

Wprowadzenie do technologii HDR

Krzysztof Gdawiec



UNIWERSYTET ŚLĄSKI
INSTYTUT INFORMATYKI

Grafika komputerowa wykorzystywana m.in. w filmach, grach, symulacjach czy zastosowaniach wojskowych wymaga coraz bardziej realistycznych obrazów wyświetlanych na ekranie komputera.

Jest to szczególnie trudne biorąc pod uwagę obrazy świata naturalnego, które przedstawiają naszemu systemowi wizualnemu szeroką gamę kolorów i intensywności.

Gwieździsta noc ma średni poziom luminancji ok. 10^{-3} cd/m², a dla dziennej sceny jest to poziom ok. 10^6 cd/m². Ludzie mogą obserwować detale w obszarach, które są pomiędzy 1 a 10^4 cd/m².

Poza przemysłem filmowym przez wiele lat nie było aż tak wielkiego nacisku, aby osiągnąć większy zakres dynamiczny.

Urządzenia do wyświetlania ograniczają zakres tego, co można przedstawić do około dwóch rzędów wielkości pomiędzy minimalną a maksymalną luminancją.

Monitory CRT osiągają luminancję ok. 100 cd/m^2 , zaś monitory LCD $250\text{--}500 \text{ cd/m}^2$, co nawet nie zbliża się do poziomu światła dziennego.

Obecnie wszystko to zmienia się dzięki obrazom HDR (ang. high dynamic range – wysoki zakres dynamiczny) i nowatorskim technologiom HDR do przechwytywania i wyświetlania.

W tej chwili oświetlenie rzeczywiste możemy rejestrować, przechowywać, przesyłać i w pełni wykorzystywać do różnych zastosowań bez jakiegokolwiek potrzeby linearyzacji sygnału i radzenia sobie z przyciętymi wartościami. Zatem bardzo ciemne i jasne obszary sceny możemy jednocześnie nagrywać dla zdjęć i filmów, unikając obszarów niedoświetlonych (ang. under-exposed) i prześwietlonych (ang. over-exposed).

Tradycyjne metody obrazowania nie wykorzystują wartości fizycznych i zazwyczaj są ograniczone przez technologię, która wykorzystuje 8 bitów na kanał koloru. Takie obrazowanie (8 bitów lub mniej na kanał) jest typowo nazywane obrazowaniem o niskim zakresie dynamicznym (ang. low dynamic range – LDR).

Obraz HDR może być również wygenerowany poprzez zarejestrowanie wielu obrazów tej samej sceny przy różnych poziomach ekspozycji i scalenie ich, aby zrekonstruować oryginalny zakres dynamiczny sceny.

Obrazy HDR oraz filmy video HDR mogą zajmować nawet cztery razy więcej miejsca w pamięci niż odpowiadające obrazy LDR. Jest to spowodowane tym, że aby poradzić sobie z pełnym zakresem oświetlenia wartości kolorów są przechowywane przy użyciu trzech liczb zmiennoprzecinkowych.

Przechowywanie wartości za pomocą liczb zmiennoprzecinkowych ma nie tylko znaczenie przy przechowywaniu i transmisji danych HDR, ale również przy przetwarzaniu ich. W konsekwencji opracowano efektywną reprezentację zmiennoprzecinkową specjalnie dla obrazów HDR. Ponadto wiele algorytmów kompresji takich jak JPEG, MPEG zostało rozszerzonych, aby obsługiwały również obraz i video HDR.

Mając obraz HDR możemy go wykorzystać do różnych celów, np.

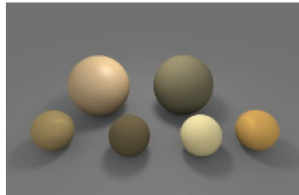
- ▶ ponowne oświetlenie obiektów syntetycznych lub rzeczywistych. Dane HDR przechowują szczegółowe informacje nt oświetlenia środowiska. Informacje te następnie mogą być wykorzystane do wykrycia źródeł światła i ponownego oświetlenia obiektów. W ten sposób wygląd obrazu jest przenoszony na obiekty otrzymując wysoki poziom realizmu;



- ▶ przechwycenie próbek do stworzenia funkcji BRDF.



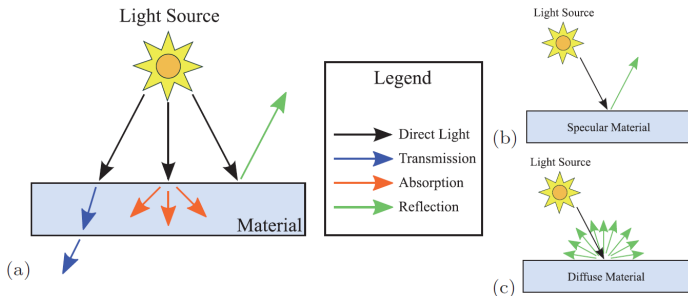
(a)



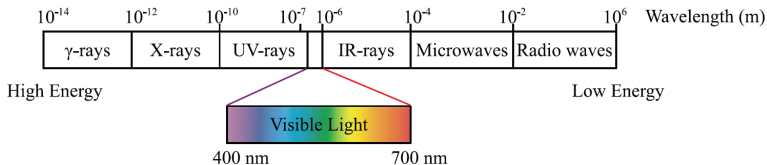
(b)

Światło

Światło widzialne jest formą energii promieniowania, która przemieszcza się w przestrzeni, wchodząc w interakcje z materiałami. Może ono zostać pochłonięte, załamane, odbite i przekazane.



Podróżujące światło może dotrzeć do ludzkich oczu i stymulować je do wywołania wrażeń wizualnych w zależności od długości fali.



Radiometria i fotometria zajmują się pomiarem światła w czasie, przestrzeni i kierunku. W radiometrii mierzone są jednostki fizyczne, zaś w fotometrii bierzemy pod uwagę ludzkie oko, gdzie wartości spektralne są ważone przez spektralną odpowiedź \bar{y} standardowego obserwatora.

Jednostki radiometryczne i fotometryczne są ustandaryzowane przez Commission Internationale de l'Eclairage (CIE).

Główne jednostki radiometryczne to:

- ▶ Energia promienista (ang. radiant energy – Ω_e). Jest to podstawowa jednostka światła, która mierzona jest w dżulach (J).
- ▶ Moc promienista (ang. radiant power – $P_e = \frac{\Omega_e}{dt}$). Jest to ilość energii, która przepływa w jednostce czasu. Mierzona jest w watach ($W = J/s$).
- ▶ Intensywność promienista (ang. radiant intensity – $I_e = \frac{dP_e}{d\omega}$). Jest to ilość mocy promienistej na jednostkę kierunku. Mierzona jest w watach na steradian (W/sr).

- ▶ Irradiancja (ang. irradiance – $E_e = \frac{dP_e}{dA_e}$). Jest to ilość mocy promienistej na jednostkę powierzchni. Mierzona jest w watach na metr kwadratowy (W/m^2).
- ▶ Radiancja (ang. radiance – $L_e = \frac{d^2P_e}{dA_e \cos \theta d\omega}$). Jest to ilość mocy promienistej przychodzącej/wychodzącej z punktu w określonym kierunku. Mierzona jest w watach na steradian na metr kwadratowy ($W/(sr \cdot m^2)$).

Główne jednostki fotometryczne to:

- ▶ Moc świetlna (ang. luminous power – P_v). Jest to ważona moc promienista. Jest mierzona w lumenach (lm). Jest to jednostka pochodna kandeli ($lm = cd \cdot sr$).
- ▶ Energia świetlna (ang. luminous energy – Q_v). Jest ona analogiczna do energii promienistej. Mierzona jest w lumenach na sekundę (lm/s).
- ▶ Intensywność świetlna (ang. luminous intensity – I_v). Jest to moc świetlna w kierunku. Jest mierzona w kandelach (cd).
- ▶ Natężenie oświetlenia (ang. illuminance – E_v). Jest ono analogiczne do irradancji. Mierzone jest luksach (lux), które są równoważne lm/m^2 .
- ▶ Luminancja (ang. luminance – L_v). Jest to ważona radiancja. Mierzona jest w kandelach na metr kwadratowy (cd/m^2) lub nitach.

Miara względnej luminancji sceny jest bardzo użyteczna, ponieważ pokazuje ona pewne własności sceny, takie jak: obecność powierzchni matowych lub połyskliwych, warunki oświetlenia itp.

Względna miara luminancji nazywana jest **kontrastem**. Formalnie kontrast jest definiowany jako stosunek pomiędzy najciemniejszą i najjaśniejszą wartością na scenie i może być obliczony na różne sposoby:

- ▶ kontrast Webera $C_W = \frac{L_{max} - L_{min}}{L_{min}}$,
- ▶ kontrast Michelsona $C_M = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$,
- ▶ współczynnik kontrastu $C_R = \frac{L_{max}}{L_{min}}$,

gdzie L_{min} i L_{max} to minimalna i maksymalna wartość luminancji na scenie.

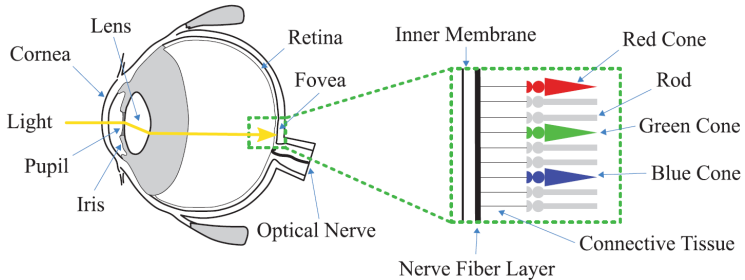
Na zajęciach będziemy korzystać ze współczynnika kontrastu.

Podstawowe wiadomości o ludzkim oku

Oko jest organem, który zbiera światło na fotoreceptorach, które następnie przekształcają je w sygnały.

Sygnały te transmitowane są przez nerw wzrokowy (ang. optical nerve) do kory wzrokowej (ang. visual cortex) czyli do obszaru mózgu, który przetwarza sygnały tworząc postrzegany obraz.

Cały ten system, który jest odpowiedzialny za wizję, nazywany jest ludzkim układem wzrokowym (ang. human visual system – HVS).



lens – soczewka, cornea – rogówka, pupil – źrenica, iris – tęczówka,
 retina – siatkówka, fovea – dołek środkowy, inner membrane –
 błona wewnętrzna, connective tissue – tkanka łączna, nerve fiber
 layer – warstwa włókien nerwowych, rod – pręcik, cone – czopek

Wyróżniamy dwa typy fotoreceptorów: czopki i pręciki. Spektralna wrażliwość fotoreceptorów przy różnych poziomach luminancji jest opisywana przez skotopowe (ang. scotopic), mezopowe (ang. mesopic) i fotopowe (ang. photopic) funkcje wydajności świetlnej. Czasami funkcje te nazywane są funkcjami czułości kontrastu (ang. contrast sensitivity functions – CSF).

Czopków jest ok. 6–7 milionów i są one wrażliwe na poziomy luminancji pomiędzy 10^{-2}cd/m^2 a 10^8cd/m^2 (widzenie fotopowe lub inaczej widzenie dzienne). Są one odpowiedzialne za postrzeganie wzorów o wysokiej częstotliwości, szybkiego ruchu i kolorów.

Za widzenie kolorowe odpowiedzialne są trzy typy czopków: o krótkiej długości fali (wrażliwe na długość ok. 435 nm), o średniej długości fali (wrażliwe na długość ok. 530 nm), o długiej długości fali (wrażliwe na długość ok. 580 nm).

Pręcików jest ok. 75–150 milionów i są one wrażliwe na poziomy luminancji pomiędzy 10^{-6}cd/m^2 a 10cd/m^2 (widzenie skotopwe lub inaczej widzenie nocne). W tym zakresie pręciki są bardziej wrażliwe niż czopki, ale nie zapewniają widzenia kolorowego. Występuje tylko jeden rodzaj pręcików.

Widzenie mezopowe lub inaczej widzenie zmierzchowe zachodzi dla poziomów luminancji pomiędzy 10^{-2}cd/m^2 a 10cd/m^2 . Wówczas zarówno czopki jak i pręciki są aktywne.

Czopki i pręciki kompresują oryginalny sygnał, redukując zakres dynamiki przychodzącego światła. Kompresja jest zgodna z funkcją w kształcie litery S:

$$\frac{R}{R_{max}} = \frac{I^n}{I^n + \sigma^n},$$

gdzie R – odpowiedź fotoreceptora, R_{max} – maksymalna odpowiedź fotoreceptora, I – intensywność światła. Zaś zmienne σ i n to półnasycenie i wykładnik kontroli czułości, które są różne dla czopków i pręcików.

Przestrzenie kolorów

Przestrzeń kolorów to matematyczny opis dla reprezentacji kolorów. Zazwyczaj reprezentowany przez trzy składowe nazywane kolorami podstawowymi. Wyróżniamy dwie klasy przestrzeni kolorów: zależne i niezależne od urządzenia.

Przestrzeń zależna od urządzenia opisuje informacje o kolorze w relacji do technologii reprodukcji koloru użytej w urządzeniu. W przypadku monitorów zależy ona od podstawowych luminoforów, a w drukarkach atramentowych od zbioru podstawowych tuszów. Wadą tej reprezentacji jest to, że kolor o tych samych współrzędnych będzie wyglądał różnie na różnych urządzeniach.

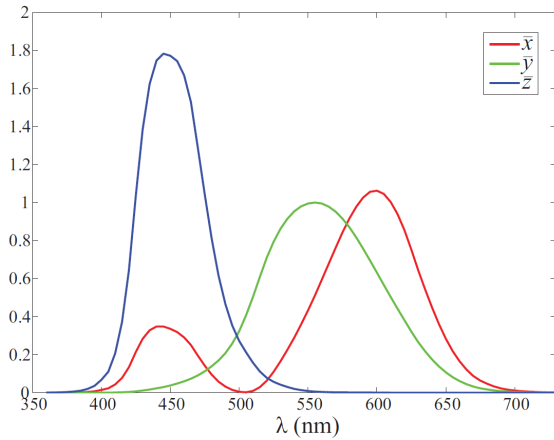
Przestrzeń niezależna od urządzenia nie zależy od charakterystyki danego urządzenia. W ten sposób kolor reprezentowany w takiej przestrzeni zawsze odpowiada tej samej informacji o kolorze.

Przykładem przestrzeni zależnej od urządzenia jest przestrzeń RGB, którą możemy reprezentować jako sześcian.

Przykładem przestrzeni niezależnej od urządzenia jest przestrzeń CIE 1931 XYZ. Jest ona formalnie zdefiniowana jako rzut widmowego rozkładu mocy I na funkcje dopasowania kolorów \bar{x} , \bar{y} i \bar{z} :

$$X = \int_{380}^{830} I(\lambda)\bar{x}(\lambda)d\lambda, \quad Y = \int_{380}^{830} I(\lambda)\bar{y}(\lambda)d\lambda, \quad Z = \int_{380}^{830} I(\lambda)\bar{z}(\lambda)d\lambda,$$

gdzie λ to długość fali, a funkcje \bar{x} , \bar{y} i \bar{z} przedstawione są na następnym slajdzie.

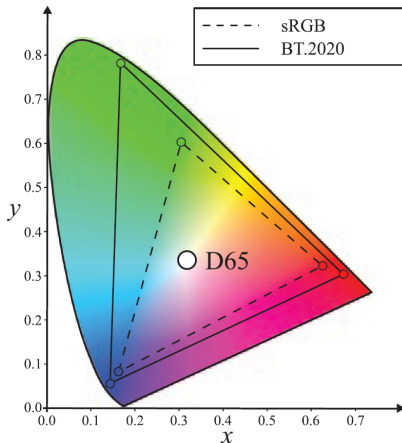


Przestrzeń kolorów XYZ została tak zaprojektowana, że składowa Y mierzy luminancję koloru.

Chromatyczność koloru obliczana jest ze składowych XYZ następująco:

$$x = \frac{X}{X + Y + Z}, \quad y = \frac{Y}{X + Y + Z}.$$

Wartości x , y mogą być naniesione na wykres tworząc tzw. diagram chromatyczności CIE xy . Diagram ten pokazuje wszystkie kolory dostrzegane przez ludzki układ wzrokowy.



Przestrzeń sRGB (używana w monitorach CRT i LCD) oraz BT.2020 (używana w telewizji UHD TV – ultra-high definition television). Przestrzeń BT.2020 ma znacznie większy gamut.

Istnieje liniowa zależność między przestrzeniami XYZ i RGB:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M \begin{bmatrix} R \\ G \\ B \end{bmatrix},$$

gdzie macierz M zależy od przyjętej przestrzeni RGB.

W przypadku przestrzeni sRGB macierz M ma postać:

$$M_{sRGB} = \begin{bmatrix} 0.412 & 0.358 & 0.181 \\ 0.213 & 0.715 & 0.072 \\ 0.019 & 0.119 & 0.950 \end{bmatrix}$$

Przestrzeń sRGB przedstawia nieliniową transformację każdego kanału koloru w celu linearyzacji sygnału wyświetlanego na ekranie LCD lub CRT.

Tą nieliniową transformację nazywamy funkcją EOTF (ang. electro-optical transfer function). Jej odwrotność nazywana jest funkcją OETF (ang. opto-electrical transfer function). Potrzeba takiej funkcji wynika z faktu, że istnieje nieliniowa zależność pomiędzy wyjściową intensywnością wygenerowaną przez wyświetlacz a napięciem wejściowym.

Zależność ta dla przestrzeni sRGB jest modelowana za pomocą:

$$C_v = \begin{cases} 12.92C, & \text{gdy } C \leq 0.0031308, \\ 1.055C^{\frac{1}{\gamma}} - 0.055, & \text{w przeciwnym przypadku,} \end{cases}$$

gdzie C to kanał koloru, $\gamma = 2.4$.

Dla przestrzeni BT.2020 macierz M ma postać:

$$M_{BT.2020} = \begin{bmatrix} 0.636958 & 0.144617 & 0.168881 \\ 0.262700 & 0.677998 & 0.059302 \\ 0.000000 & 0.028073 & 1.060985 \end{bmatrix}$$

Zaś funkcja EOTF modelowana jest wzorem

$$C_v = \begin{cases} 4.5C, & \text{gdy } C \leq \beta, \\ \alpha C^{\frac{1}{\gamma}} - (\alpha - 1), & \text{w przeciwnym przypadku,} \end{cases}$$

gdzie $\gamma = 2.2$, $\alpha = 1.099$, $\beta = 0.018$ dla 10-bitowego systemu, a $\alpha = 1.0993$, $\beta = 0.0181$ dla 12-bitowego systemu.

Wiele obliczeń związanych z HDR wykonywanych jest dla kanału luminancji Y w modelu XYZ, który zazwyczaj oznaczany jest przez L .

Dodatkowo wykorzystywane są różne statystyki związane z luminancją:

- ▶ minimalna L_{min} i maksymalna L_{max} wartość luminancji,
- ▶ arytmetyczna średnia wartość luminancji

$$L_{avg} = \frac{1}{N} \sum_{i=1}^N L(x_i),$$

- ▶ logarytmiczna średnia wartość luminancji

$$L_H = \exp \left(\frac{1}{N} \sum_{i=1}^N \log(L(x_i) + \varepsilon) \right),$$

gdzie x_i – współrzędne i -tego piksela, $\varepsilon > 0$ – mała stała w celu uniknięcia osobliwości dla funkcji logarytmicznej, tzn. $\log 0$.