ITE4052 Computer Vision

Light

Dong-Jin Kim

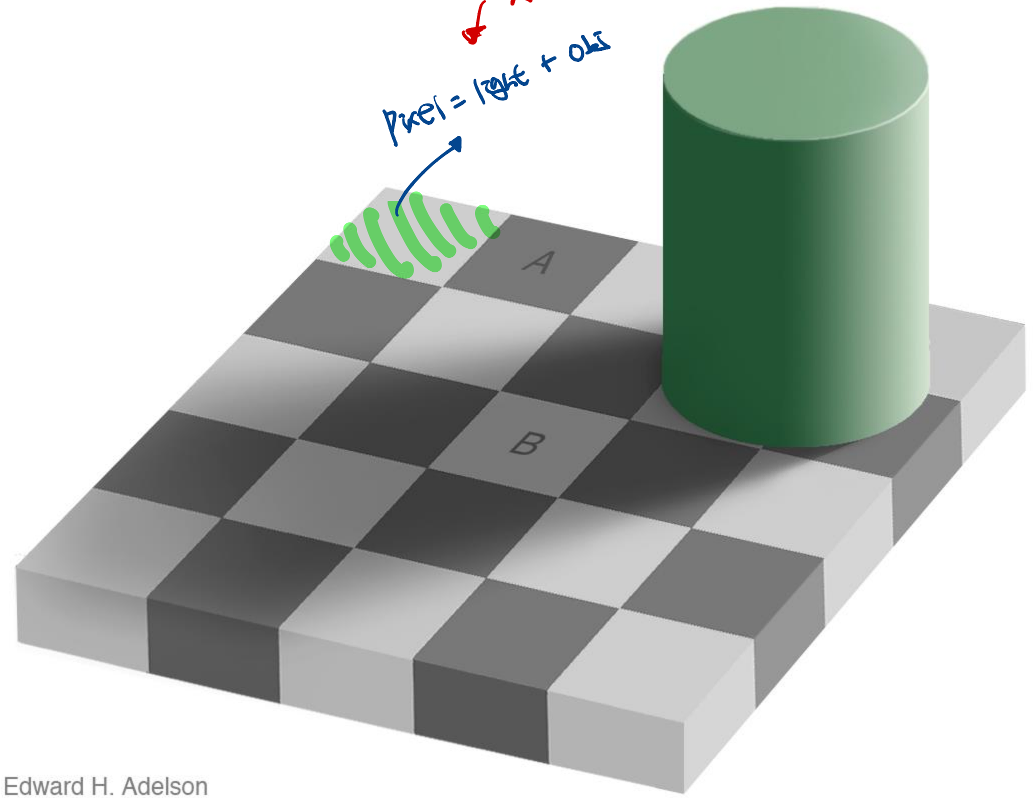
Spring 2025

HANYANG UNIVERSITY

Light

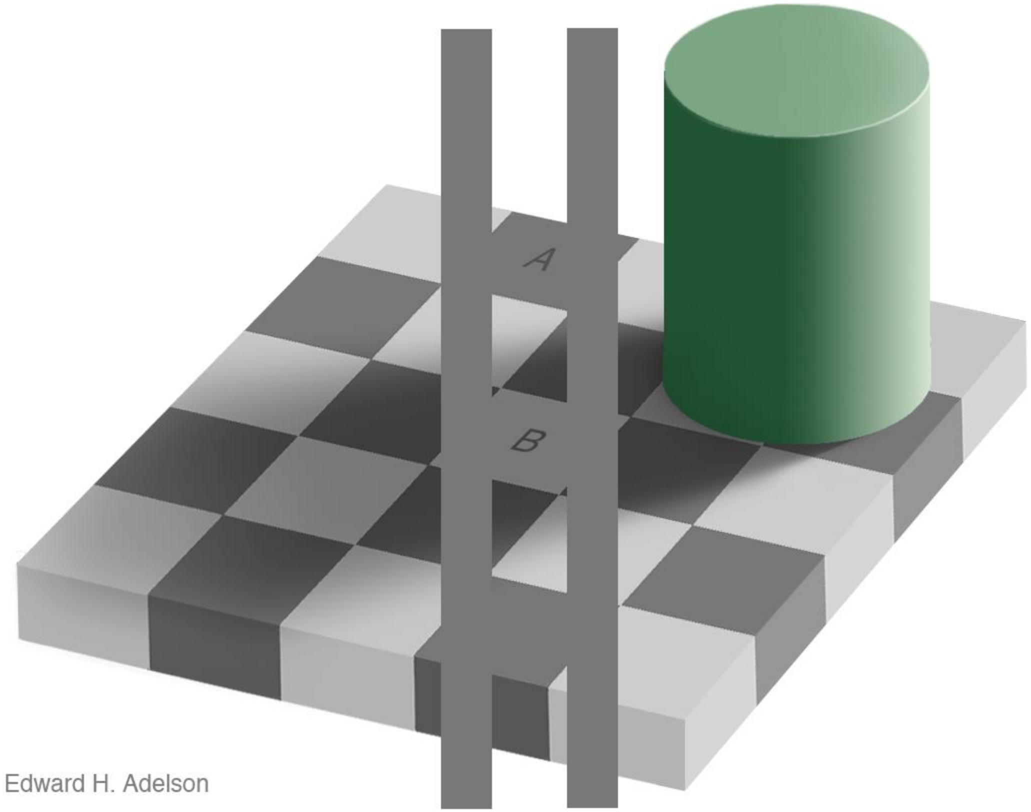
A와 B가 같은 색처럼 느껴지는 이유

Pixel = light + obj



Edward H. Adelson

Light



What color is the dress?

White + Gold?

사람은 Light의 Color가 예측

→ 같은 object에 대해 사람마다 다른 색으로 느끼는 이유

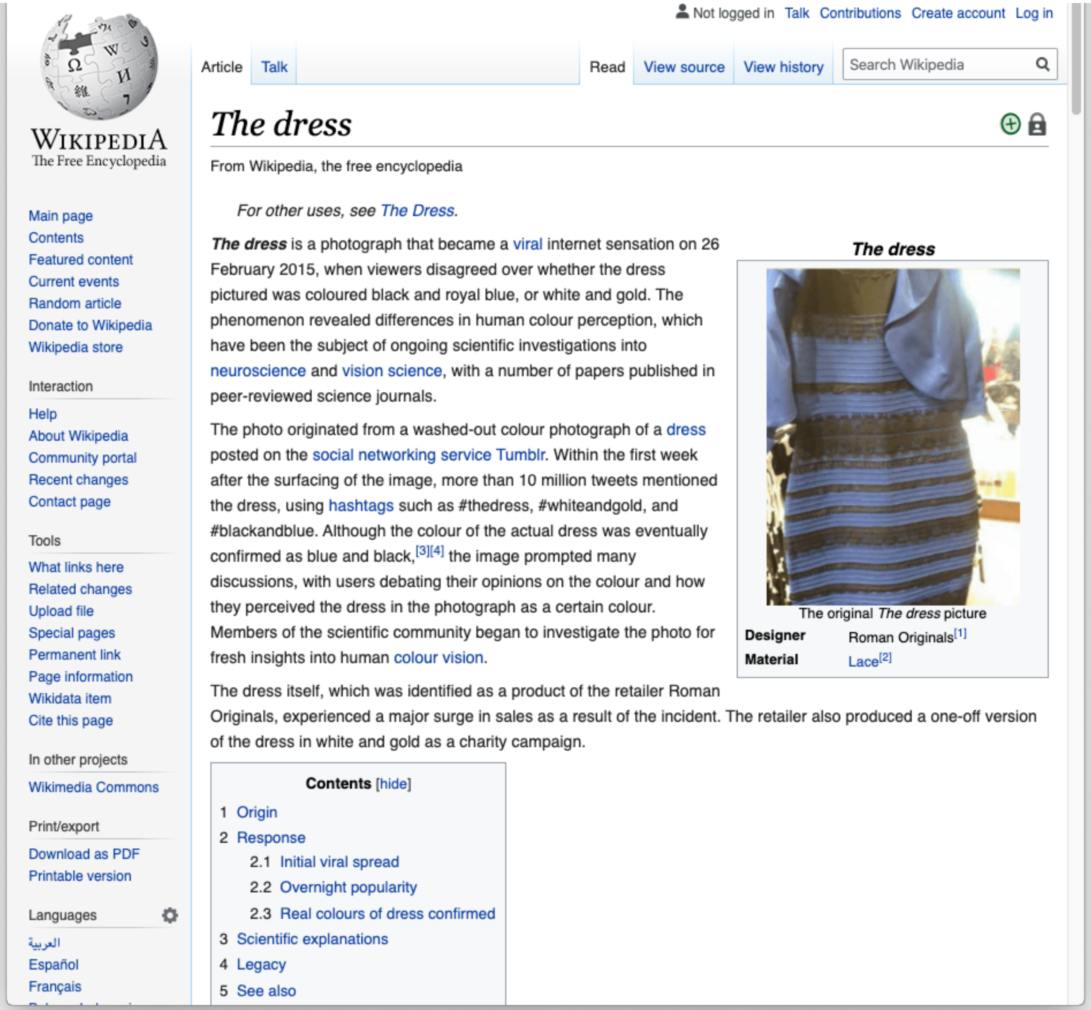
Black + Blue?

Reflectance and Illumination

Pixel = light + object

Light에 따라 보이는 것이 달라진다.

빛과 물체의 색을 합쳐 인식



Not logged in Talk Contributions Create account Log in

WIKIPEDIA
The Free Encyclopedia

Article Talk | Read View source View history | Search Wikipedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

In other projects
Wikimedia Commons

Print/export
Download as PDF
Printable version

Languages
العربية
Español
Français

The dress

From Wikipedia, the free encyclopedia

For other uses, see The Dress.

The dress is a photograph that became a viral internet sensation on 26 February 2015, when viewers disagreed over whether the dress pictured was coloured black and royal blue, or white and gold. The phenomenon revealed differences in human colour perception, which have been the subject of ongoing scientific investigations into neuroscience and vision science, with a number of papers published in peer-reviewed science journals.

The photo originated from a washed-out colour photograph of a dress posted on the social networking service Tumblr. Within the first week after the surfacing of the image, more than 10 million tweets mentioned the dress, using hashtags such as #thedress, #whiteandgold, and #blackandblue. Although the colour of the actual dress was eventually confirmed as blue and black,[3][4] the image prompted many discussions, with users debating their opinions on the colour and how they perceived the dress in the photograph as a certain colour. Members of the scientific community began to investigate the photo for fresh insights into human colour vision.

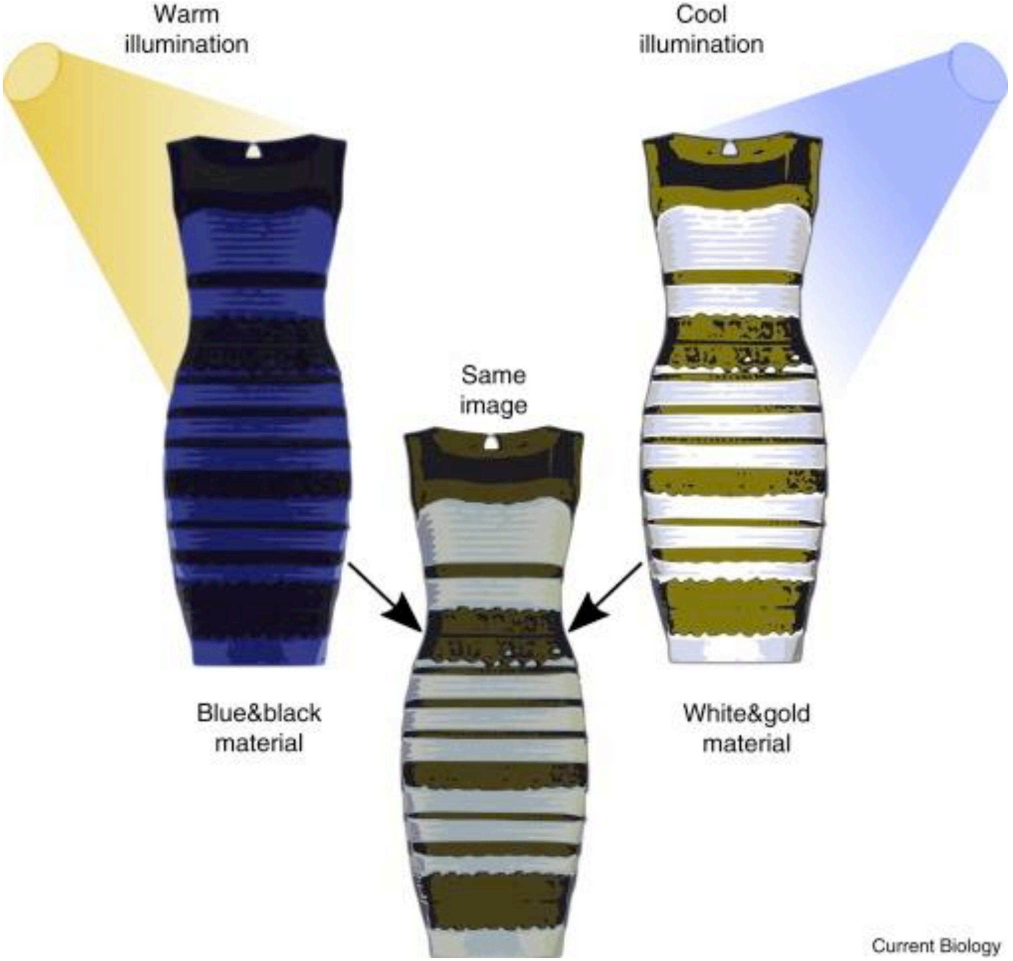
The dress itself, which was identified as a product of the retailer Roman Originals, experienced a major surge in sales as a result of the incident. The retailer also produced a one-off version of the dress in white and gold as a charity campaign.

The dress

The original *The dress* picture

Designer	Roman Originals[1]
Material	Lace[2]

Contents [hide]



What color is the center ball?

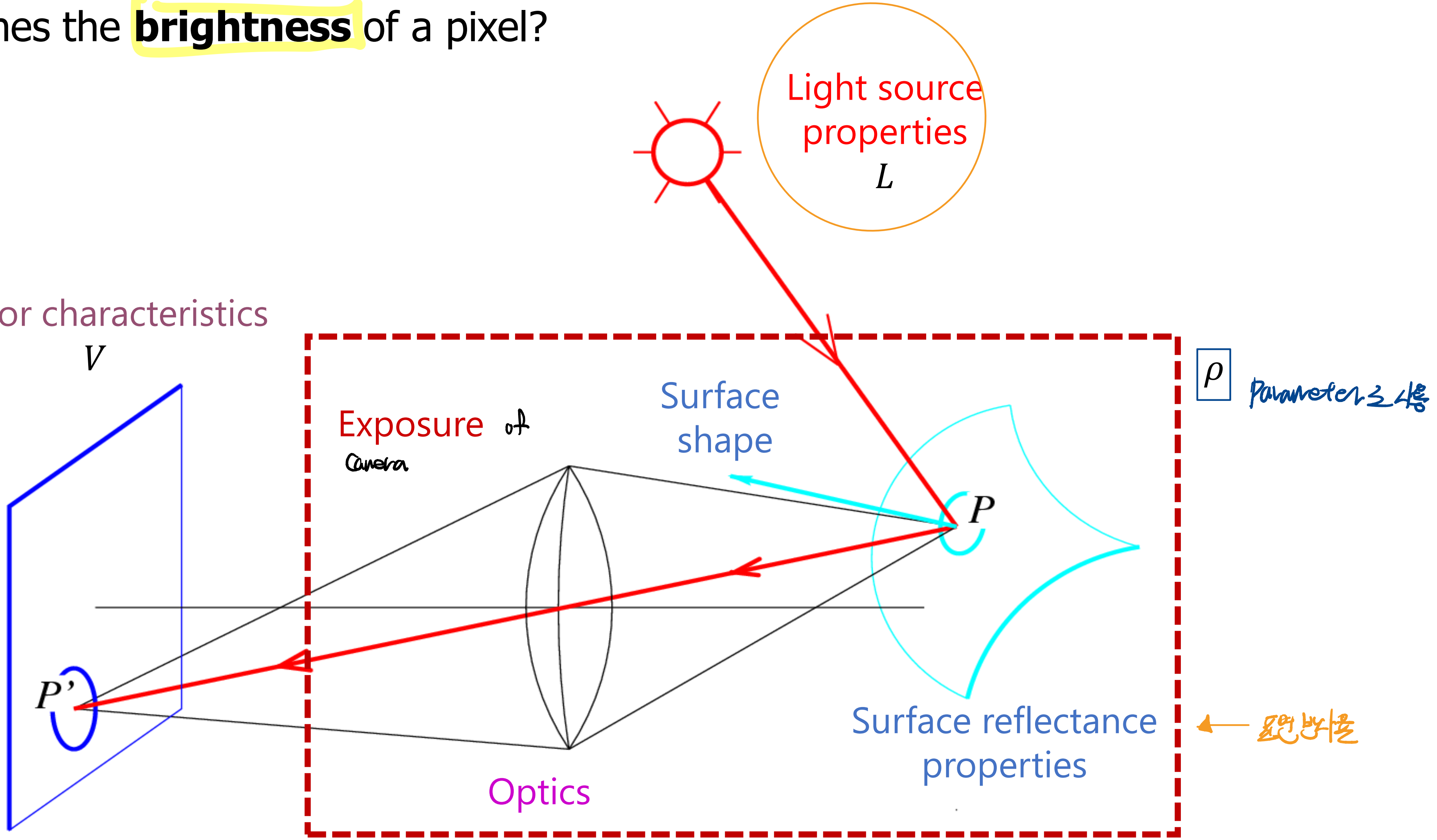
gray = cyan light + red object

What color is the center ball?

Radiometry

$\rho \rightarrow$ 하나의 parameter로 모조정보

What determines the **brightness** of a pixel?



Human Eye

Light received from an object (Illuminance):

물체에서 반사된 빛의 양

$$I(\lambda) = \underbrace{\rho(\lambda)}_{\text{무엇이 보일지 결정하는 요소}} L(\lambda)$$

λ: 빛의 파장

L(λ):

- The amount of **light** incident on a surface.
- Incident **energy** distribution

→ 물체 표면에 도달하는 빛의 양

Reflectance (albedo) ρ(λ):

반사율

- The proportion of incident light that is **reflected** from a surface .
- The reflectivity or transmissivity of the object

→ 반사율, Surface shape

Luminance

하나의 물체에서 반사하는 빛의 양 (특정 파장의)

$I(\lambda) = \rho(\lambda) \cdot L(\lambda)$ (for all pixel)

사람이 인식하는 빛의 양

$L(\lambda) = \int I(\lambda) \cdot V(\lambda)\, d\lambda$

휘도 (강도) Pixel의 값: $f(x,y)$

Luminance (or Intensity): The energy a visual system receives from a light source

→ 광원으로부터 실제 받아들이는 Energy, 빛의 양

$$f(x,y) = \int_0^\infty I(x,y,\lambda)\, V(\lambda)\, d\lambda \quad \text{where } I(\lambda) = \rho(\lambda)\, L(\lambda)$$

(x,y) pixel의 Energy

물체의 이 위치에 반사하는 ~

인간은 모든 파장을 동일하게 느끼지 못한다. 이를 적용하는 함수 (파장대별로 가중치)

적분?: 모든 파장 영역에 대해서 계산

- I: **Light** distribution received from an object
- V: **Luminosity** function (Relative luminous efficiency function of the visual system)

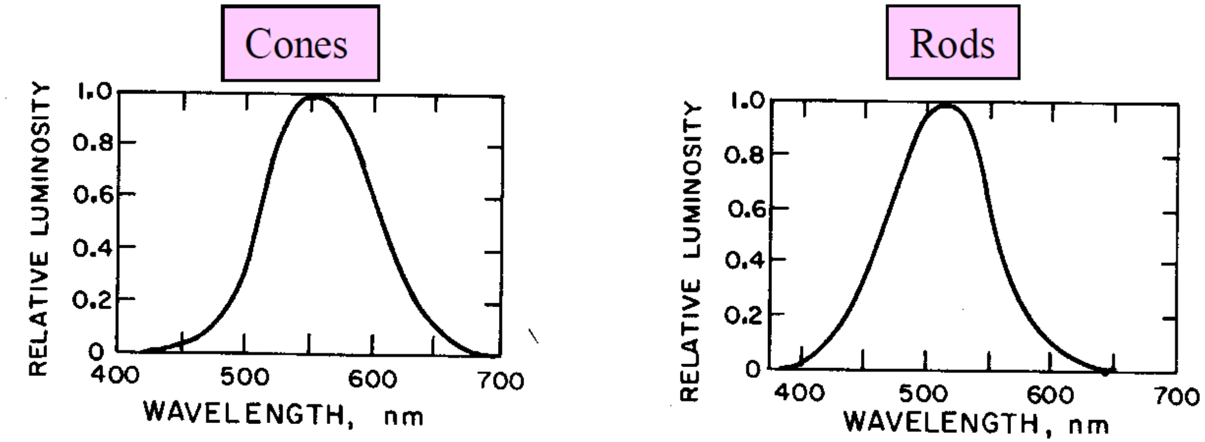
Photosensitive Cells

Color perception

- Retina contains photosensitive cells (**rods** and **cones**)

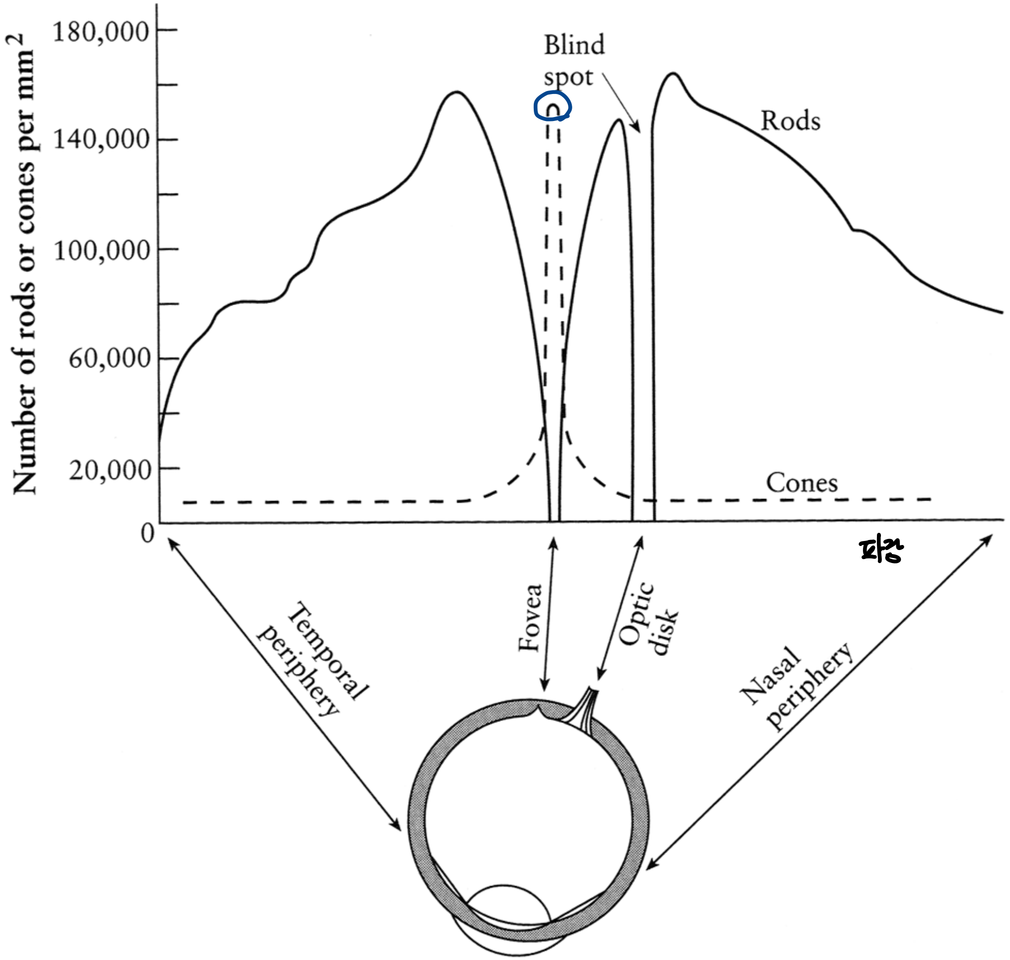
- **Rods** (간상체)→ Intensity (명암, 밝기 차이 등)
- **Cones** (추상체) → Color

→ 파장에 따라 감도가 다름 : V(λ)



(a) Photopic luminosity function (b) Scotopic luminosity function

FIGURE 3.1-4. Relative luminous efficiency functions (9).



Brightness Contrast and Constancy

Brightness → Pixel value

흑백 이미지에선 흑 — 백 정도가 Brightness

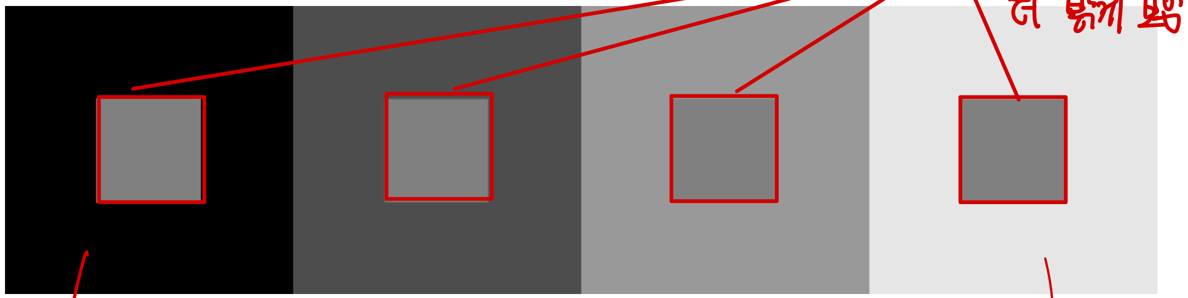
→ 색을 다르게 인식하는 것은 Brightness의 착각이다.

The apparent **brightness** depends on the surrounding region

- **brightness contrast**: a constant region seems lighter or darker depending on the surroundings

  → 같은 색 주변에 따라 다르게 인식

  같은 밝기인데 주변부가 어두울수록 더 밝게 보임



- **brightness constancy**: a surface looks the same under widely varying lighting conditions.

  → 조명이 달라지면 인식되는 색이 달라져야 하지만 인간은 조명이 달라져도 같은 색으로 인식

  → Pixel 값이 바뀌어도 다른 조건에 의해 같은 Pixel로 인식

light↓ → real color of object ↑

light ↑ → real color of object ↓

Brightness / Contrast

사람이 느끼는 (다른 조건에 의해 변할 수 있음)

Brightness (what we really perceive)

- Perceived **luminance** depending on the luminance of surround

  사람이 느끼는

Contrast : 물체와 배경의 밝기 차이가 얼마나 인지할 정도인가?

- Weber's Law: If object luminance is just noticeably different from the surround luminance,

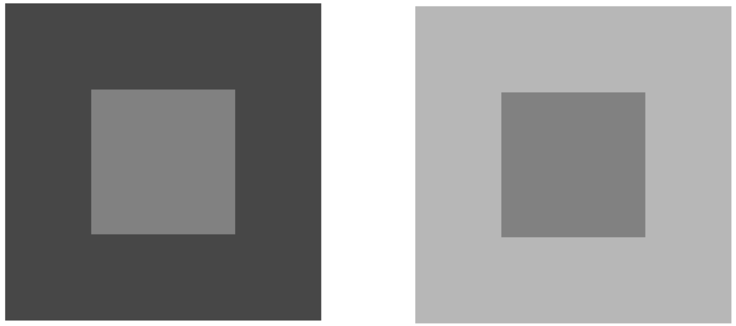
$$\frac{\text{Brightness} - \text{Luminance}}{\text{Luminance}}$$

$$\frac{|f_{surround} - f_{object}|}{f_{object}} = \frac{|\Delta f|}{f} = \Delta c \cong 0.02.$$

← 일정시간에 안함

$f_{object} \times 2\%$ 라면 차이를 인식

- Simultaneous contrast effect: Spatial interaction in brightness



1. Color와 Brightness의 관계

RGB 모델에서 "색(color)"과 "밝기(brightness)"는 서로 다른 성분입니다.

- **RGB**는 색을 나타내는 **세 개의 채널 (R, G, B)** 을 조합해 전체 색을 만듦
- **밝기(Brightness)** 는 이 RGB가 **얼마나 밝은지**, 즉 **에너지 총량(빛의 세기)** 에 해당

2. 밝기를 RGB에 대해 세 번 적용한 게 색인가?

→ 아니요. 반대입니다.

- **색(color)** 은 R, G, B의 **비율**
- **밝기(brightness)** 는 RGB 값들의 **총량(크기)**

즉, **같은 색(color)** 이라도 **밝기(brightness)** 를 달리하면:

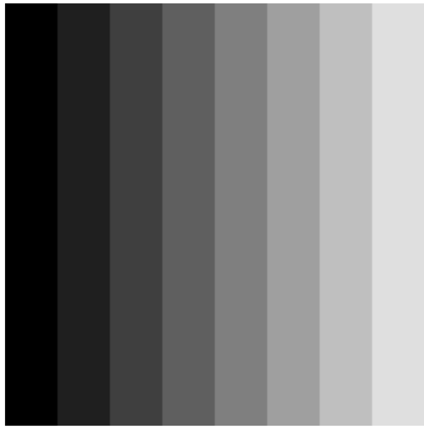
: $(R, G, B) = (100, 50, 50)$

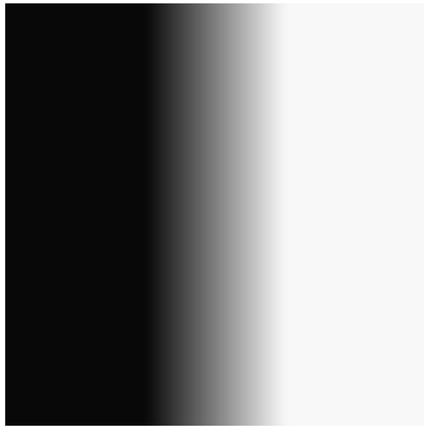
2 : $(200, 100, 100)$

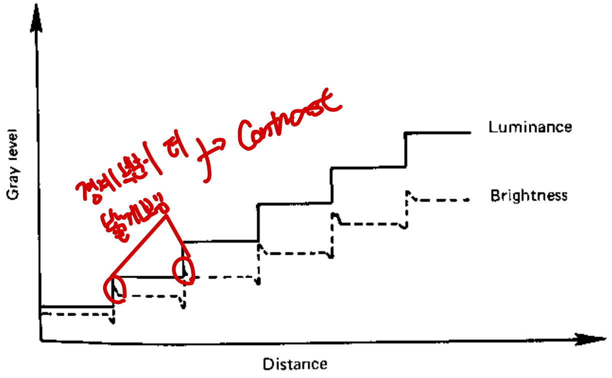
, 2

Contrast

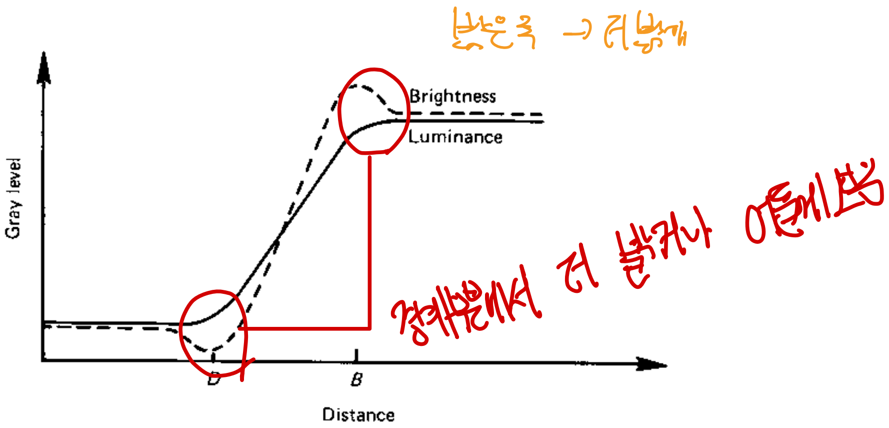
Lateral Inhibition







Luminance versus brightness.



Mach band effect.

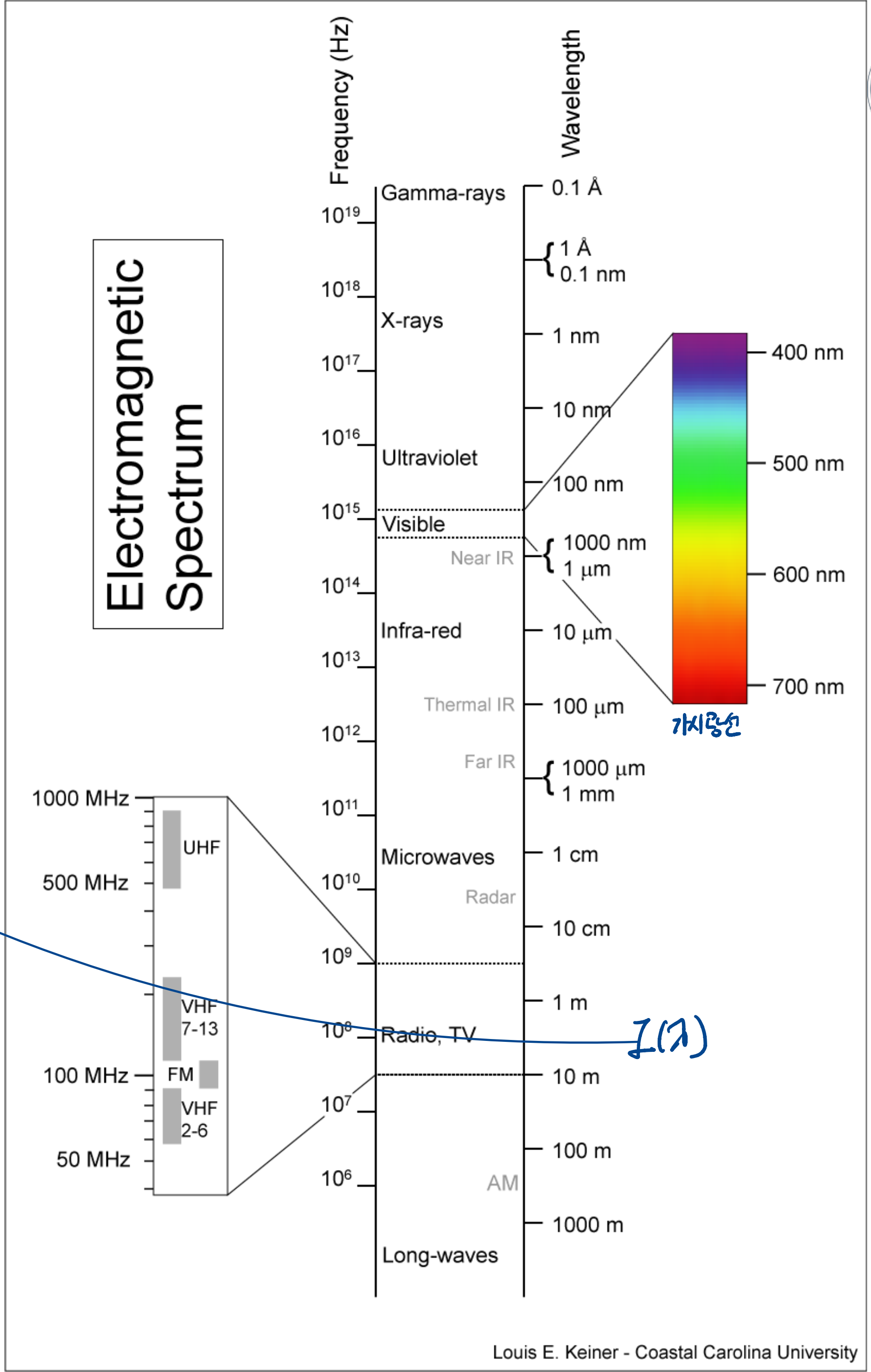
Visible Light

We “see” light in a range of **wavelengths**

Power Spectrum

- “How much power (or energy) at each 파장 **wavelength**”
  강도 표현

Our visual system converts a light spectrum into **“color”**



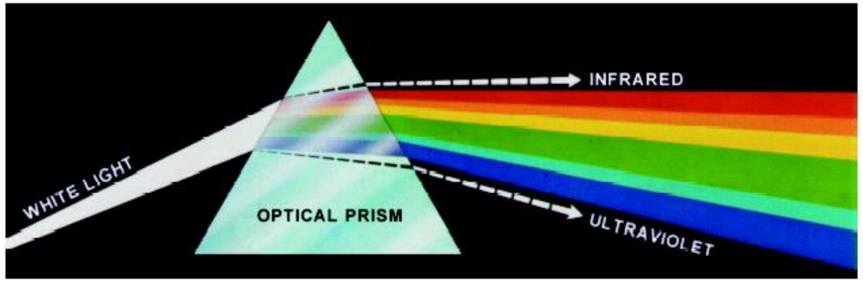
단

장

Color

Wavelength of the electromagnetic **spectrum**

빛 → Spectrum



$\lambda\uparrow \rightarrow E\downarrow$

$\lambda\downarrow \rightarrow E\uparrow$

FIGURE 6.1 Color spectrum seen by passing white light through a prism. (Courtesy of the General Electric Co., Lamp Business Division.)

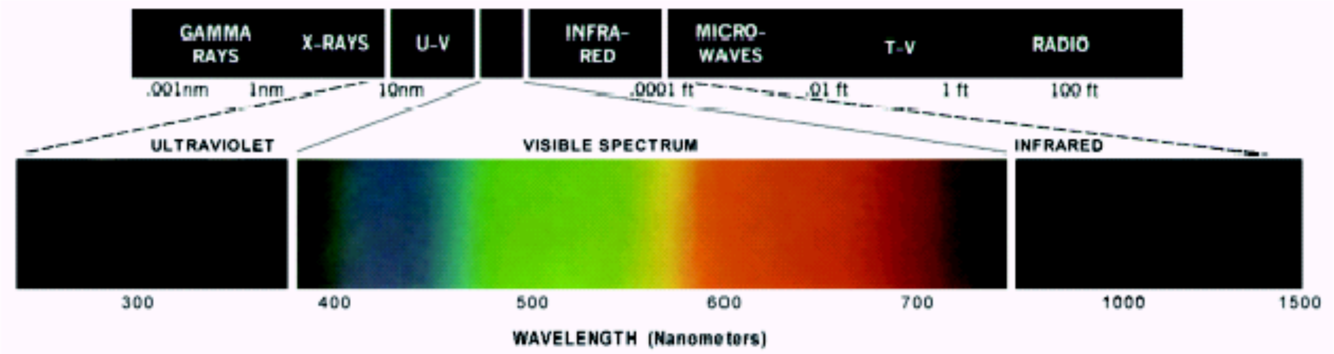


FIGURE 6.2 Wavelengths comprising the visible range of the electromagnetic spectrum. (Courtesy of the General Electric Co., Lamp Business Division.)

3 Types of Cones

$$f(x,y) = \int_0^\infty I(x,y,\lambda)V(\lambda)d\lambda$$

R,G,B 별로 각각의 function이 정의되어져 있다.

Each is sensitive in a different region of the spectrum (**red**/**green**/**blue**, but overlapping)

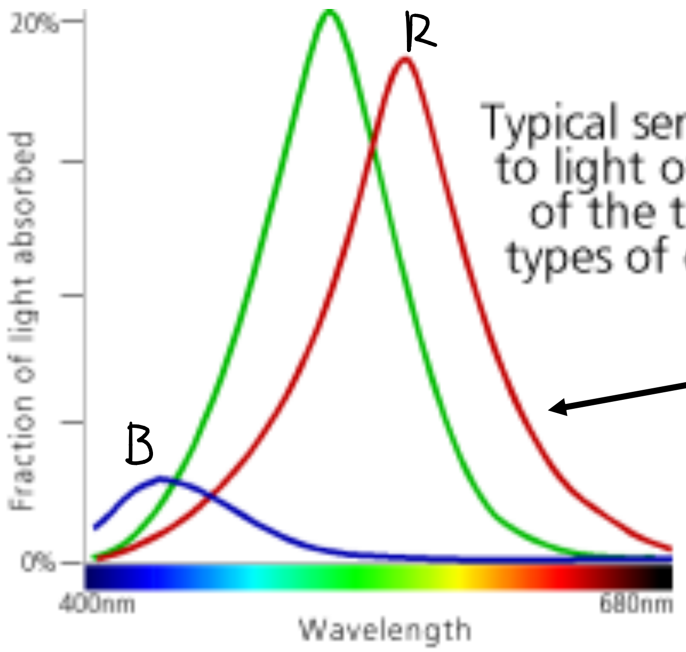
3가지 Cones 세포

Long — L-Cones : Red에 가장 민감

Medium — M-Cones : Green에 가장 민감

Short — S-Cones : Blue에 가장 민감

각각 민감한 파장 영역이 있다. 완전히 독립적이지 않고 반응 영역이 겹친다.



Typical sensitivity to light of each of the three types of cones

V response curve

$\boxed{\text{Sensitivity function} \rightarrow V(\lambda)}$

세 가지 Cones의 조합

사람이 인식하는 빛 : 물체가 반사하는 빛

물체가 반사하는 빛? 물체 표면과 광원의 영향을 받는다.

$$I(\lambda) = \underbrace{\rho(\lambda)}_{\text{물체 표면}} \cdot \underbrace{L(\lambda)}_{\text{광원}}$$

사람의 눈은 파장대마다 빛을 인식하는 정도가 다르다.

$$f(x,y) = \int \underbrace{I(\lambda)}_{\text{빛의 양}} \cdot \underbrace{V(\lambda)}_{\text{사람의 인식 정도}} \, d\lambda$$

특정 pixel에서 인식하는 빛은 전 파장 영역에서 사람시스템이 물체에서 반사되는 빛을 어떻게 인식하는지를 물체에서 반사된 빛의 양과 곱한 값이다.

$V(\lambda)$: Human Visual System

Rods : 밝기만 인식 (어두운 환경)

Cones : 색을 인식 (L, M, S Cone이 존재 : 각각 R, G, B 파장에 민감)

Luminance

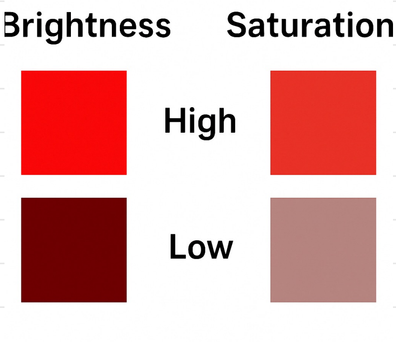
가중합

- $f_R + f_G + f_B$ (밝은 환경)

$f_R = \int I(\lambda) \cdot V_R(\lambda) \, d\lambda$, $f_G = \int I(\lambda) \cdot V_G(\lambda) \, d\lambda$, $f_B = \int I(\lambda) \cdot V_B(\lambda) \, d\lambda$

- f_{Rod} (어두운 환경)

$\Rightarrow \int I(\lambda) \cdot V_{Rod}(\lambda) \, d\lambda$



Color

- f_R, f_G, f_B의 비율

$f_R : f_G : f_B$ → 같은 비율은 같은 색, 그러나 밝기가 다름

Brightness : 실제로 느끼는 밝기 (주변 환경에 따라)

Contrast : 환경이 달라 밝기로 실제와 다름

Constancy : 환경이 달라도 밝기가 실제와 동일

Color Perception

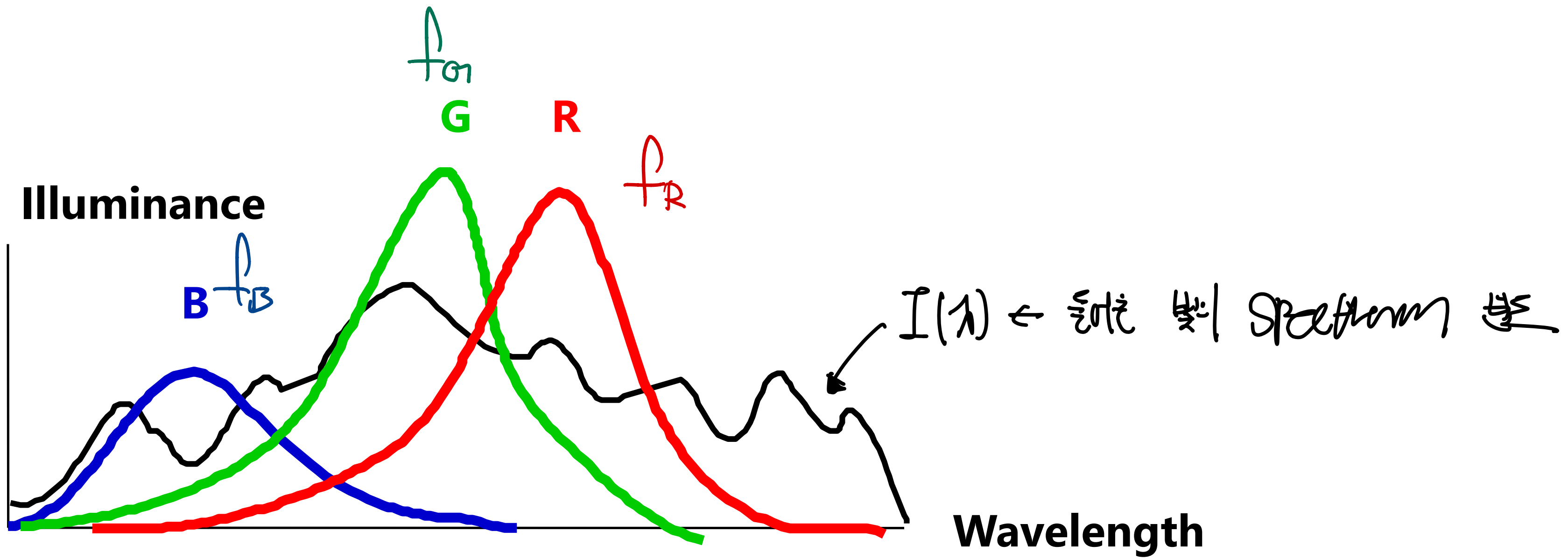
$$f(x,y) = \int_0^\infty \underline{I(x,y,\lambda)V(\lambda)}\,d\lambda$$

스펙트럼

적분 → 각 R,G,B값이 나옴

각 파장 영역에 대해 빛을 얼마나 감지하는지

- **Rods** and **cones** act as $\boxed{\text{filters on the spectrum}}$
  - **multiply** its response curve (V) by the spectrum (I), integrate over all wavelengths

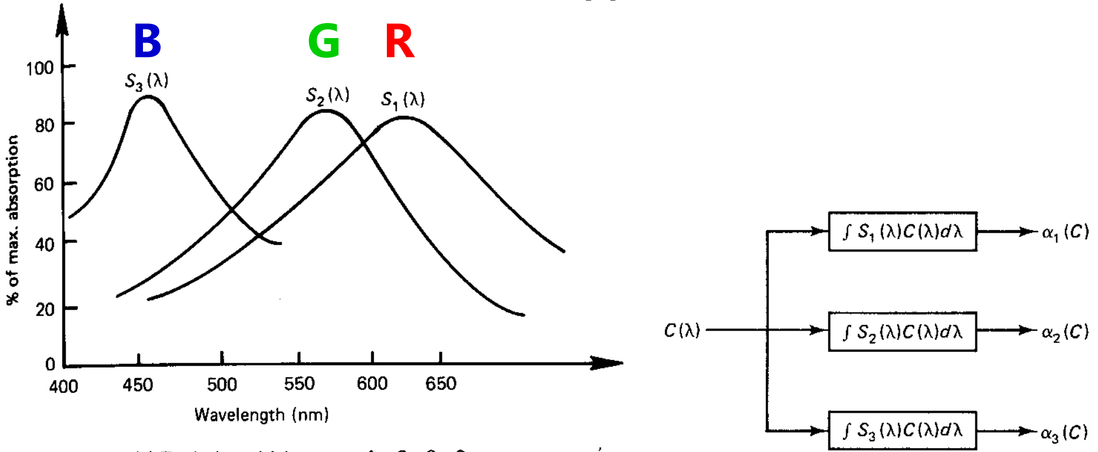


Human Visual System

$$f(x,y)=\int_{0}^{\infty} I(x,y,\lambda)V(\lambda)d\lambda$$

- **Tri-chromatic Model**: Human retina has 3 types of cones.

3원색
→ 3개의 Cone



(a) Typical sensitivity curves for S_1, S_2, S_3 (not to scale).

(b) Three receptor model for color representation

Figure 3.11 (a) Typical absorption spectra of the three types of cones in the human retina; (b) three-receptor model for color representation.

- i-th cone response

가시광선: 380 ~ 700

↳ i번째 Cone이 어떻게 반응하는지

R,G,B 값이 나온다

$$\boxed{\alpha_i(C)}=\int_{\boxed{\lambda_{\min}}}^{\boxed{\lambda_{\max}}} \underbrace{S_i(\lambda)}_{\text{V (response)}}\underbrace{C(\lambda)d\lambda}_{\text{I (light)}}, \qquad i=1,2,3$$

들어온 빛의 Spectrum

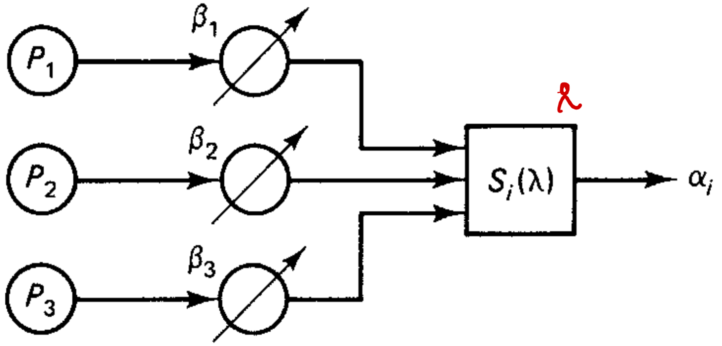
i=1 → R
i=2 → G
i=3 → B

↳ Response function을 알아야 색 가능

"Should we know the **response function** to represent the color?"

Color Matching : 임의의 C(λ)은 세가지 primary light sources로 표현해보자

사람은 빛의 합에서 해당 빛의 색을 primary light으로 재현한다



$\int, \Sigma$ is linear

- 3 primary light sources (k=1,2,3)

Basis → Linearly Independent!

새로운 light 정의

$$[C(\lambda)] = [\beta_1 P_1(\lambda) + \beta_2 P_2(\lambda) + \beta_3 P_3(\lambda)] \qquad \int P_k(\lambda)d\lambda = 1$$

(β: scale, P: distribution)

P_1, P_2, P_3 → Three different distribution

Color representation

$$\alpha_i(C) = \int \boxed{\sum_{k=1}^{3} \beta_k P_k(\lambda)} S_i(\lambda)d\lambda = \sum_k \int \beta_k P_k S_i d\lambda$$

* 우리는 $S_i(\lambda)$, $P_i(\lambda)$, $C(\lambda)$를 미리 알고있어야 한다.

β를 구하자!

각각 R,G,B 같은 sensor의 출력값

↳ i번째 sensor에 k번째 prior light source를 넣었을 때

$$\alpha_i(C) = \sum_{k=1}^{3} \beta_k \int \boxed{S_i(\lambda)P_k(\lambda)d\lambda}$$

$$\begin{bmatrix} a \end{bmatrix} \cdot \begin{bmatrix} \beta \end{bmatrix} = \begin{bmatrix} \alpha \end{bmatrix}$$

$$\boxed{a_{i,k}} \equiv \alpha_i(P_k) = \int S_i(\lambda)P_k(\lambda)d\lambda$$

→ S,P를 알아 구할 수 있다.

Primary response on **sensor** $\boldsymbol{i}$ from the **light source** $\boldsymbol{k}$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \alpha_1(C) \\ \alpha_2(C) \\ \alpha_3(C) \end{bmatrix}$$

$a_{i,k}$

$$\boxed{\sum_{k=1}^{3} \beta_k a_{i,k} = \alpha_i(C)} = \int S_i(\lambda)C(\lambda)d\lambda$$

Sensor response as a "**combination**" $\boldsymbol{\beta}$ of primary responses

(Color Matching Equation)

↳ Matrix 가 Invertible 하면

"Unless $\{a_{i,k}$, i,k=1,2,3$\}$ is a singular matrix, we can find $\{\beta_k$ k=1,2,3$\}$ for any C."

세 가지 주요 빛을 사용하여 어떤 색이든 만들어보자

이 빛은 총 3개 해당하는 빛을 조합 : $C(\lambda) = \beta_1 P_1(\lambda) + \beta_2 P_2(\lambda) + \beta_3 P_3(\lambda)$

원래의 색과 다르다 같다. when $\underbrace{S = [V_R, V_G, V_B]}_{\text{Sensor}}$, $\underbrace{\alpha = [f_R, f_G, f_B]}_{\text{Sensor response}}$

$$\alpha = \int C(\lambda) \cdot S(\lambda)\, d\lambda$$

$$= \int C(\lambda) \cdot S(\lambda)\, d\lambda$$

$$\alpha_i = \int \left(\sum_{k=1}^{N} \beta_k P_k\right) \cdot S_i(\lambda)\, d\lambda$$ ← R, G, B 중 하나의 값

$$\alpha_{i,k} = \int P_k \cdot S_i(\lambda)\, d\lambda$$ ← Primary light 중 하나의 R, G, B 중 하나

$$\rightarrow \sum_{k=1}^{N} (\beta_k \cdot \alpha_{i,k}) = \alpha_i$$

→ Primary light와 Sensor를 알면 β를 알고, 어떤 Color라면 Primary

Primary light에 대한 Sensor Response

light의 선형 조합으로 만들어낼 수 있다.

+) $\alpha_{i,k}$가 Not Scalar

그래서 우리가 R, G, B 같은 primary light을 알면 어떤 Color 든지

Primary light의 Linear Combination으로 표현 가능

$$\boxed{a} \cdot \boxed{\beta} = \boxed{\alpha}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \alpha_1(C) \\ \alpha_2(C) \\ \alpha_3(C) \end{bmatrix}$$

$C(\lambda)$를 $P(\lambda)$의 Linear Combination으로 재현

- $C(\lambda)$
- $P(\lambda)$
- Sensitive function

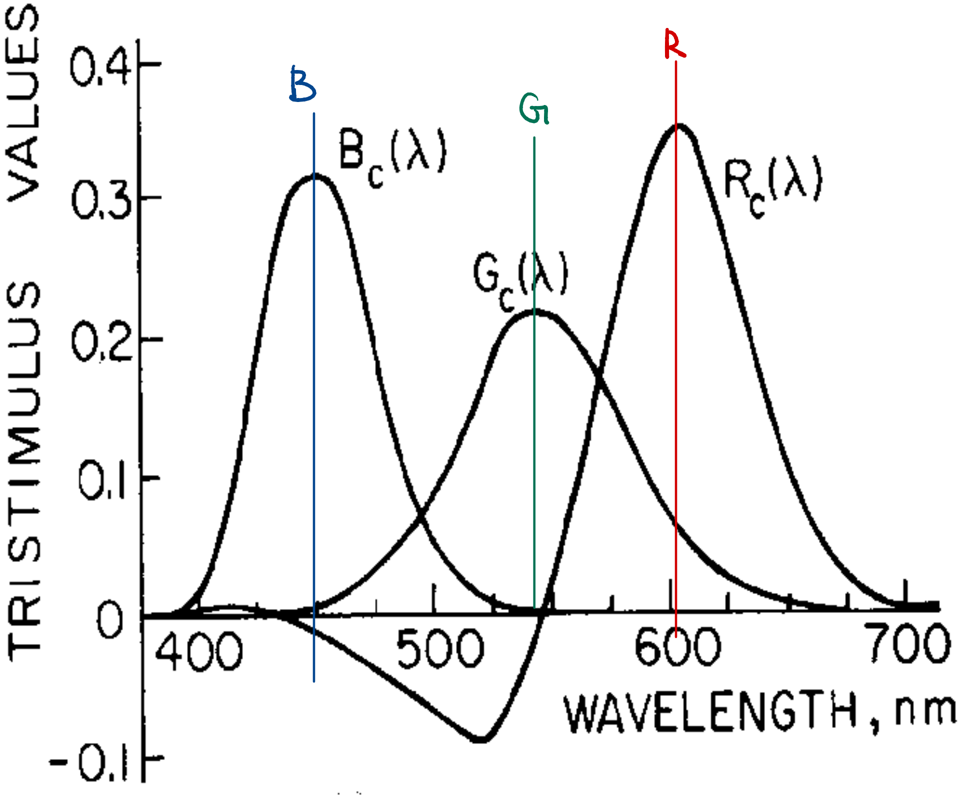
알고 있어야 하는 전제

Color Matching

The CIE spectral primary system

- $P_1(\lambda) = \delta(\lambda - \lambda_1),\ \lambda_1 = 700nm$ (red)
- $P_2(\lambda) = \delta(\lambda - \lambda_2),\ \lambda_2 = 546.1nm$ (green)
- $P_3(\lambda) = \delta(\lambda - \lambda_3),\ \lambda_3 = 435.8nm$ (blue)

CIE에서 사용한 primary light sources



Color Matching

$f(x,y) = \int I(x,y,\lambda) V(\lambda) d\lambda$

$f(x,y) = \int I(x,y,\lambda) V(\lambda) d\lambda$

휘도

빛의 양

인간의 감도함수

색상 하나에서 얻을 수 있는 밝기

$f(c) = \int C(\lambda) S(\lambda) d\lambda$

sensor

색당 하나

Light is linear:

The same intensity

각 밝기의 합 = 혼합한 빛의 밝기

Laws of **color matching** (Grassman's laws)

- Any color can be matched by mixing 3 primary light sources.

어느 색깔이든지 만들어낼 수 있다.

- **Luminance** of a color mixture == sum of the **luminance** of its components

색의 조합 (여러 빛(색)이 섞임)

$$\text{If } C = C_1 + C_2 + \cdots + C_N, \quad \text{then } \boxed{Y(C) = \sum_{n=1}^{N} Y(C_n)}$$

light is linear

각 밝기의 선형 조합이다.

$C = \beta_1 P_1 + \beta_2 P_2 + \beta_3 P_3$ 일때,

$$\Rightarrow \quad Y = Y(C) = \beta_1 L(P_1) + \beta_2 L(P_2) + \beta_3 L(P_3)$$

각 빛의 밝기

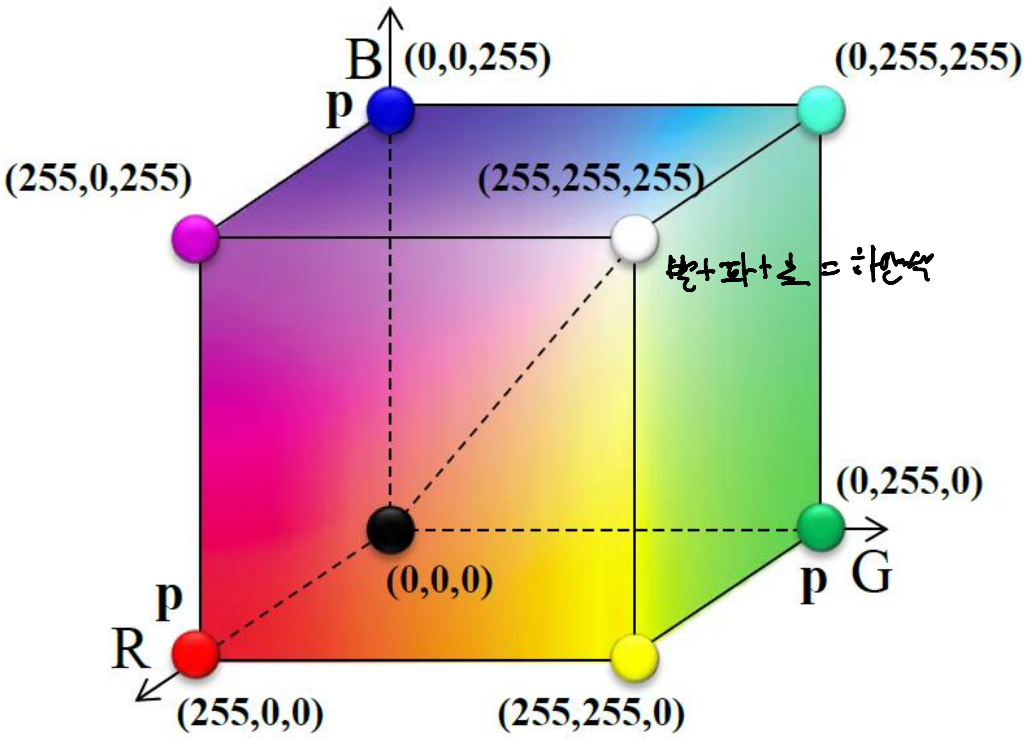
Light = linear combination of 3 light source

- Note $L = L(C) = \int C(\lambda) V(\lambda)\, d\lambda$

Prove) $L(C) = \int C(\lambda) \cdot V(\lambda) d\lambda$

$= \int \left(\sum_{n=1}^{N} C_n(\lambda)\right) \cdot V(\lambda) d\lambda = \sum_{n=1}^{N} \int C_n(\lambda) \cdot V(\lambda) d\lambda = \sum_{n=1}^{N} C_n(\lambda)$

RGB-CMY

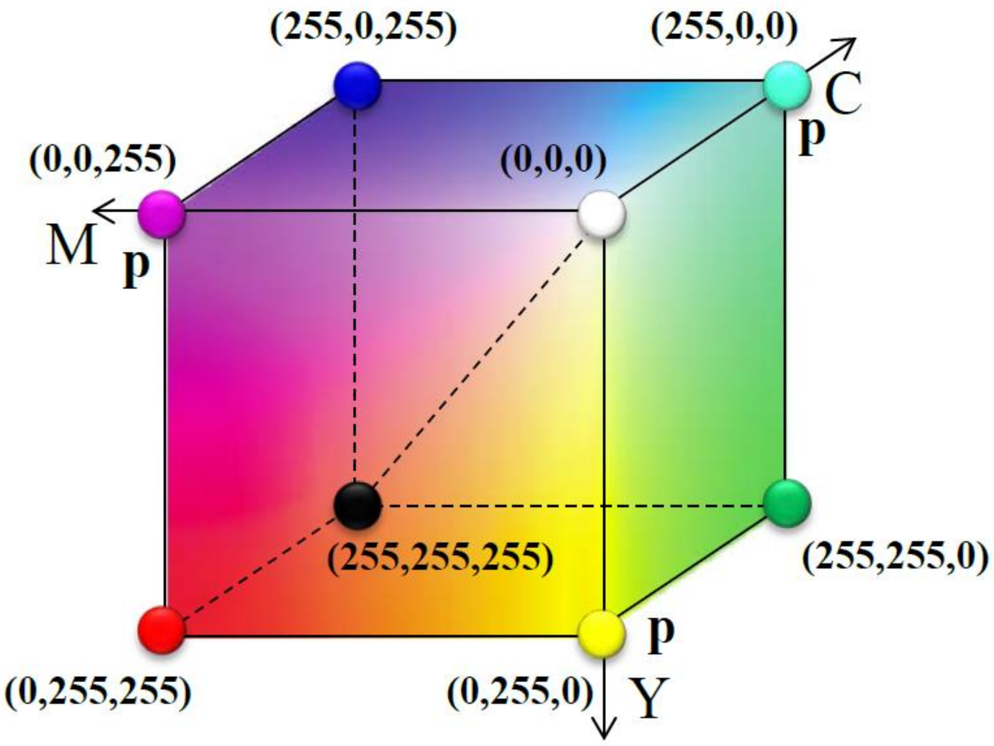


빨+파+초 = 하얀색

빨+초 = 노

RGB
Primary colors of light

빛의 3원색



CMY
(cyan, magenta, yellow)
Primary colors of pigments

잉크의 3원색

Intensity, Saturation, and Hue

H: Hue, I: Intensity, S: Saturation 으로 이루어진 색공간 모델

Hue: θ
Saturation: r
→ Intensity: z

HIS System: Non-linear 한 color space ← Polar Coordinate 처럼 표현한다.

$$I = \frac{1}{3}\left(R_N + G_N + B_N\right) \qquad S = 1 - \frac{\min\left(R_N, G_N, B_N\right)}{I}$$

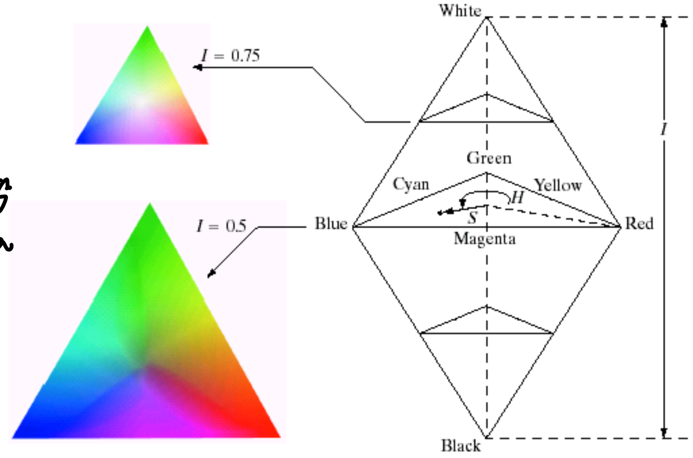
→ R,G,B=1 동일 : 흰색나온다

min(R_N, G_N, B_N) 이 클수록 색이 탁함
→ $R_N = G_N = B_N$
→ S=0 회색

$$H = \frac{1}{2\pi}\cos^{-1}\left\{\frac{2R_N - G_N - B_N}{2\sqrt{(R_N - G_N)^2 + (R_N - B_N)(G_N - B_N)}}\right\}$$

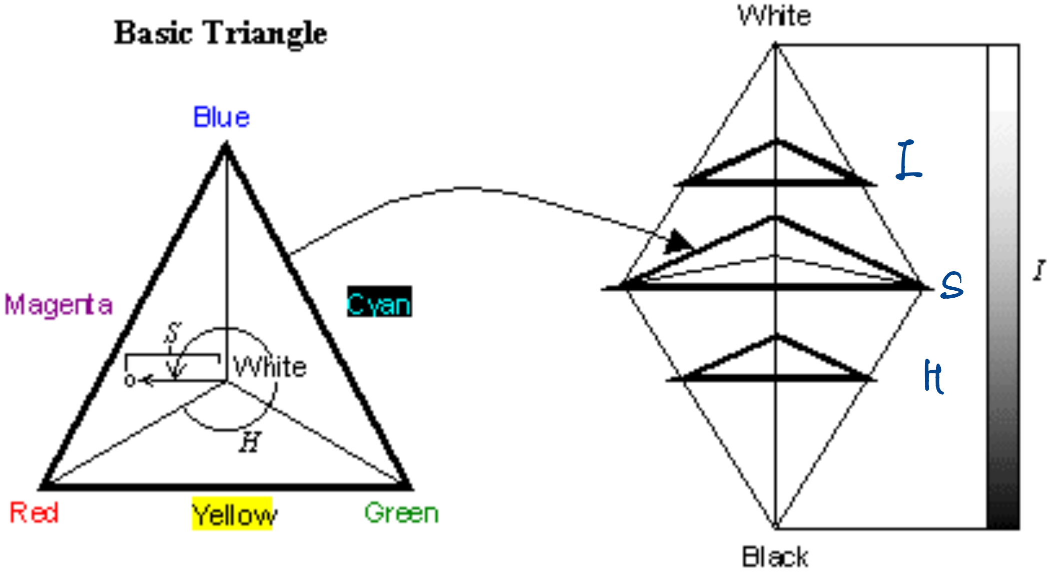
→ Red를 기준으로 각도 계산

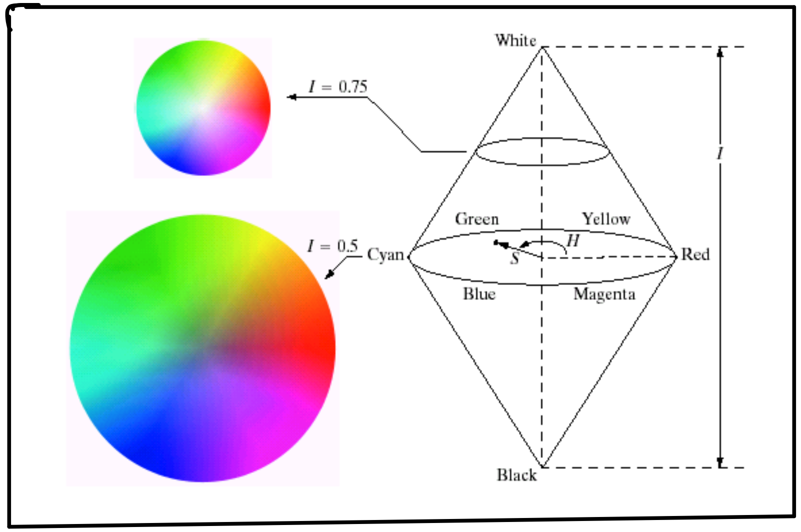
RGB to HSI
: 선형 변환 불가능



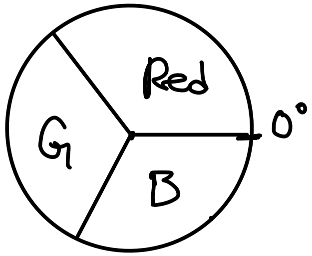
Color Space

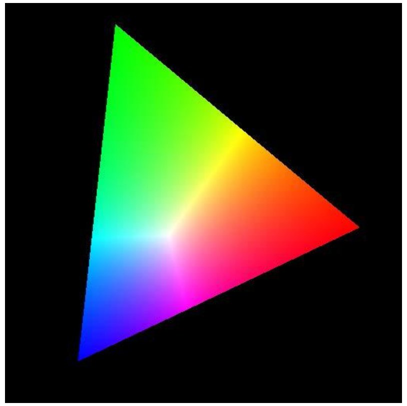
Basic Triangle



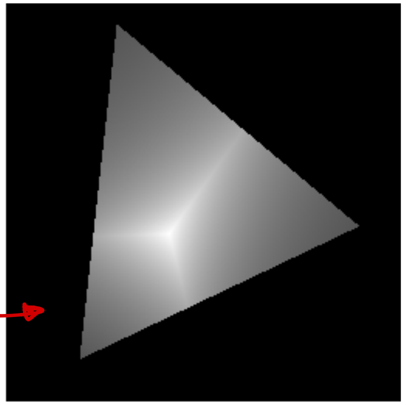


Intensity, Saturation, and Hue



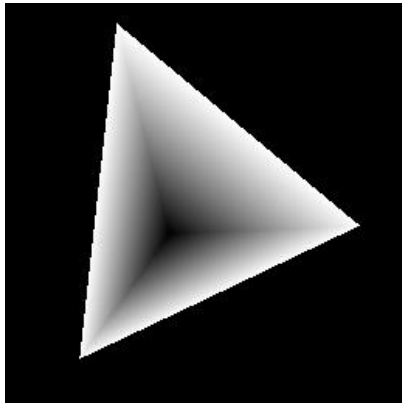


가운데 Y 모양 → 밝다.



I Channel

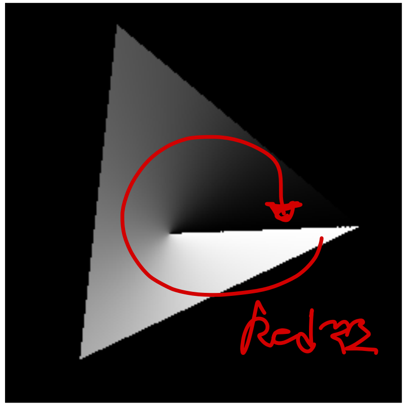
Intensity : 밝기



S Channel

Saturation

: 색 |선명함|



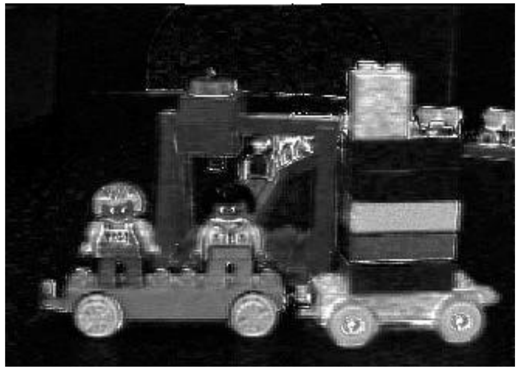
H Channel

Hue : 색상

An Example Image in HSI

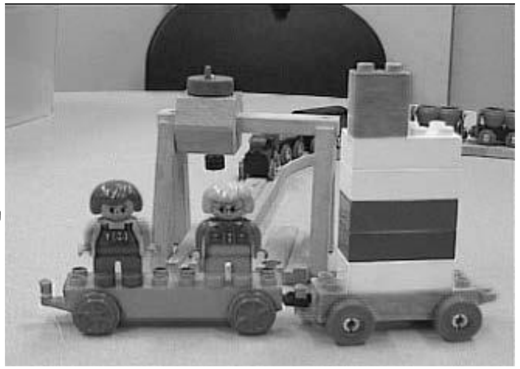
Hue

검정 : 낮은 Hue (빨강 계열)
흰색 : 높은 Hue (보라계열)



Saturation

채도↑
→ 선명한 색
→ 밝게 보임
채도↓
→ 검은색, 흰색, 회색
→ 검게 보임



Intensity

흑백이미지와 동일
흰배경, 밝은 부분 : 흰색

HSI Manipulation

Hue 값을 바꿔서
이미지를 얻을 수 있다.

RGB → HSI → RGB

Hue

Saturation

Intensity

YCbCr Color Coordinate System

영상 압축 등에서 주로 사용

노랑 계열이 b

CL↑

회색 → Cb = Cr = 128

- There are various **color coordinate systems** (XYZ, UVW, YIQ, etc.)
- Most of them can be computed with a linear transformation.

RGB to YCbCr:

정규화 (RGB: 0 ~ 255 이기 때문) Transform matrix

밝기	Luminance
파란색 − Y	blue
빨간색 − Y	red

$$\begin{bmatrix} Y' \\ C_B \\ C_R \end{bmatrix} = \frac{1}{256} \boxed{\begin{bmatrix} 77 & 150 & 29 \\ -43 & -84 & 127 \\ 127 & -106 & -21 \end{bmatrix}} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}$$

RGB to YCbCr matrix

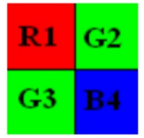
색이 적음 ← 128 → 색이 많음

중심을 128으로 맞춤 → 8-bit에 저장하기 위해 unsigned로

Luminance : 원색 >>> 중간색

$Y = \beta_1 R + \beta_2 G + \beta_3 B$ ← 형태 $\begin{bmatrix} R \\ G \\ B \end{bmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

$C_B = 200 - Y$
$C_R = 100 - Y$
→ 파랑과 빨강이 밝기보다 강한가? 강하다면 파랑, 빨강 계열



2x2 Bayer Pattern block

Top left pixel: (R,G,B) = (R1,G2,B4)
Top right pixel: (R,G,B) = (R1,G2,B4)
Bottom left pixel: (R,G,B) = (R1,G3,B4)
Bottom right pixel: (R,G,B) = (R1,G3,B4)

Colors in **Cameras**

카메라 센서 → 밝기만 측정
↳ Color filter
Filter에서 각 Pixel은 R,G,B 중 하나의 값만 측정하는 형태

- A mosaic of color filters (a **Bayer filter**) in front of the sensor

Image 앞에 있다.

- **Demosaicking**: interpolate colors to make a full-resolution color images

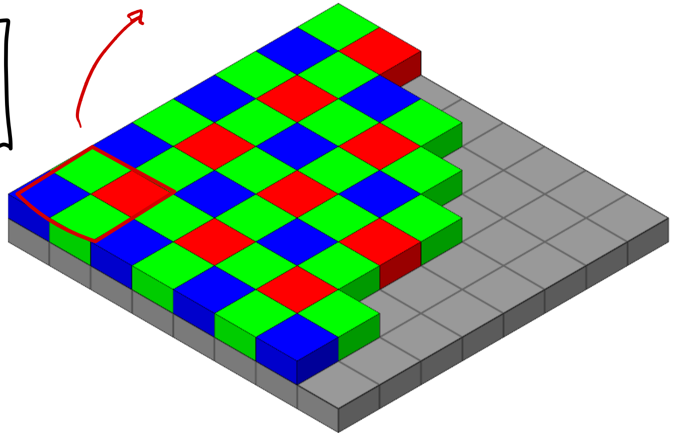
Filter에서 각 Pixel은 R,G,B 중 하나의 값만 인식. 한 Pixel에서 filter로 측정하지 못한 값은 주변값을 통해 예측 (Interpolation 기법)

Demosaicking

일반적으로 2x2 filter R 2G B

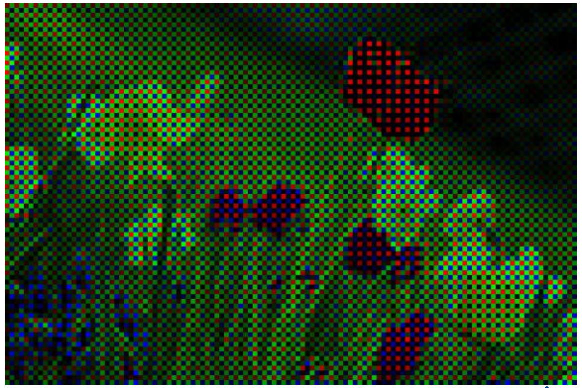
$$\begin{bmatrix} G & R \\ B & G \end{bmatrix}$$

→ 사람눈은 초록색에 가장 민감하기 때문



Bayer filter pattern in front of sensor

2W x 2H x 1 $\xrightarrow{\text{하나의 센서로 부터 한 Pixel이 R,G,B 값을 모두 가지게 할 수 있다.}}$ W x H x 3 → R,G,B 채널



What the camera sees
("raw" image)

Pixel 4 통합

Demosaicked image

✅ 예시 (4×4 RAW → 2×2 RGB)

RAW 값 (4×4, 단일 채널):

```
[
 [ 60 120  55 130 ]
 [ 40  90  42 100 ]
 [ 65 115  58 125 ]
 [ 38  92  40  98 ]
]
```

Bayer 패턴 (G R / B G):

```css
G R G R
B G B G
G R G R
B G B G
```

좌상단 2×2 블록 예)

```makefile
G:60  R:120
B:40  G:90
```

→ 보간된 RGB = [R=120, G=(60+90)/2=75, B=40]

✅ 최종 RGB 결과 (2×2×3)

```lua
[[[120, 75,  40],   [130,  75, 100]],
 [[115, 71,  38],   [125,  75,  98]]]
```

Naïve Color Photos

→ 초기에는

Before color film, 3 separate exposures with 3 different color filters (RGB)

정렬

Blue, Green, Red exposures

Combined color image (1911)

Image **Colorization** [1]: 딥러닝 기반으로 흑백이미지 → 컬러이미지

L 채널 → 밝기 정보   딥러닝이 예측한 색상 정보   최종 이미지

$\boxed{W \times H \times 1} + \boxed{W \times H \times 2} \rightarrow \boxed{W \times H \times 3}$

[1] Zhang, Richard, Phillip Isola, and Alexei A. Efros. "Colorful image colorization." *ECCV* 2016.

$\mathcal{F}$

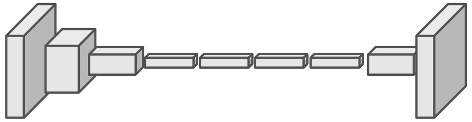
Deep learning
Model

Grayscale image: *L* channel

$$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$$

Color information: *ab* channel

$$\hat{\mathbf{Y}} \in \mathbb{R}^{H \times W \times 2}$$ → 색상 정보

L →  → **ab**

a: Green - Red

G 0 R

-128 128

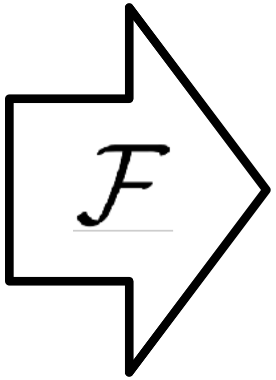
b: Blue - Yellow

B 0 Y

-128 128

왜 Lab? 밝기와 색의 분리 → 색 예측에 유리

$\mathcal{F}$



Grayscale image: *L* channel

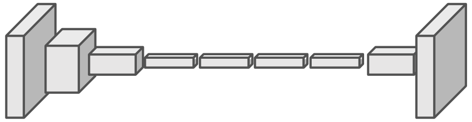
$$\mathbf{X} \in \mathbb{R}^{H \times W \times 1}$$

Concatenate (L,ab)

$$(\mathbf{X}, \widehat{\mathbf{Y}})$$

L과 ab를 융합

→ W × H × 3

L →  → ab

Lab Color space

→ 이를 RGB로 바꾸면 Color Image를 얻을 수 있다.

Loss Functions

Input: 흑백 → output: ab channel (예측)

a: green to red
b: blue to yellow

- Regression with L2 loss inadequate

$$L_2(\widehat{\mathbf{Y}}, \mathbf{Y}) = \frac{1}{2} \sum_{h,w} \| \mathbf{Y}_{h,w} - \widehat{\mathbf{Y}}_{h,w} \|_2^2$$

정답

실제 a,b값 　 예측 a,b값

- Use **multinomial classification** : 분류 기반 Loss function

$$L(\widehat{\mathbf{Z}}, \mathbf{Z}) = -\frac{1}{HW} \sum_{h,w} \sum_{q} \mathbf{Z}_{h,w,q} \log(\widehat{\mathbf{Z}}_{h,w,q})$$

예측

q → 모든 클래스

→ 정답 class에 대해 예측한 확률이 높으면 Loss ↓

H: 이미지의 세로 길이

W: 이미지의 가로 길이

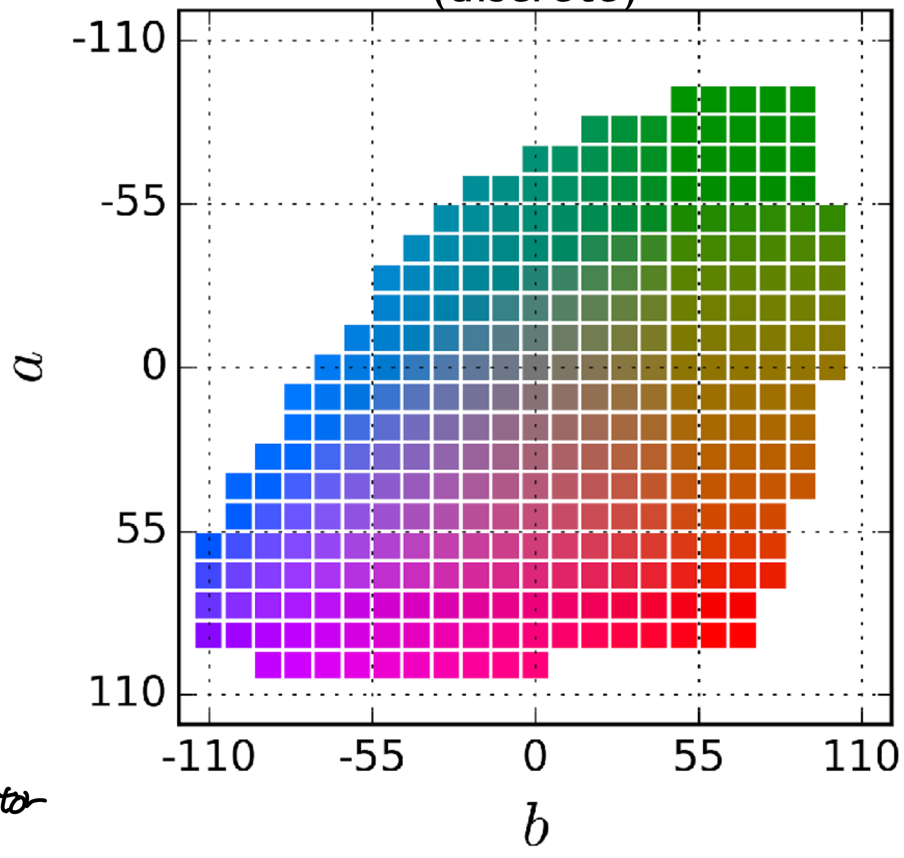
→ 오른쪽 그림에서 각 pixel을 class로 본다.

$Z_{h,w,q}$: 정답 클래스 = 1 인 one-hot vector
[0, 0, 0, 1, 0 ⋯ , 0]

$\hat{Z}_{h,w,q}$: 한 pixel가 어떤 클래스에 속할 확률 벡터
[0.1 0.03 0.09 ---]

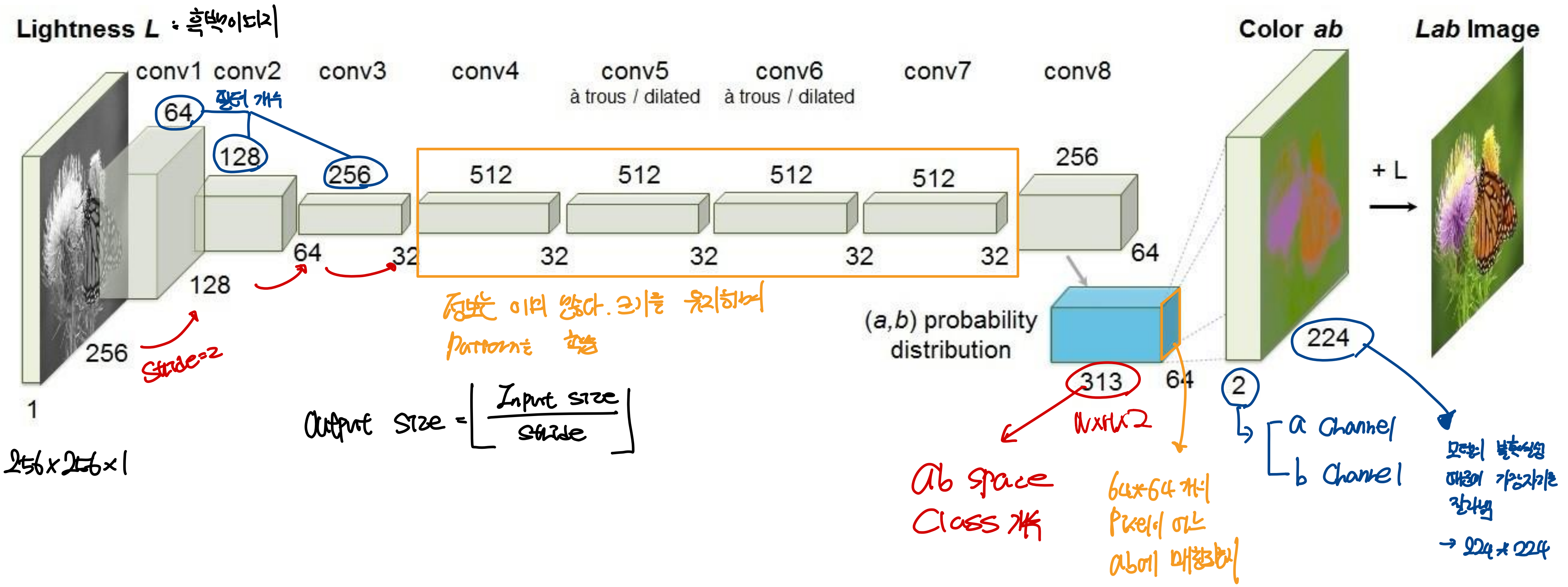
Colors in *ab* space

(discrete)



→ 313개의 Class

Architecture



$$\text{Output size} = \left\lfloor \frac{\text{Input size}}{\text{stride}} \right\rfloor$$

Colorization Results

Results on legacy BW photos.

Inherent Ambiguity

Grayscale

Inherent Ambiguity : 흑백이미지에서 컬러이미지로 변환할 때, L채널에는 색 정보가 없다.
객관적인 정답이 없기 때문에 유사하지만 파란색이 빨간색으로 완전히 변하는 등의 상황이 발생할 수 있다.

Our Output

Ground Truth : 실제 정답

More Recent Colorization Works

First Solution : Stochastic

Diverse Colorization (VAE or GAN) [1] : 여러 결과를 생성해 가장 흡사한 것 선택

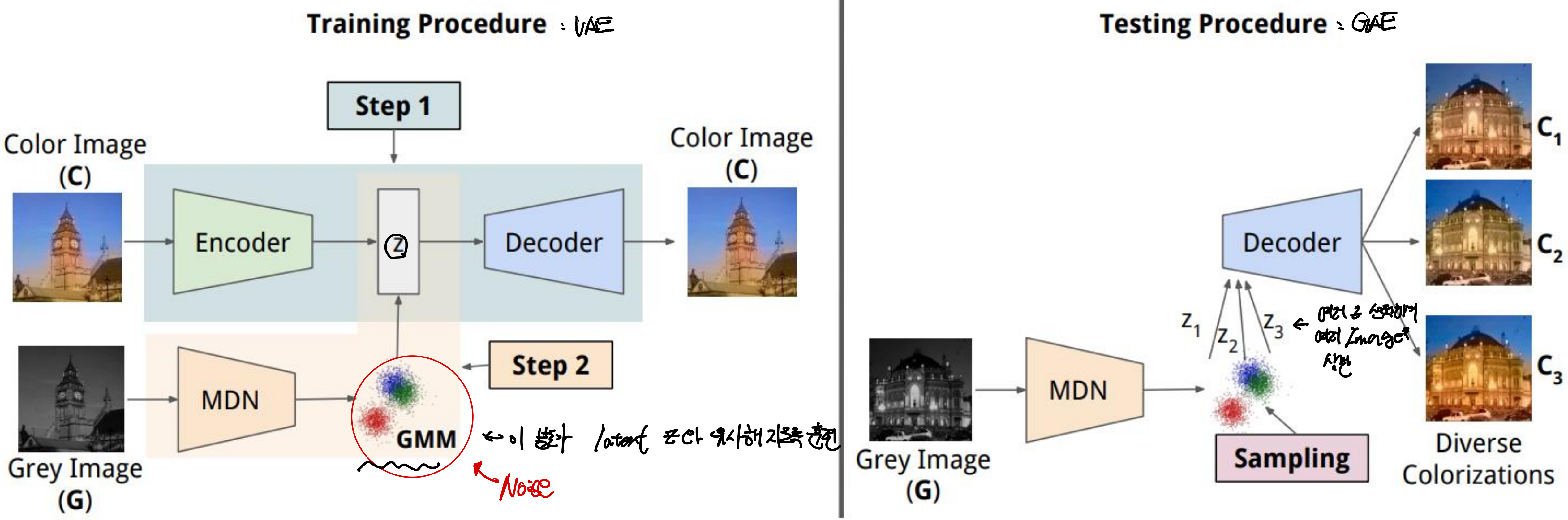
Ours GT

[1] Deshpande et al., Learning diverse image colorization. CVPR 2017.

More Recent Colorization Works

Details about those generative models?
Coming Soon.

Diverse Colorization (VAE or GAN) [1]



[1] Deshpande et al., Learning diverse image colorization. CVPR 2017.

More Recent Colorization Works

② Controllable

Language Based Colorization [1] ; 명령문 추가

1896. "Lightship Scotland." | *There is a white ship with two red sign on the sea.*

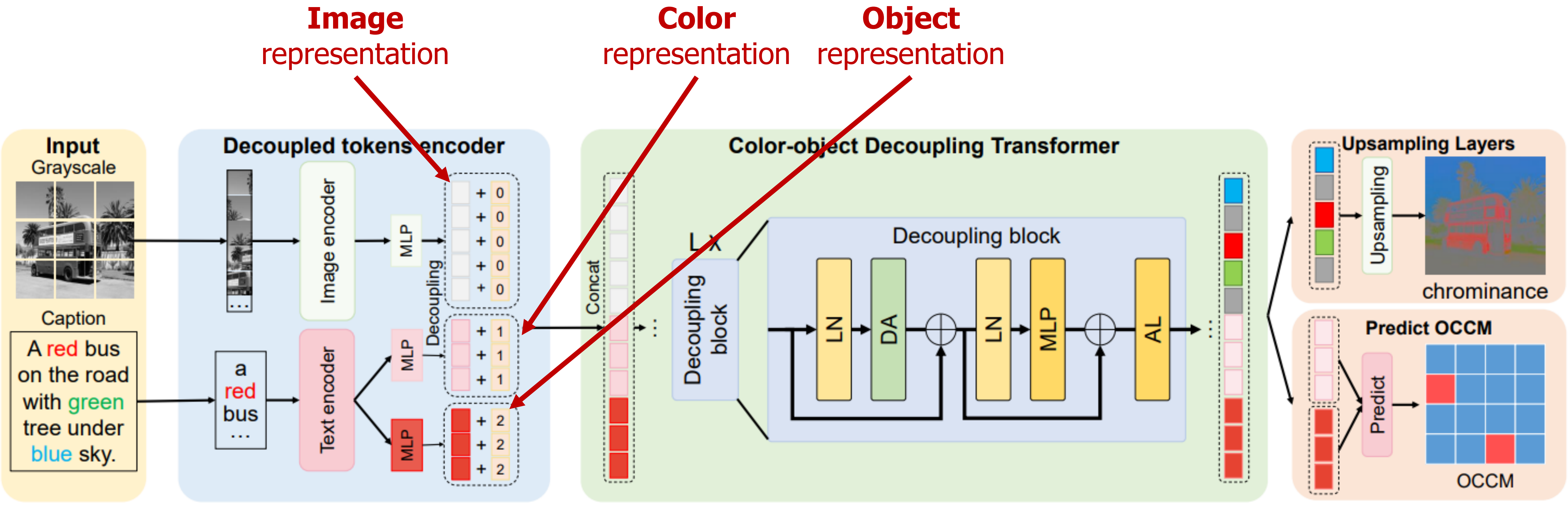
1937. "Dad jokes, the McNallys at dinner." | *Man in blue shirt with his family in front of tan wall.*

1940. "Old mine office. Virginia City, Nevada." | *A red house surrounded by yellow grass.*

[1] Chang et al., L-CoDer: Language-Based Colorization with Color-Object Decoupling Transformer. ECCV 2022.

More Recent Colorization Works

Language Based Colorization [1]



[1] Chang et al., L-CoDer: Language-Based Colorization with Color-Object Decoupling Transformer. ECCV 2022.

More Recent Colorization Works

Exemplar Based Colorization [1]

Target Reference 1 Result 1 Reference 2 Result 2 Reference 3 Result 3

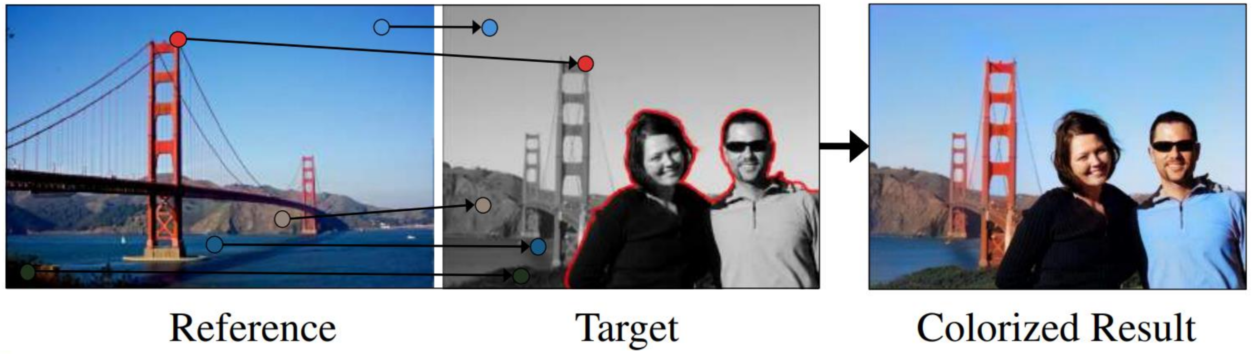
적용

Example

[1] He et al., Deep Exemplar-based Colorization. TOG 2018.

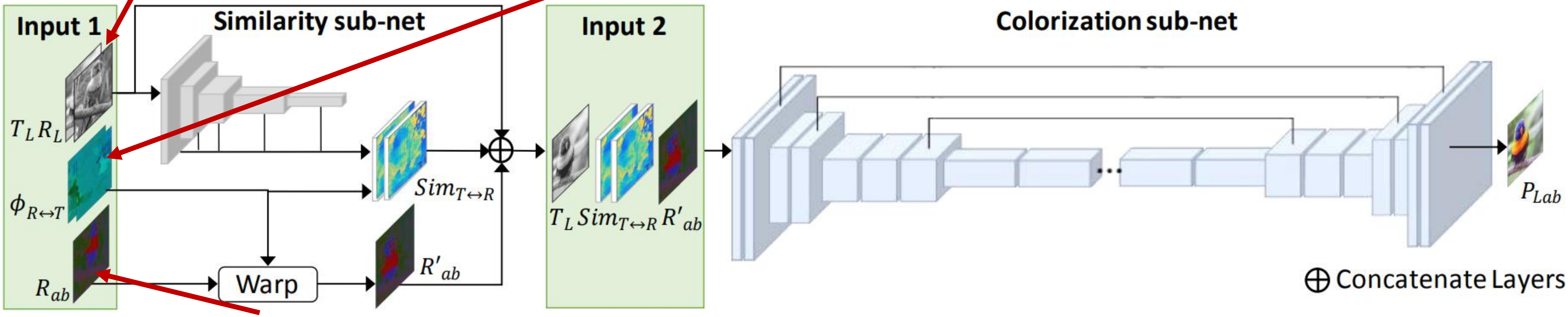
More Recent Colorization Works

Exemplar Based Colorization [1]



Intensity of
Target / Reference Images

Correspondence
between R&T



Chrominance (a,b) of the Reference

[1] He et al., Deep Exemplar-based Colorization. TOG 2018.

ITE4052 Computer Vision

3D Vision

Dong-Jin Kim

Spring 2025