An image and its colours: which colour matters?

Dr Pietro Giudice
(research associate, Influencerdb)

Data Science Meetup Muenster
07/June/2017

The problem

- find the (dominant) colors present in an image:

- the amount of color
- name

Why are we interested in this problem?

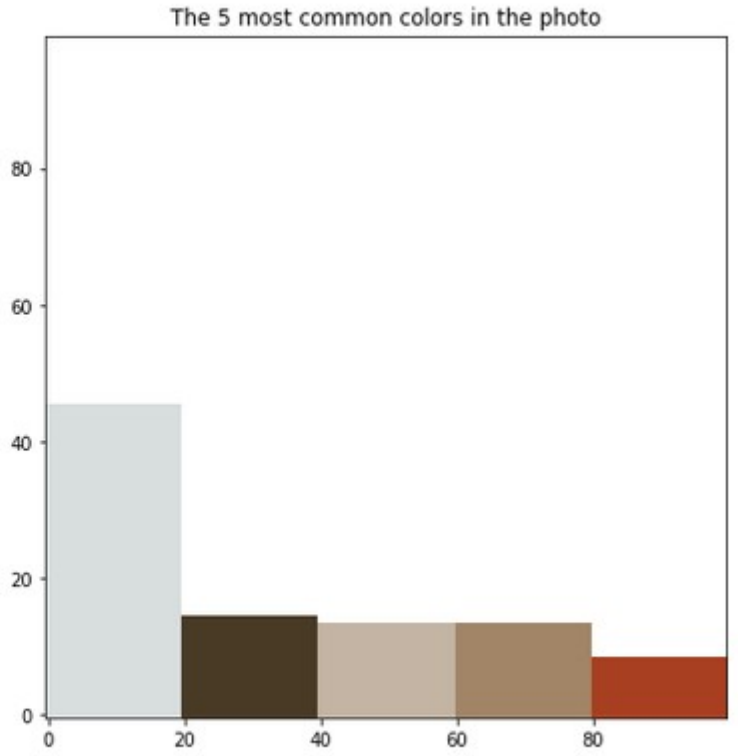
- search for a specific environment (snow, desert, sky, …)
- in a photo-sharing social networking : correlation with social metrics

How do we tackle the problem?

- average color of all pixels
- color quantization (well known in computer vision)
- k-means clustering (k is fixed, not deterministic)

An example


0
100
200
300
400
500
0
50
100
150
200
250
300
350
400


The 5 most common colors in the photo
0
20
40
60
80
0
20
40
60
80

Color names

So far we just determined clusters of pixels and we determined their number in each cluster
BUT: what is the name of he colors ?!!?

```
import webcolors
import matplotlib.colors

len(matplotlib.colors.XKCD_COLORS) #949
#'xkcd:cloudy blue', 'xkcd:dark pastel green', 'xkcd:dust', 'xkcd:electric lime', ...

len(webcolors.css3_hex_to_names)  #138
#'aliceblue', 'antiquewhite', 'cyan', 'aquamarine', 'azure', 'beige', 'bisque', 'black',...

my definition: #10
#white, black, gray, blue, green, yellow, brown, orange, red, pink.
```

Here there is a problem: concept of distance of colors!

Distance between two colors

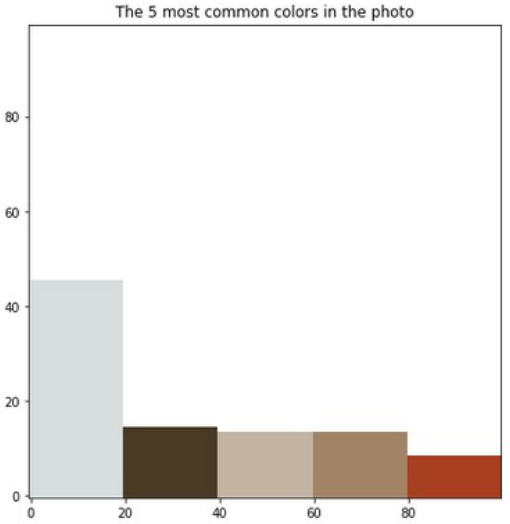
There have been many attempts to weight RGB values to better fit human perception

Different possibilities:

- Euclidean $\sqrt{(R_2 - R_1)^2 + (G_2 - G_1)^2 + (B_2 - B_1)^2}$
- Rule 2-4-3 $\sqrt{2 \times \Delta R^2 + 4 \times \Delta G^2 + 3 \times \Delta B^2}$
- The International Commission on Illumination (CIE):

```
from colormath.color_diff import delta_e_cie1976, delta_e_cie1994,
delta_e_cie2000
```

Let’s give a name!



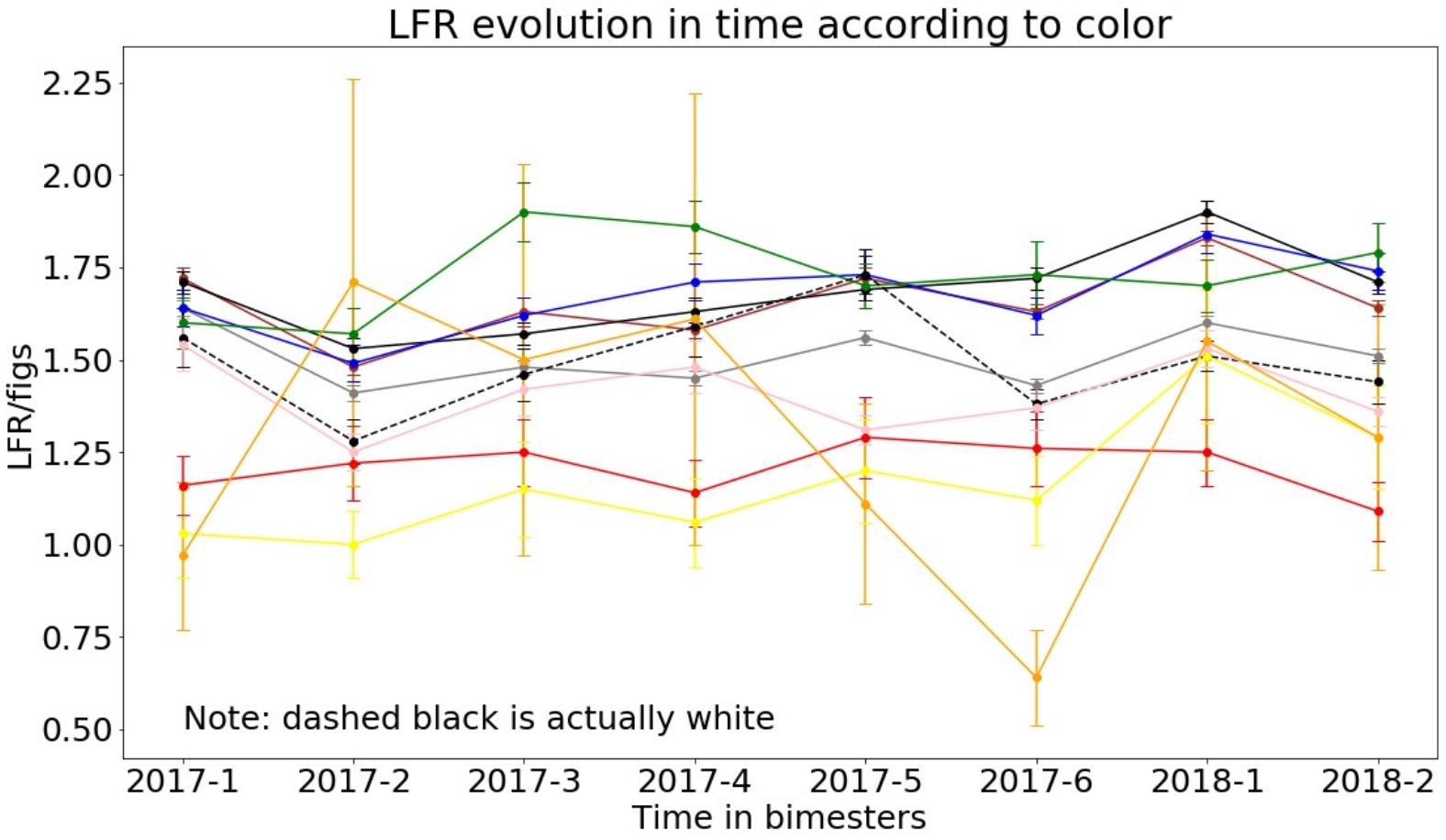

```
Clusters: 0.46, ( 216, 221, 222),   azure                  xkcd:cool grey          gray296
Clusters: 0.15, (  72,  59,  37),   darkgoldenrod          xkcd:mud brown          brown45
Clusters: 0.14, ( 162, 133, 106),   peru                   xkcd:pale brown         brown4
Clusters: 0.14, ( 197, 181, 165),   linen                  xkcd:mushroom           gray345
Clusters: 0.10, ( 170,  66,  34),   sienna                 xkcd:rusty red          brown40

All colors [('gray', 0.46), ('brown', 0.15), ('brown', 0.14), ('gray', 0.14), ('brown', 0.1)]

Sum Colors [('gray', 0.6), ('brown', 0.39)]

Colour dominant:  gray
```

Color in time



Thanks!

Thanks to InfluencerDB for data!

- My email: pietro.giudice@libertymail.net
- You can find me on Xing and Linkedin!