

Vass Gergely

HDRI

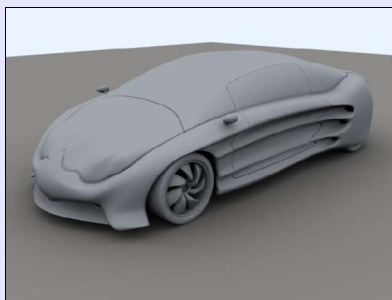
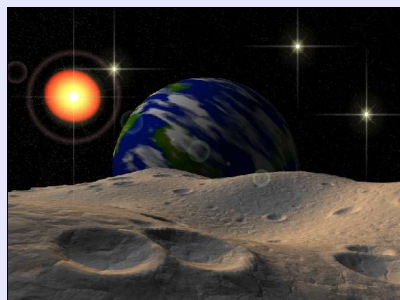
Divat vagy forradalom?

5-6 évvel ezelőtt szinte minden számítógéppel előállított 3D-s képen láthattunk „lens flare” effektet, azaz a lencsék becsillanásának – általában igen gyenge – szintetikus utánezatát. Természetesen divatok jönnek, mennek, így a lens flare csillaga is leáldozott, és elfoglalta méltó helyét a ritkábban használatos eszközök sorában. A globális illuminációs renderelő programok elterjedésével – melyek a szórt, felületekről visszaverődő fényeket is képesek kezelni – minden Internetes galéria megtelt a szürke síkon álló szürke modellekkel, melyeket egy homogén szürke „égbolt”, vagy „dóm-fény” világított meg. A divatokkal persze nincsen semmi baj, hiszen mindig hoznak valami újat, valami frisset, ami némi idő elteltével szervesen be is épül a számítógépes grafika világába.

Mára úgy tűnik a szürke dóm-fény éra lassan elmúlik. Napjainkban a netre felkerülő képek nagy részén ugyanis észrevehetjük, hogy az alkotó nem egy homogén égbolttal világít, hanem az Internetről letölthető egyik, ún. HDR képet használ a modell megvilágítására. A

legtöbb ifjú 3D-s titán a számítógépes grafika nagy tudósai közül szinte csak Paul Debevec nevét ismeri, aki a HDR képek gyakorlati alkalmazásának egyik első kidolgozója. A HDRI technológia szele azonban nemcsak a hobbistákat, hanem a professzionális stúdiókat is

megcsapta. Lehet, hogy nem is egy szimpla divatról van szó? Cikkünkben körüljárjuk, hogy mi is az a HDRI, mire jó, mire nem jó és hogy tényleg hozhat-e alapvető változásokat a 3D grafikusok mindennapjaiban?



Divatok a 3D világban: a lens flare effekt, a dómfényvel megvilágított majd a HDR képpel megvilágított modellek. (Holger Schömann, Barta Gergely, Barta Gergely)

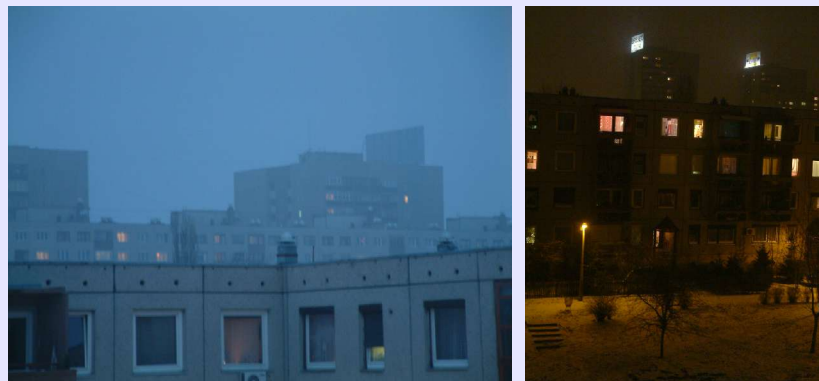
Először is: mi a fene az a HDR vagy HDRI? A betűszó feloldása a következő: High Dynamic Range Image, azaz „nagy dinamikájú kép”. Éppen ezért ne azt mondjuk, hogy „HDRI kép” hiszen annak jelentése „nagy dinamikájú kép kép”. De mit is jelent ez a nagy dinamika? Képek esetében dinamika alatt azt értjük, hogy az eltárolható lehetséges legsötétebb és legvilágosabb szín (ill. annak világossága) mennyire nagy tartományt „fog közre”, pontosabban a legvilágosabb fényesség mennyivel világosabb mint a legsötétebb. Borús napon egy homokbányáról készült kép alacsony dinamikájú (LDR), míg egy sötétebb szobában készült kép, ahol az ablakokon át a tűző napot látjuk, nagy dinamikájú lesz (HDR). Az első esetben nincs nagy gond, hiszen az alacsony dinamikájú képeket probléma nélkül el tudjuk tárolni, fel tudjuk dolgozni és meg tudjuk jeleníteni a mindennapos eszközökkel is. A nagy dinamikájú képekkel azonban meggyűlik a bajunk!

A számábrázolásról

A képpontok színét meghatározó számok – általában három (r,g,b) – ábrázolása a számítógépes világban digitálisan történik. Ezeknek az értékeknek fix, előre megadott helyen el kell férniük, ami sajnos azt is jelenti, hogy nem tudunk tetszőleges számokat ábrázolni. Hogy pontosan milyen értékeket tudunk eltárolni, azt alapvetően két dolog határozza meg:

- mekkora helyen tároljuk a számokat (pl. 8 bit, 10 bit, 16 bit, 32 bit);
- hogyan tároljuk a számokat (pozitív egészek, előjeles egészek, lebegő pontos ábrázolás stb.)
- Adott számábrázolás mellett két korlátba ütközhetünk a fényességértékek tárolásánál:
- előre adott az eltárolható maximális és minimális érték;
- adott az ábrázolás „finomsága”, azaz a szomszédos számok közötti különbség.

A manapság használatos képek szinte kivétel nélkül csatornánként 8

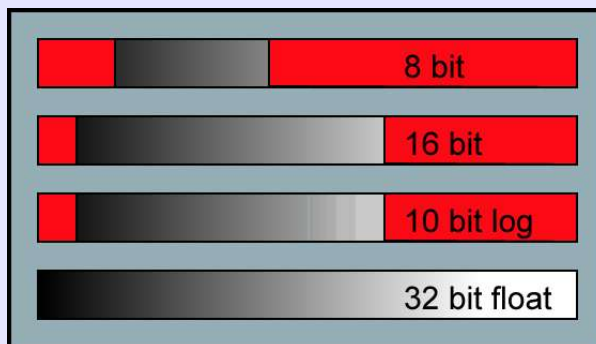


A szürkületkor készült kép alacsony dinamikájú, az este készült verzió a látható fényforrások miatt magas dinamikájú.

bit helyen kerülnek eltárolásra, ahol természetesen az intenzitást pozitív értékekkel jellemezzük. Ez pedig azt jelenti, hogy csatornánként 0 és 255 közötti egész számokat tudunk csak eltárolni. Persze arról, hogy pl. a 156-os érték pontosan milyen fényességet jelent, még nem szóltunk. Logikusnak látszik a kép legvilágosabb pontját 255-el jellemezni, a feketét pedig 0-val. Vagy inkább a kép legsötétebb pontja legyen a 0? És mi van, ha egy sötét képet akarunk eltárolni? Akkor is a maximális értékkel jellemezzük a legvilágosabb, de amúgy sötét pontot? Ha nem, akkor mi legyen a 255-nek megfelelő „referencia” fehér? Ha meghatároztuk a 0-hoz és 255-höz tartozó fényességet, akkor hogyan kapjuk meg a 127-nek megfelelő világosságot? Valamilyen egyenletes, vagy éppen nem lineáris átmenetet

használnunk a sötét és a világos között? Ezekre a kérdésekre nem lehet és nem is kell egyértelmű választ adni, hiszen – elvileg – minden képnek és megjelenítő eszköznek beállítható a kontrasztja, fényessége, gammája stb. melyek meghatározzák a kód – fényesség megfeleltetést.

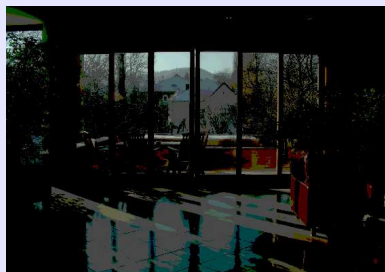
Akkor hol a gond? Miért ne tehetném meg, hogy a bevezetőben említett nagy dinamikájú képet 8 bites számokkal tárolom el? Csupán annyit kell beállítanom, hogy a 255-ös szám legyen a vakító napfény, a 0-ás kód legyen a sötét szoba árnyékos sarkának legsötétebb pontja! Sajnos ez a megoldás – mint azt az egyik ábra is mutatja – nem járható! A 256 különböző lehetséges érték túl kevés a kép korrekt eltárolásához, és lényegesen különböző pixelek is



A különböző dinamikatartományok viszonya (szemléltető jelleggel)



Különböző expozíciós idővel készült felvételeken a nagy dinamikatartomány más-más szeletét láthatjuk. A HDR képek képesek mindezt eltárolni.



Ha a nagy dinamikájú, csonkítatlan képet 8 bites formátumba elmentjük, valami ilyesmi, borzasztó képet kapunk.

ugyanolyan színeképet kapunk. Mit lehet hát tenni? A válasz kézenfekvő: ki kell választani azt a világosság-tartományt, mely a számunkra fontos (abb) információkat tartalmazza, és az ennél világosabb ill. sötétebb értékeket le kell vágni. Gyakorlatilag fényképezésnél is ugyanezt csináljuk – manuálisan vagy az automatika segítségével – mikor a megfelelő záridőt és blendét megválasztjuk. A túl világos részek, melyek „kilógnak” a választott dinamika-tartományból kiégnek, azaz a lehető legvilágosabb kódot kapják, a túl sötét részletek pedig feketék lesznek. Így kétségtelenül eldobunk sok információt, ami később, a kép további feldolgozása során bizony jól jöhetne.

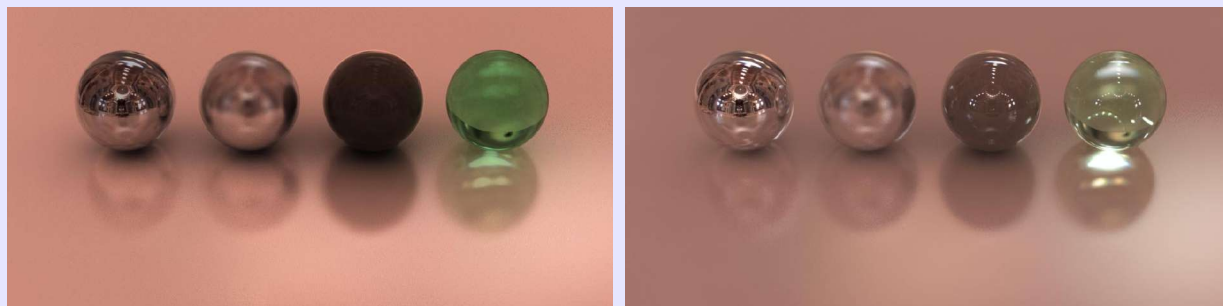
Nagyobb dinamika

Jogos lehet a kérdés, hogy miért nem merült fel előbb az olyan képformátumok iránt az igény, melyek jóval nagyobb dinamikával rendelkeznek? Miért elégszünk meg a – nem túl sok – 256 féle különböző fényességgel? A válasz egyszerű: az input és output eszközök döntő többsége alacsony dinamikájú. A monitorok, a projektorok, az átlagos nyomtatók sajnos nem képesek 256-nál többféle intenzitás visszaadására. Bár elsőre úgy gondolhatjuk, hogy erre az átlagembereknek nincs is szüksége, de gondoljunk csak a monitorok kapcsán a mai számítógépes játékokra! Mennyivel élvezetesebbek lennének a kalandok, ha a sötét, komor barlangok, sejtelmes folyosók után a nyílt téren játszódó részek sokkal izzóbbak, világosabbak, élőbbek lennének. A 2003-as Siggraph kiállításon San Diegoban már be is mutatták a nagy dinamikájú monitor prototípusát, és kétségtelenül szebb képet adott mint egy hagyományos monitor.

Professzionális alkalmazásoknál természetesen már korábban is léteztek nagyobb dinamikájú formátumok, de ezek nem terjedtek el a 3D grafikusok körében. A jobb digitális fényképezőgépek például

képesek un. „raw”, azaz nyers formátumba menteni, mely általában 16 bites képet jelent. Ezek a képek sokkal több információt tartalmaznak mint 8 bites testvéreik, hiszen csatornánként 65536 féle világosság tárolható! A gyakorlatban az ilyen fotók számítógépes feldolgozás első lépése a file konvertálása 8 bitre úgy, hogy csak az értékes intenzitás-tartomány essen a 8 bites „ablakba”. Ez lehetőséget biztosít arra, hogy utólag manuálisan korrigáljuk az alul- ill. felülexponált képeket és ne hagyatkozzunk a fényképezőgép automatikájára. A 16 bites képek szerkesztése, feldolgozása, retusálása még nem igazán egyszerű, hiszen nemcsak a korrekt megjelenítés probléma, hanem – a szerző tapasztalata szerint – a mai programok sincsenek erre felkészítve. Meg kell jegyezni azonban, hogy drámai fejlődésre számíthatunk ezen a téren.

Filmes körökben szintén kevésnek bizonyult a 8 bit, itt ugyanis – a valamivel nagyobb dinamikájú – celluloid szalag a forrás és a kimeneti hordozó. A levilágító berendezések már képesek 256-nál többféle intenzitást is előállítani, így az utómunkát végzők a monitoron csak „nagyából” azt látják amit a nézők majd a moziban. A 16 bites képek már képesek a filmszalagon rögzíthető információ eltárolására, de ezeknél a viszonylag nagy méret miatt egy kicsit trükkösebb megoldás terjedt el. Az emberi látás sajátosságai miatt sötétebb színeknél érzékenyebbek vagyunk az intenzitás változására, mint a nagyon világos árnyalatoknál. Éppen ezért a filmes utómunkában használt un. logaritmikus formátumok nem a fényességgel arányos skálát használnak, hanem a filmszalag fényáteresztő képességét jellemző számot tárolják el. Ez azt eredményezi, hogy a sötétebb színek pontosabban kerülnek tárolásra mint a világosak, és így elegendő a 10 bites tárolás is.



A bal képen alacsony, a jobb képen magas dinamikájú képet használtunk a tükröződéshez és a világításhoz. A különbség jól látható. (Kép: Barta Gergely)

HDRI

Ha léteznek már 16 bites formátumok, melyek igen sok, 65536 különböző fényességet képesek tárolni, miért akarna bárki is ennél nagyobb dinamikájú képformátumot? Gyakorlatilag ma még nem nagyon ismerünk olyan output eszközt, mely képes lenne ilyen képet megjeleníteni. De ne felejtjük el, a 3D grafikában képeket nem csupán pusztán megjelenítésre szokás használni, hanem sok egyéb funkciójuk lehet: használhatjuk őket mint textúra, tükröződésben megjeleníthetjük vagy éppen világíthatunk is vele. Az utóbbi két esetben a kép elsődleges funkciója nem a néző gyönyörködtetése, hanem fizikailag pontos számításokhoz adatok szolgáltatása a virtuális világ fényviszonyairól.

A legegyszerűbb példa a többé-kevésbé tükröző golyó esete nagy dinamikájú környezetben. Hogy miért is jelent ez gondot a gyakorlatban? Vegyünk például egy szobát, ahol az ablakon keresztül besüt a nap, de a plafonon is ég egy 60W-os izzó. Ha erről a környezetről készítünk egy hagyományos, alacsony dinamikájú fotót, akkor mind a nap, mind az izzó fénye ki fog égni, azaz ugyan olyan, maximális fényerő kerül eltárolásra. (Sajnos ezt nem nagyon tudjuk elkerülni, legfeljebb csak úgy, hogy a sötétebb részeket eldobjuk, akár az alulexponált fényképeknél.)

Ha ezt az alacsony dinamikájú képet használjuk a tükröződések megjelenítésére úgy, hogy a tükröző felület tökéletes tükör, nincsen semmi baj. Ha viszont a modell felülete csak kevésbé tükröz és elnyeli a fény egy részét, rögtön lebukunk: mind a nap, mind a sokkal gyengébb izzó fénye ugyanúgy fog csökkenni, gyengülni a tükröződésben. A valóságban persze a nap tükröződését még bőven kellene látni, mikor a lámpa tükröződése már nem látható.

Talán nem igényel különösebb magyarázatot, hogy ha világításra használjuk a képet – azaz a virtuális világot megvilágító, különböző irányokból beeső fényt egy textúrából olvassuk ki – mégfontosabb, hogy a valóságos fényerő (radiancia) kerüljön eltárolásra és ne egy csonkított érték. A nap fénye dimenziókkal erősebb mint egy izzócska fénye, így a kép kiszámításánál fontos, hogy ez az információ ne vesszen el.

Milyen számábrázolást használjunk hát valódi fényviszonyok tárolására?



Ingyenesen letölthető HDRI gömbtükrös fotók a www.debevec.org-ról.



Felső sor: hagyományos LDR képek esetében a fényerő utólag nehezen állítható.

Alsó sor: HDR képeknél bőven van tartalék a sötét és a világos részeken is. (Mindig a középső az eredeti kép)

A 65536 különböző érték nem tűnik elegendőnek, hiszen nem ritkán a legvilágosabb pixel többszázszerez (!!!) világosabb mint a legsötétebb pont. Éppen ezért egy jóval rugalmasabb és sokkal nagyobb tartományt lefogó ábrázolást terjedt el, a 32 bites, lebegőpontos ábrázolás. Ez azt jelenti, hogy nem csak iszonyúan nagy számokat vagyunk képesek tárolni, de a kicsi számokat is sokkal pontosabban. A HDRI formátumok nem megfoghatatlan, a pixelek világosságával valamilyen módon arányos értéket tárolnak (mint a 8 bites képek), hanem szigorúan az adott pixelre jellemző fény mennyiséggel arányos számot.

Azaz a nap színe tényleg akár sokmilliószor világosabb lesz mint a 60 wattos izzóé.

HDR képek előállítás

A 32 bites, nagy dinamikájú képek képesek szinte tetszőlegesen sötét és világos pixeleket eltárolni. De hogyan állítjuk elő ezeket a képeket? Nem létezik olyan (elérhető) fényképezőgép, mely a koromsötét szoba sarkát és ugyanakkor az ablakban látható napos ég részleteit egyidőben rögzíteni tudja. Az analóg gépekben használatos filmnek és a digitális CCD-nek is véges a dinamikai tartománya. 12 vagy 16 bites input képeknél jobbakra sajnos

nem nagyon számíthatunk! Jelenleg az egyik felkapott kutatási téma a nagy dinamikájú videó rögzítése, a 2003-as Siggraph konferencián pont erről szólt az egyik népszerű prezentáció. Az átlagos felhasználó számára elérhető legegyszerűbb megoldás Paul Debevec ingyenesen letölthető programja, a HDRShop használata.

Ez a programcska képes arra, hogy több, különböző expozíciós idővel és fix blendével készült képből generáljon egy, nagy dinamikájú képet. Hogy pontosan hány képre van szükség, azt két dolog is befolyásolja. Az egyik az, hogy az algoritmusnak ki kell találnia a film,



Balról jobbra: eredeti kép, alacsony dinamikájú elmosva, magas dinamikájú elmosva.

vagy a digitális kamera érzékenységi görbéjét (response curve vagy response function), amihez elvileg elegendő 2 kép, ha eléggé „változatosak” a pixel értékek. A gyakorlatban azonban ajánlott minimum 3-4 képet használni. A másik fontos tényező az elérni kívánt és az input képek eredeti dinamika tartományának nagysága. Pongyolán fogalmazva, ha a 1-től 1000-ig tartó dinamika tartományt szeretnénk elérni, de a fényképezőgépünk segítségével készült képek csak 100 „széles” tartományt fednek le, akkor 1000/100, azaz 10 különböző képet kell csinálnunk, hogy az egész tartományt lefedjük. Átfogalmazva: annyi fotót kell készítenünk, hogy ne legyen olyan pixel, mely az összes képen vagy kiégett vagy teljesen fekete maradt, hiszen ekkor az adott pixel eredeti intenzitása nem állítható vissza. Összefoglalva: érdemes minimum 4 képet készíteni, de nagyon nagy dinamika esetében sokkal több, akár 8-10 kép is szükséges lehet.

Sajnos a HDR képekkel is az a helyzet mint sok más adat-típussal, hogy nem létezik egy, mindenki által elfogadott és használt file formátum. Az egyik formátum a Radiance program formátuma, a .HDR file, de

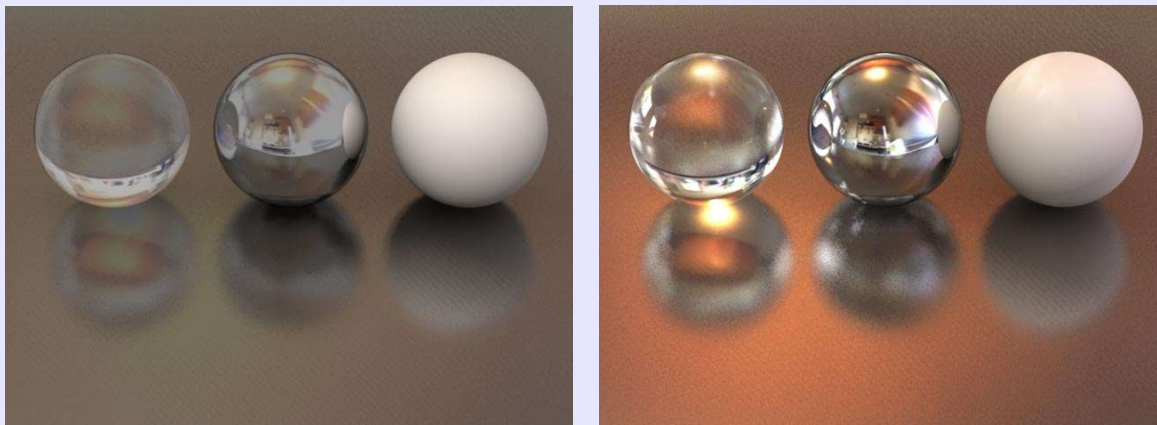
nagy dinamikájú képeket 16 bites vagy lebegőpontos TIFF formátumban is elmenthetünk. Szintén használatos az ún. „portable float map”, azaz a .PFM, valamint a nyers RAW formátum. Az egyik legnagyobb amerikai utómunka stúdió, az Industrial Light and Magic (ILM) jelenleg mindent megtesz, hogy saját nyílt formátumát, az openEXR-t is népszerűvé tegye, ami képes gyakorlatilag tetszőleges szám-ábrázolású képeket elmenteni.

Gyakorlatban

A HDR képek tehát, az LDR képekkel szemben, alkalmasak arra, hogy a valódi világ fényviszonyait a 3D-s, virtuális világban reprezentálja. Ez a képek korrekt renderelésénél feltétlenül szükséges. HDR képek készítésével könnyedén meg tudjuk örökíteni például filmek forgatásánál vagy termék-fotózásnál a fényviszonyokat, így később csak virtuálisan létező modelleket is le tudunk fotózni az eredeti környezetben. Ez valóban forradalmi újdonság, hiszen ezentúl nem kell (feltétlenül) hosszú órákat vacakolni a fények állítgatásával, hogy sikerüljön az eredeti fényviszonyokat utánozni. Két dolgot azonban ne feledjünk: csak nagyon ritkán ússzuk

meg, hogy ne kelljen kézzel világítani a 3D térben. Egy fiatal 3D grafikus nagy hátrányban lesz akkor, ha csak az Internetről letölthető HDRI állományokkal világítja meg modelljeit, hiszen nem tanulja meg a fények, árnyékok kezelését. Egy 3D-s képen legalább olyan fontos a világítás kidolgozása és minősége mint a modellé. A másik fontos dolog az, hogy HDR képeket azért nem olyan egyszerű csinálni. Ritkán lesz olyan szerencsénk, hogy az exponálások között a környezet semmit ne változzon, főleg ha szabad ég alatt vagy élő szereplőkkel dolgozunk.

De mire másra lehet jó még a HDRI technológia? Van értelme álló- vagy mozgóképeinket úgy elmenteni, hogy megőrizzük a nagy dinamika tartományt? A válasz: igen. Sok képfeldolgozásban gyakran használt szűrő, speciális effektus csak HDR képek esetén ad korrekt eredményt. Az egyik ilyen az igen gyakran használt elmosás, vagy blur, illetve a különböző defókusz szűrők. Gondoljuk végig mi történik, ha karácsonyi képeinket, melyek hagyományos LDR képek, a számítógépen utólag elmoszuk? A karácsonyfa izzói, és a szaloncukrok



LDR (bal) és HDR (jobb) kép használata a virtuális környezet reprezentálására. (Kép: Barta Gergely)

elmosás után szinte ugyanolyan világos foltok lesznek. A valóságban azonban az elmosott izzók nem halovány homályos pacák lesznek, hanem intenzíven izzó „foltokká” válnak, hiszen fényerejük nagyon nagy. (Lásd ábrán!) Az álló- vagy mozgóképek utólagos színekorekciója jó minőségben szintén csak úgy képzelhető el, ha a feldolgozott képek dinamika tartománya nagyobb mint a megjeleníthető. Ez azért fontos, mert ellenkező esetben a világosság enyhe növelése vagy csökkentése csúnya hibákat eredményezne. Gondoljunk végig mi történik egy LDR képpel, ha csökkentjük a brightness értékét? A kép minden pontja sötétebb lesz, vagyis az elvileg izzó napkorong is

beszürkül. HDR képek esetében bőven van tartalék, így a fényerő csökkentésével az izzó részek továbbra is izzani fognak, a fényerő növelésével pedig a sötét részekben újabb részletek „bukkanak” elő. (Lásd ábra!)

Mindezek mellett nem tudhatjuk biztosan, hogy a 32 bites HDR képek mennyire terjednek el a képfeldolgozásban vagy a DTP világában. Igazából az utóbb említett előnyei a nagy dinamikájú képeknek már a 16 bites képeknél is kihasználhatóak, ezekben is van annyi „nem megjeleníthető” tartalék, mely elegendő a hétköznapi feladatok korrekt megoldásában.

Persze annak ellenére, hogy a 16 bites képeket (még) nem szokták HDR képeknek hívni, ezek a file-ok is teljes joggal nevezhetők nagy dinamikájúaknak, és nyugodtan számíthatunk elterjedésükre. A legtöbb input eszköz, a szkennerek, a digitális fényképezőgépek és a renderelő programok is „belül” 12 vagy 16 (sőt néha még több) bitben gondolkodnak, így az elmentett 8 bites képek már csonkítottak. Biztos ami biztos, készüljünk fel arra, hogy a jövőben valószínűleg több információ lesz az általunk használt képekben, mint amennyit a monitoron látunk! És akkor majd jó lesz nekünk!