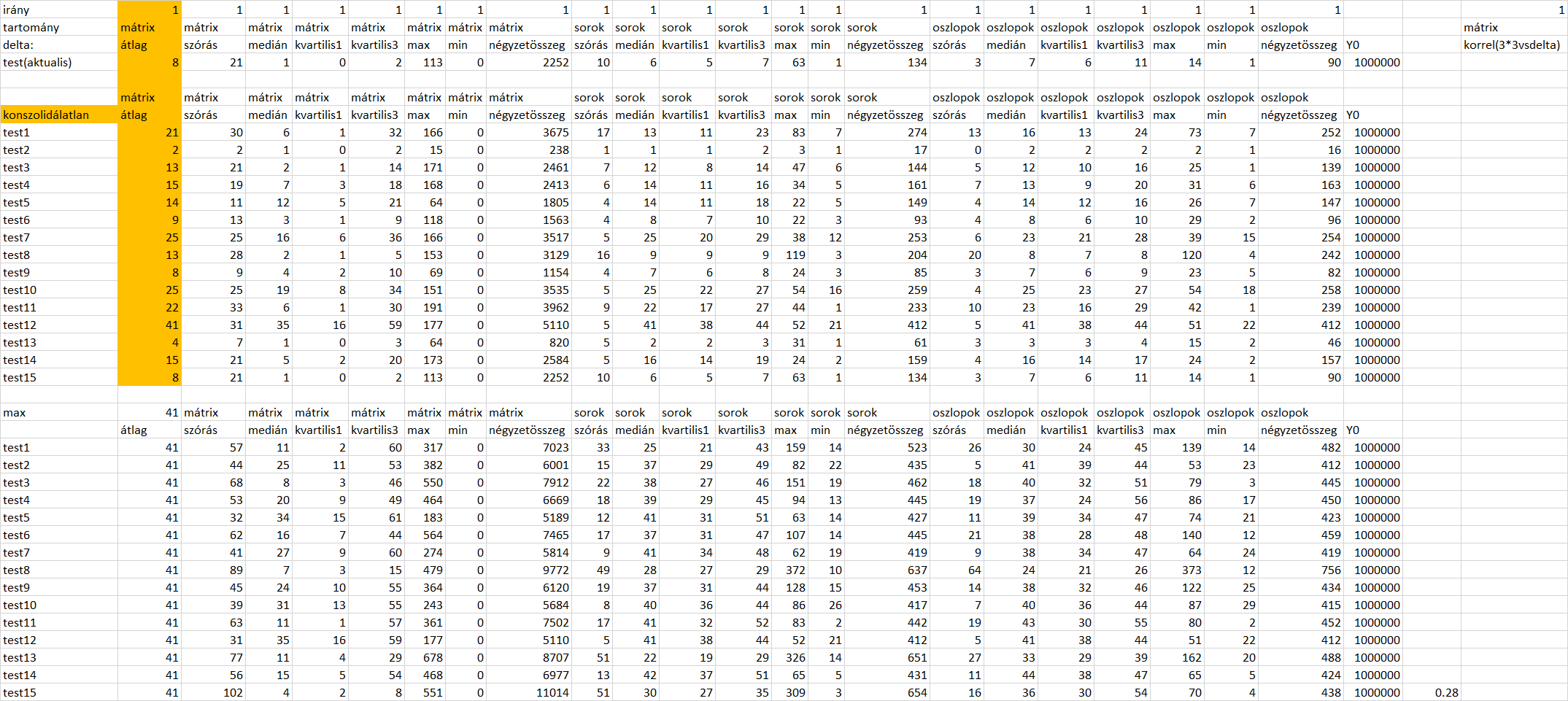
A COCO MCM ezzel szemben a teljes kép „együttrezgését” használja fel a képpontcsoportok érték-rétegeinek és rétegösszegeinek becslésére, mint ahogy Beethoven érzékeli a zenét a cipészinashoz képest. Jelenleg azonban a robotizált cipészinas szépség-ideálja és a leginkább antidiszkriminatív/robosztus humán értékelés kerül összevetésre.

A 6. ábra világosan jelzi, hogy a 3\*3-as leképezés nem tud elszakadni az alapmotívumtól, szemben a robotesztéta I (COCO MCM) megoldásának relatíve magas homogenitásától…

Akár cipészinas, akár udvari zeneszerző áll a háttérben, a végső adathalmazt a tényleges kép szürke árnyalatos kódjai és a becsült kép szürkeárnyalatos kódjai közötti eltérések statisztikai mutatói adják, melyek mindegyikének iránya (1) a minél kisebb, annál szebb elvet követi:

* mátrix-szint
  + (átlag)
  + szórás
  + kvartilis1
  + kvartilis3
  + medián
  + max
  + min
  + négyzetösszege
* sor-összegek szintje
  + (átlag)
  + szórás
  + kvartilis1
  + kvartilis3
  + medián
  + max
  + min
  + négyzetösszege
* oszlop-összegek szintje
  + (átlag)
  + szórás
  + kvartilis1
  + kvartilis3
  + medián
  + max
  + min
  + négyzetösszege

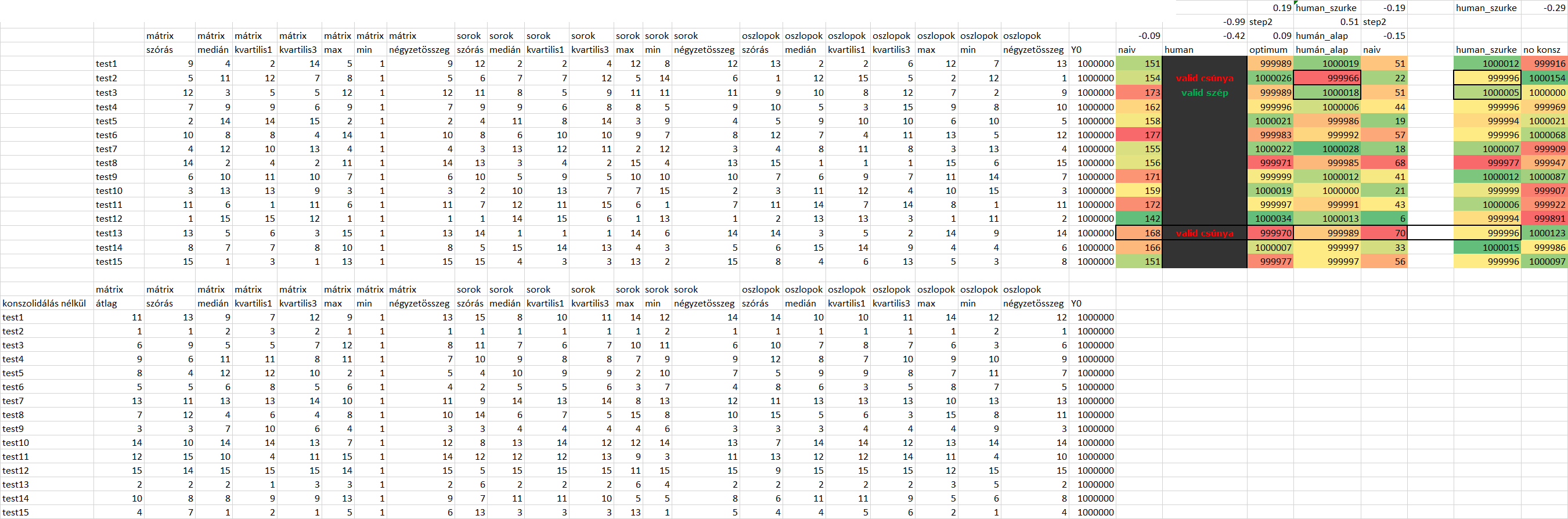
Értelemszerűen a mátrix-átlag, mely konszolidációs célra kerül eleve felhasználásra, hogy a szürkeárnyalatos kódok átlagmennyiségének azonossága mellett arányosan torzítsuk az egyes képek egyéb statisztikáit, a sor-összegek átlaga és az oszlop-összegek átlaga azonos objektum-rangsort kell, hogy eredményezzen, így az átlagok a konszolidáció mellett végső soron csak ellenőrzési célt szolgáltak (vö. 7. ábra):



1. Ábra: A 3\*3-as robotszem szépség-OAM-ja nyers (fent) és konszolidált (lent) formában (forrás: saját számítások), ahol a konszolidálás lényege a mátrix-átlag arányos egalizálása…

A 8. ábrán felül a jobb oldalon látható a nem konszolidált alapokon nyugvó szépség-becslés és a step2 oszlop-fejléc alatt (vö. 9. ábra) a konszolidált elemzés eredménye (az első minden-másként-egyforma-részmegoldást követő második réteg formájában). Ezek kapcsolata: -0.29-es korrelációt mutat, ami azt jelzi, hogy a konszolidáció alapjaiban érinti a jel-értelmezést, a szépség fogalmát.

A humán-alap-oszlop az 5. ábra becsléseit mutatja újra: vagyis ezek a zöld-sárga-piros jelek a szépség-jelek a véletlentől sok esetben jelképesen különböző emberi szavazatok alapján.



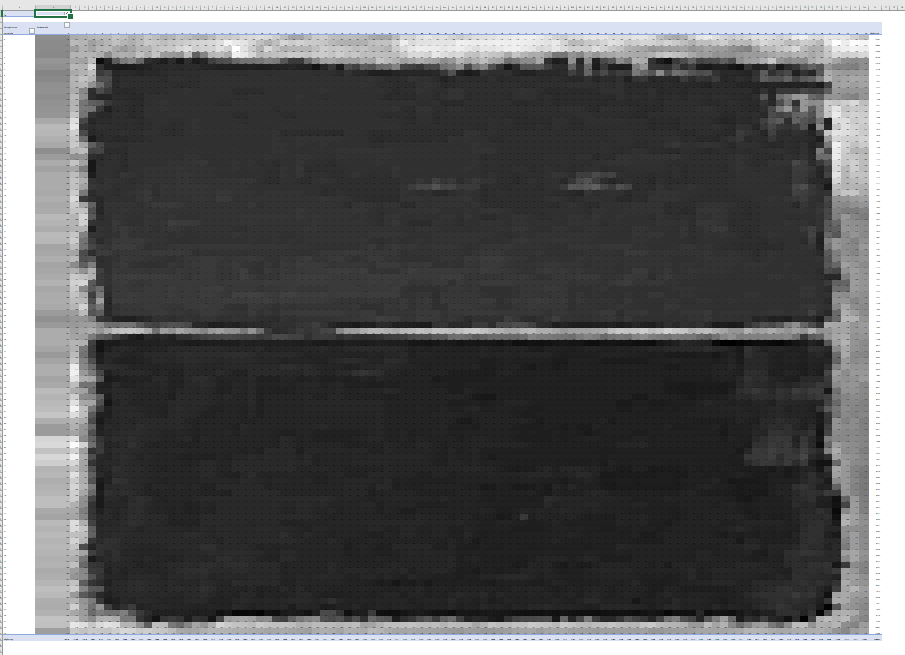
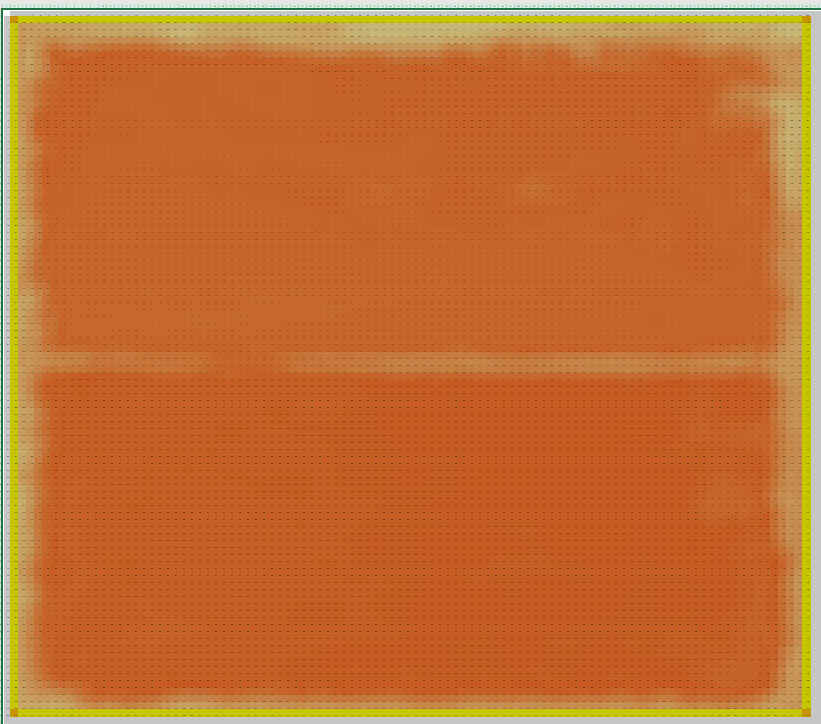
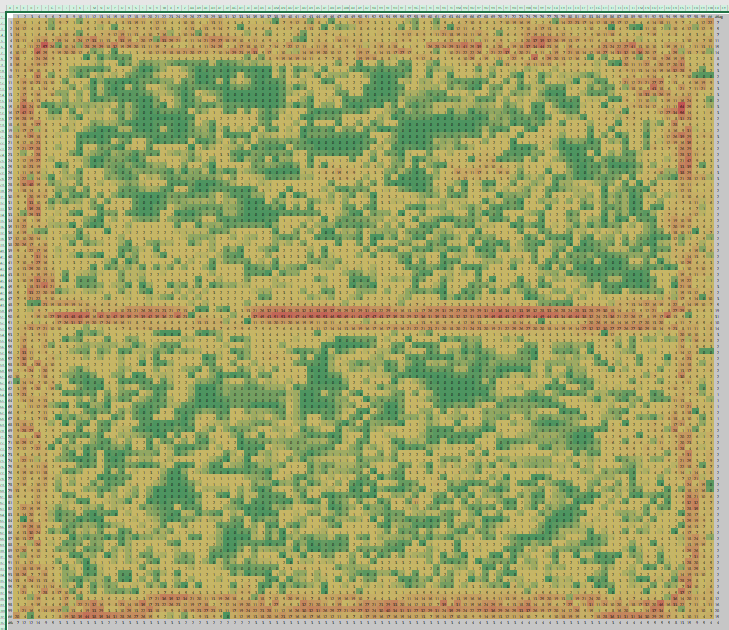
1. Ábra: A konszolidált sorszámozott OAM és eredményei (felül), ill. a konszolidálatlan OAM (alul) – forrás: saját ábrázolás

A 8. ábra világosan jelzi, hogy a test13 (a 13. kép) esetében van „csak” a valid csúnyaság kapcsán konzisztens eredmény, melyek a humán szavazatok „csak szürke” (human\_szurke) rétege is megerősít, ami fontos, mert a 3\*3-as naiv becslési logikát követő robotszem csak szürkeárnyalatos jelekkel dolgozott egyelőre.

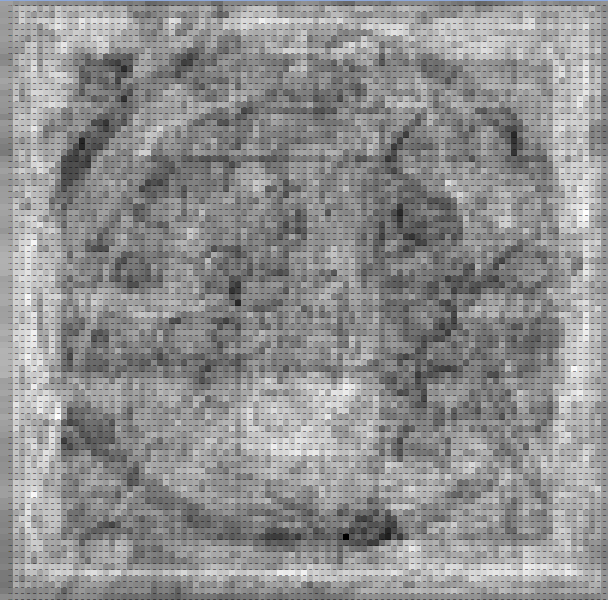
# Konklúziók

A végeredmény alapján a Turing-teszt sikere megkérdőjelezhető, melynek oka egyelőre inkább a humán oldalon, mint sem a robot-oldalon (a robot-naivitás oldalán) keresendő.

Az egy közös, konzisztens objektum fellelése látszólag kevés, de az, hogy van olyan objektum egyáltalán, mely képes volt minden ellentmondás-potenciált kizárni, az önmagában is eredmény – hiszen így az összehasonlító elemzés nem torkollik a nem-tudom-rendszerválaszba.

   (kép13)

A valid rel. csúnya ábra tehát a 13. kép. Sajnos nincs valid szép konzisztens konklúzió ebben a mintában, hogy a tény-becslés-eltéréseket a naiv emberi szemek is megkísérelhessék értelmezni…

 (kép1-15+tesztkép robotesztéta-I)

A kutatásnak egy melléktermékeként egy fajta vizuális titkosírás operacionalizált lehetősége is felismerésre került:

A fenti kép értelmében számos kép (pl. 100\*100) pl. szürkeárnyalatos kódjainak eredője egy új kép, ami megfelelő rétegbontással egy teljesen dekódolhatatlan titkosírás, hiszen az eredőkép, mint kép válik olvashatóvá az emberi agy számára, s nincs is benne karakter, mint olyan – a felületen mégis létre tud jönni a betű élménye... Így a rétegeket tartalmazó, ebben is csak a szürke árnyalat-kódokat X-Y-koordinátákhoz megadó az adatbázis akár egy teljesen kamu (szöveg/)szám-halmaz is lehet, ahol a szöveg (betű) pl. ASCII-kódja a számot helyettesítő jel. Ha megvan a kódolandó üzenet, akkor akárhány rétegre bontani véletlenszámokkal is lehet, mert a szétbontás irányába nem kell, hogy értelmű legyen, csak az összevonás irányába kell az ellenőrzést végrehajtani – s egy adott üzenet n-féle módon kódolható, ahol n quasi végtelen, így nincs esély (részlegesen) azonos tartalmak ismétlődéseinek felismerésére sem...

Az eredő kép lehet új képsor része is, vagyis n-rétegben lehet több forrásból, csatornán töredékesen érkező értelmetlen képi részleteket egy helyen is összeilleszteni…

A kapott kép maga lehet más módon kódolt üzenet, azaz a grafikus burok bármilyen tartalmat és technológiát képes integrálni.

A kapott kép nem egésze kell, hogy az üzenet legyen.

A közölt adatbázisok egy része lehet feldolgozást vezérlő paraméter is.

Vagyis a felismert lehetőség kombinatorikai tere quasi végtelen a hibridizáció kapcsán is.

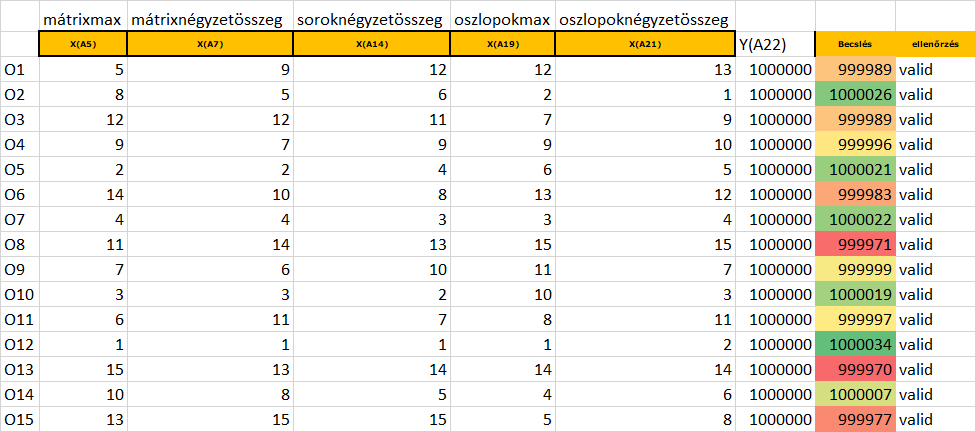
A több réteg úgy is előálítható, hogy a maximális rétegszámon belül eleve több eredménytermék (kép-üzenet) áll elő, akár a főüzenet alternatíváiként, hogy azt se lehessen tudni az összes megfejtése esetén sem, melyik a valódi üzenet. Sőt, az alternatív üzenetek betűiből speciális válogatással állhat elő a végső üzenet (vö. hibridizáció).

A titkosírás rétegei számos szereplőnél is elhelyezhetők, akik csak együtt képesek a végső tartalmat kikényszeríteni. A rétegek (eltérő szereplőknél, helyeken őrizve) tartalmazhatnak felesleges rétegeket, sőt termináló rétegeket, melyek megjelenése a teljes dekódolási folyamatot megállítják – ahol a rétegek sorrendje is fontos lehet…

# Források

…szövegközben…

# Mellékletek



1. Ábra: A step2 mögötti aktív attribútumok OAM-ja – forrás: saját ábrázolás

A 9. ábra alapján látható, hogy a tényleges és becsült képpontok közötti érdemi információt a maximumok és a négyzetösszegek hordozzák, ami megerősíti a konszolidáció szükségszerűségét, vagyis az átlagos jelerősség standardizálásának fontosságát.