

Vectors

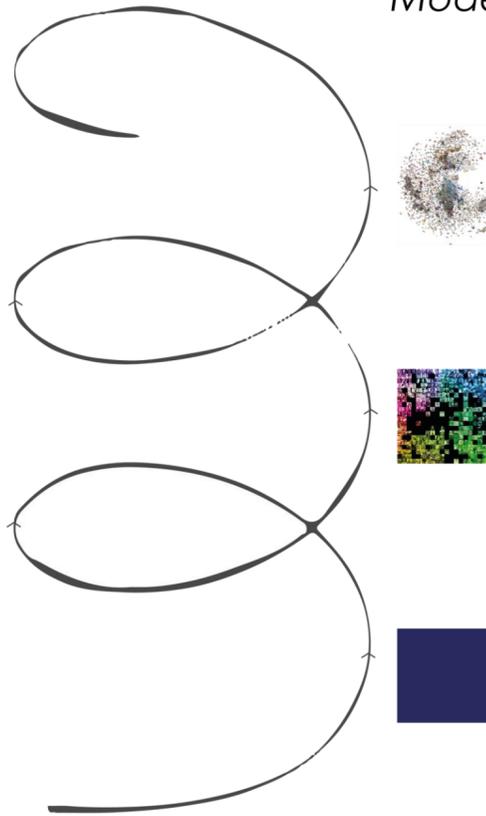
Map

Model

{72},{73},{172},{230},
{273},{274},{275},{276},
{277},{278},{279},{294},
{295},{296},{438},{519},
{520}

0101010101
0010101010
1001010101
010101

41°24'12.2N
2°10'26.5 E



VECTORS

A vector is a list of numbers

Vectors

$$\vec{a} = (2,3)$$

An example in 2 dimensions

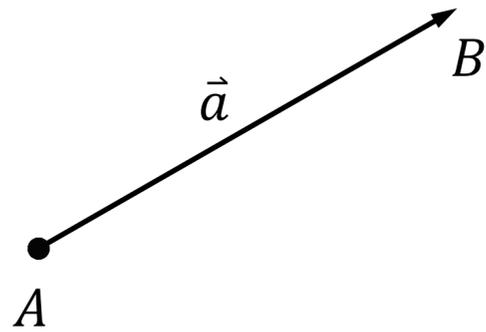
$$\vec{a} = (a_1, a_2, \dots, a_n)$$

Generalize to n-dimensional Euclidean space

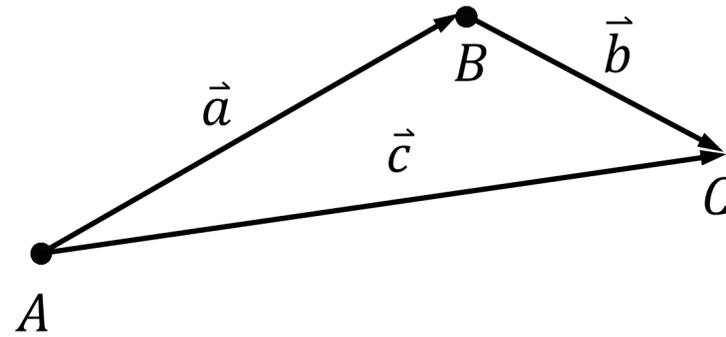
$$\|\vec{a}\| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}$$

Magnitude (length, or norm) of a vector

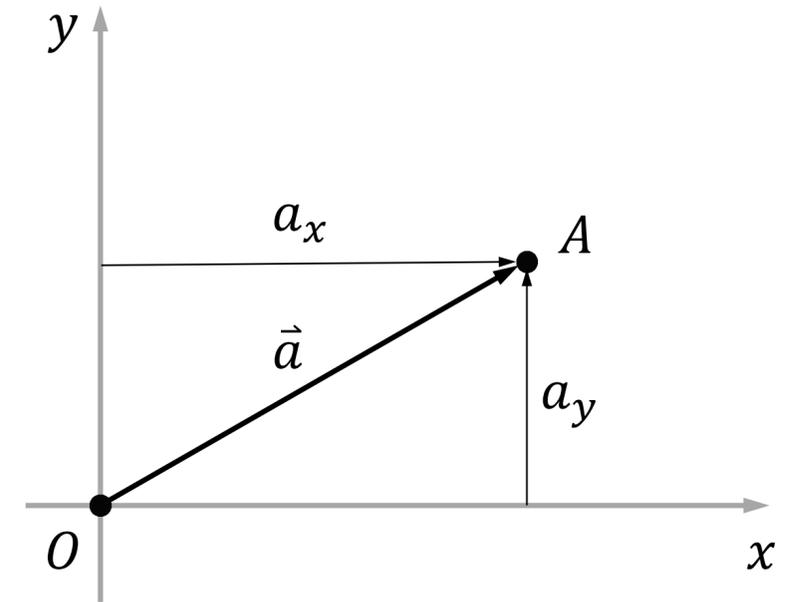
Vectors in geometry



A vector pointing from point A to B



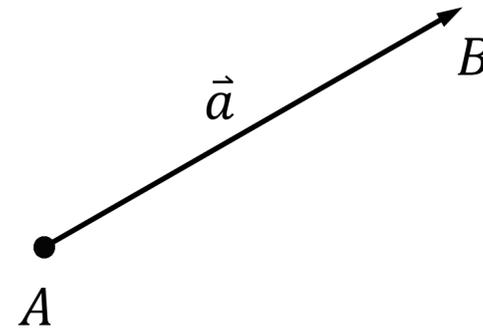
Addition and subtraction



Represent the coordinate of a point in Cartesian coordinate system

Euclidean distance

distance between two points in Euclidean space



$$\|AB\| = \sqrt{(B_1 - A_1)^2 + (B_2 - A_2)^2 + \dots + (B_n - A_n)^2} = \|\vec{a}\|$$

FEATURE VECTOR

In the context of machine learning vectors are called feature vectors as each of these values corresponds to some features

Examples of Feature Vector

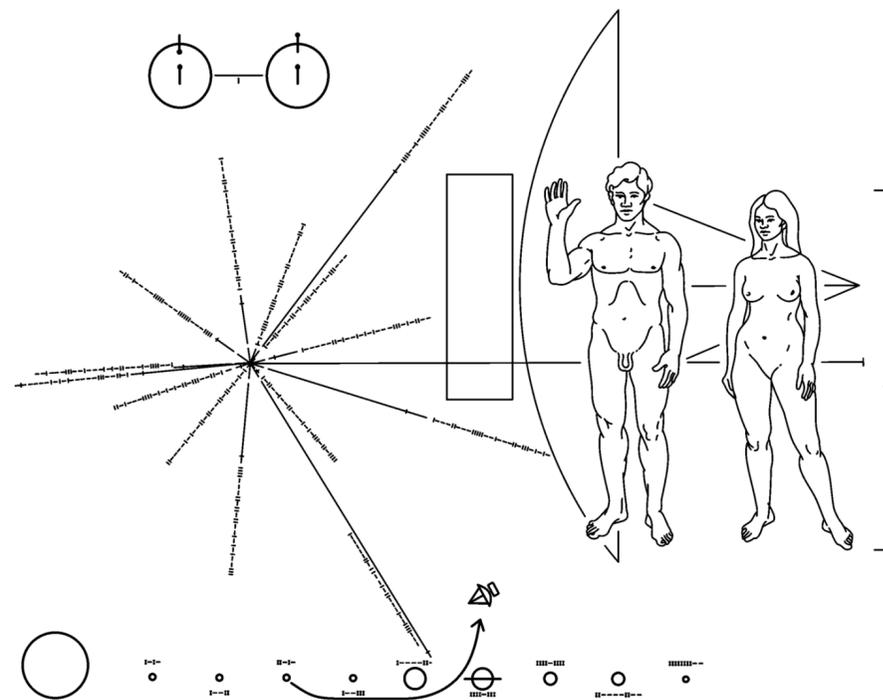
Numerical features that represent some object



An apple

(color, size, weight, sweetness)

(0 red, 12.3 cm, 180 g, 2 very)



(sex, age, weight, height)

(0 male, 25, 70 kg, 182 cm)

@Pioneer plaque

A person

Feature Vector + Euclidean Distance

A Common Ground for Comparing Objects



A

(color, size, weight, sweetness)

(0 red, 12.3 cm, 180 g, 2 very)



B

(color, size, weight, sweetness)

(1 orange, 11.7 cm, 170 g, 1 yes)

$$\|AB\| = \sqrt{(1 - 0)^2 + (11.7 - 12.3)^2 + (170 - 180)^2 + (2 - 1)^2} = 10.1173..$$

Cross Dimensionality

Communication between Euclidean spaces of different dimensionality

$$\begin{bmatrix} a_1 & a_2 \end{bmatrix} \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \end{bmatrix} = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix}$$



(color, size, weight, sweetness)

(0 red, 12.3 cm, 180 g, 2 very)



(price, do I like it?)

(0.75 Fr, 1 yes)

There are operations allow us to project from one dimensionality to another

Map and Models



(color, size, weight, sweetness)

(0 red, 12.3 cm, 180 g, 2 very)

Features

Values

Model

Map

Map and Models

Pixel values (obtained by digital camera)

255	127	...	133
132	86	...	144
...
178	254	...	233

(0 red, 12.3 cm, 180 g, 2 very)

Map

Digital photograph



(color, size, weight, sweetness)

Model

FEATURE VECTOR FOR DIGITAL OBJECTS

Models of Map

Images and Texts

Models that talk about the word

The image shows a screenshot of a Twitter search page for the hashtag #zurich. The interface includes a search bar at the top with the query #zurich, and buttons for 'Log in' and 'Sign up'. The main content area displays a list of tweets:

- Parkplatz Online Zürich Oerlikon** (@ParkplatzO) - 2h: Cheers! bleibt Gesund und zu Hause, damit wir bald wieder zum Alltag zurückkehren können! #oerlikon #binzmühlestrasse #zürichnord #parkplatzsuche #onlineparkplatz #zurich #basel #bern #stgallen #schaffhausen #luzern... [instagram.com/p/B-ett1Oji2C/...](#)
- mlnang** (@habicucu) - 2h: Die Situation im #Tessin ist ernst und schwierig, aber auch in anderen Kantonen wie #Vaud #Valais #Graubünden, #Genf #Zurich. Jetzt bleiben wir alle in unserem Kanton zu Hause. Es wird nicht für immer sein! Bis bald und Danke! #ZuhauseBleiben #stiamoacasa #RestonsChezNous
- Tactica Games** (@TacticaGames) - 2h: Wir haben noch viele Rollenspielbücher. D&D5e, DSAS, Pathfinder, Starfinder, noch viele mehr und in deutsch & englisch! Viel Spass beim einkaufen 😊 [tacticagames.ch/shop/ki/Rollen...](#) #rpg #roleplay #miniatures #books #bleibzu Hause #pathfinder #DnD #dungeonsanddragons #zurich #starfinder
- Ruby Carli** (@environ_2020) - 2h: #shape the future of #earth @environ_2020 in #zurich #Switzerland #ClimateChange #pollution #globalwarming #environment #RenewableEnergy #conferencecall
- Felere.com** (@FelereCom) - 3h: La mode vestimentaire est la manière de se vêtir [felere.com/vetement-moins](#) ... Vêtement robe de nuit pour femme #Felere #Mode #Shopping #femmes #Zurich #Genève #Bâle #Lausanne #Berne #Olten #Köniz

Below the tweets, there is a 'New to Twitter?' section with a 'Sign up' button. To the right, there are 'Search filters' for People (From anyone, People you follow) and Location (Anywhere, Near you). Below that is a 'Trends' section listing trending topics: 1. News - Trending **Afrique** (62.8K Tweets), 2. News - Trending **#webinar** (8,550 Tweets), 3. Trending **tessin**, 4. Trending **Pâques** (4,299 Tweets), 5. Trending **Klar** (11.6K Tweets). At the bottom, there are links for 'Terms', 'Privacy policy', 'Cookies', 'Ads info', and 'More', along with the copyright notice '© 2020 Twitter, Inc.'

Feature Vectors for Digital Objects

Models of Map

255 127 ... 133
132 86 ... 144
...
178 254 ... 233

Image (Pixel values)

Map



Colors, Objects, Edges in the Image, ...

Model

4.3, 7.5, 2.2, 1.7, 0.8,

Feature Vector

Map

WHY ?

As images are already numbers, why do we convert them to some other numbers?

An Example with Images

Group by interior / exterior photos



?



?



```
EuclideanDistance [  
  Flatten@ImageData@ImageResize [#1, {32, 32}],  
  Flatten@ImageData@ImageResize [#2, {32, 32}]] & @@
```



Out[2]= 21.6916

```
EuclideanDistance [  
  Flatten@ImageData@ImageResize [#1, {32, 32}],  
  Flatten@ImageData@ImageResize [#2, {32, 32}]] & @@
```



Out[3]= 16.6223

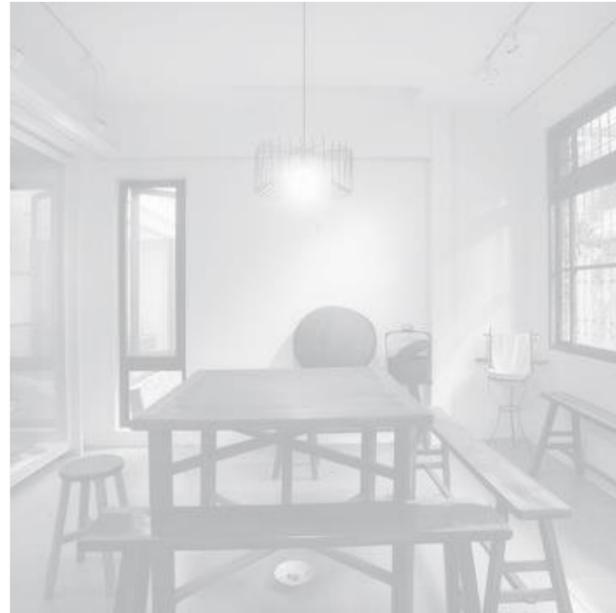
```
EuclideanDistance [  
  Flatten@ImageData@ImageResize [#1, {32, 32}],  
  Flatten@ImageData@ImageResize [#2, {32, 32}]] & @@
```



Out[4]= 19.1647

An Example with Images

Group by interior / exterior photos



Colors sometimes cannot get what we want

Similarly, we should not do word-by-word comparison for texts

FEATURE EXTRACTION

Computational mechanism for getting feature vectors



Non-probabilistic

Edge Detection

Fourier Transform

Hough Transform

...

Probabilistic (machine learning)

Object Detection

Image Captioning

...

To be, or not to be,--that is the
question:-- Whether 'tis nobler in the
mind to suffer The slings and arrows of
outrageous fortune Or to take arms against
a sea of troubles, And by opposing end them?

Word Count / Word Histogram

...

Word2Vec (Word Embedding)

Sentiment Analysis

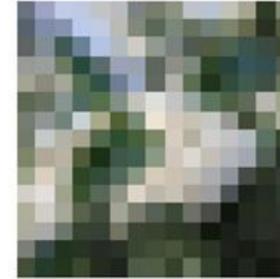
...

FEATURE EXTRACTION

Demonstration with the previous collected tweets



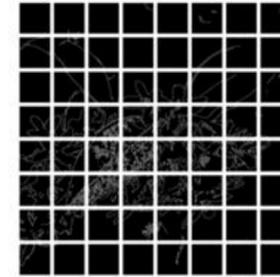
-> "colors"



```
1600  0  1600  0  1600
  0  1600  0  1600  0
1600  0  1516  84  1536
 64  1600  0  1562  38
1477 123 1494 106 1558
 42  1600  0  1600  0
1600  0  1600  0  1512
```



-> "edges"



```
0.439216 0.470588 0.541176
0.694118 0.733333 0.862745
0.780392 0.811765 0.898039
0.733333 0.741176 0.74902
0.462745 0.490196 0.458824
0.505882 0.533333 0.458824
0.662745 0.694118 0.701961
```



-> "feature extraction"

common privet

common jasmine

laurel

California laurel

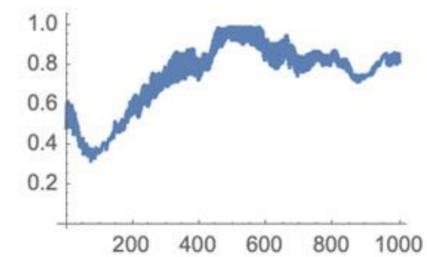
fruit tree

pride-of-rochester

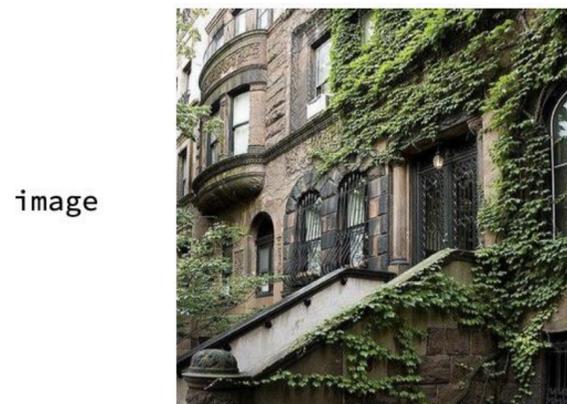
```
8.22561 -8.62608 28.0713
-3.54004 12.2321 3.18221
-6.41089 1.24627 -2.62163
```



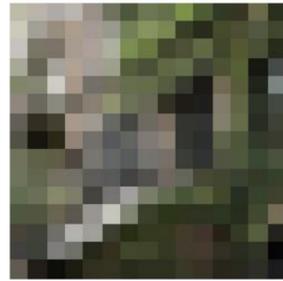
-> "fourier"



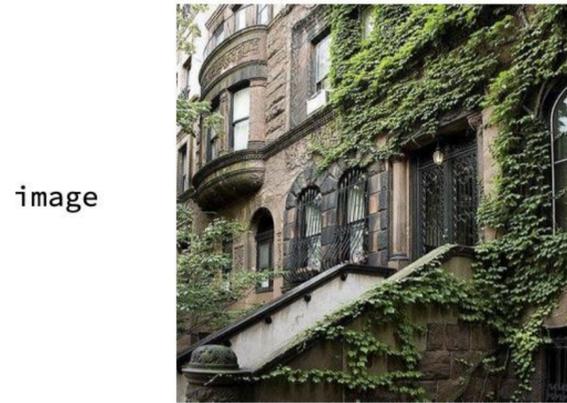
```
311.196 + 0. i      -2.8447 + 8.53134 i
-2.8447 - 8.53134 i  -18.8027 + 15.8672 i
0.0507741 - 1.67722 i -0.855818 - 0.620666 i
11.2859 + 32.356 i   2.20177 + 1.97652 i
1.98793 + 0.293197 i -2.49415 + 16.4691 i
1.08693 - 0.0374585 i -0.728638 - 0.404993 i
-7.28062 + 7.17637 i 0.616701 + 0.874254 i
```



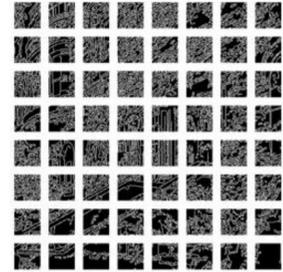
-> "colors"



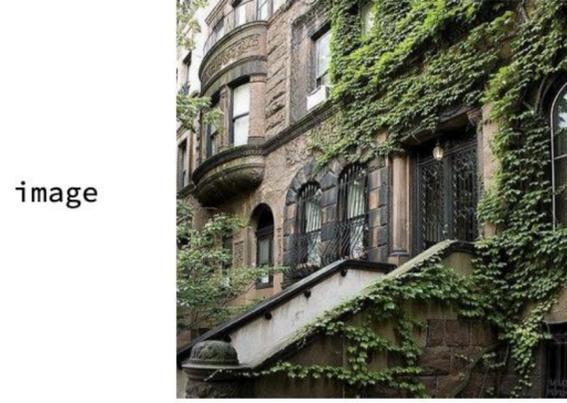
0.745098 0.741176 0.662745 0.694118 0.666667
0.635294 0.705882 0.658824 0.627451 0.494118
0.490196 0.423529 0.368627 0.427451 0.278431
0.447059 0.517647 0.313725 0.34902 0.423529
0.196078 0.203922 0.235294 0.133333 0.301961
0.305882 0.196078 0.254902 0.282353 0.172549



-> "edges"



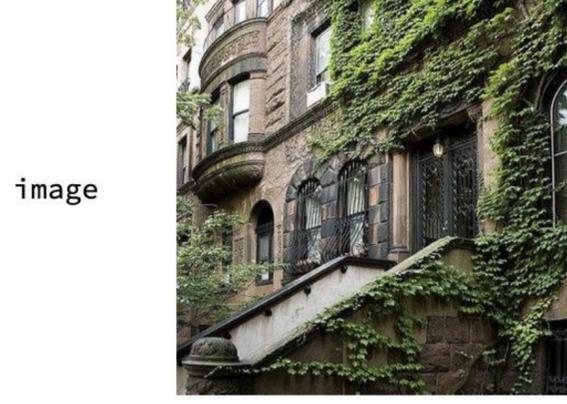
1146 454 1138
462 1057 543
1107 493 1064
536 991 609
1156 444 1204



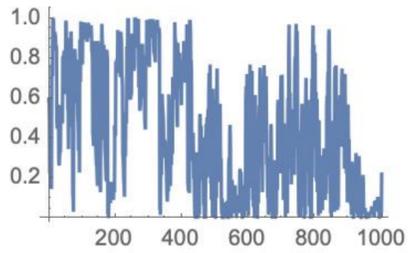
-> "feature extraction"

- Boston ivy
- vine
- vascular plant
- flora
- support
- stair

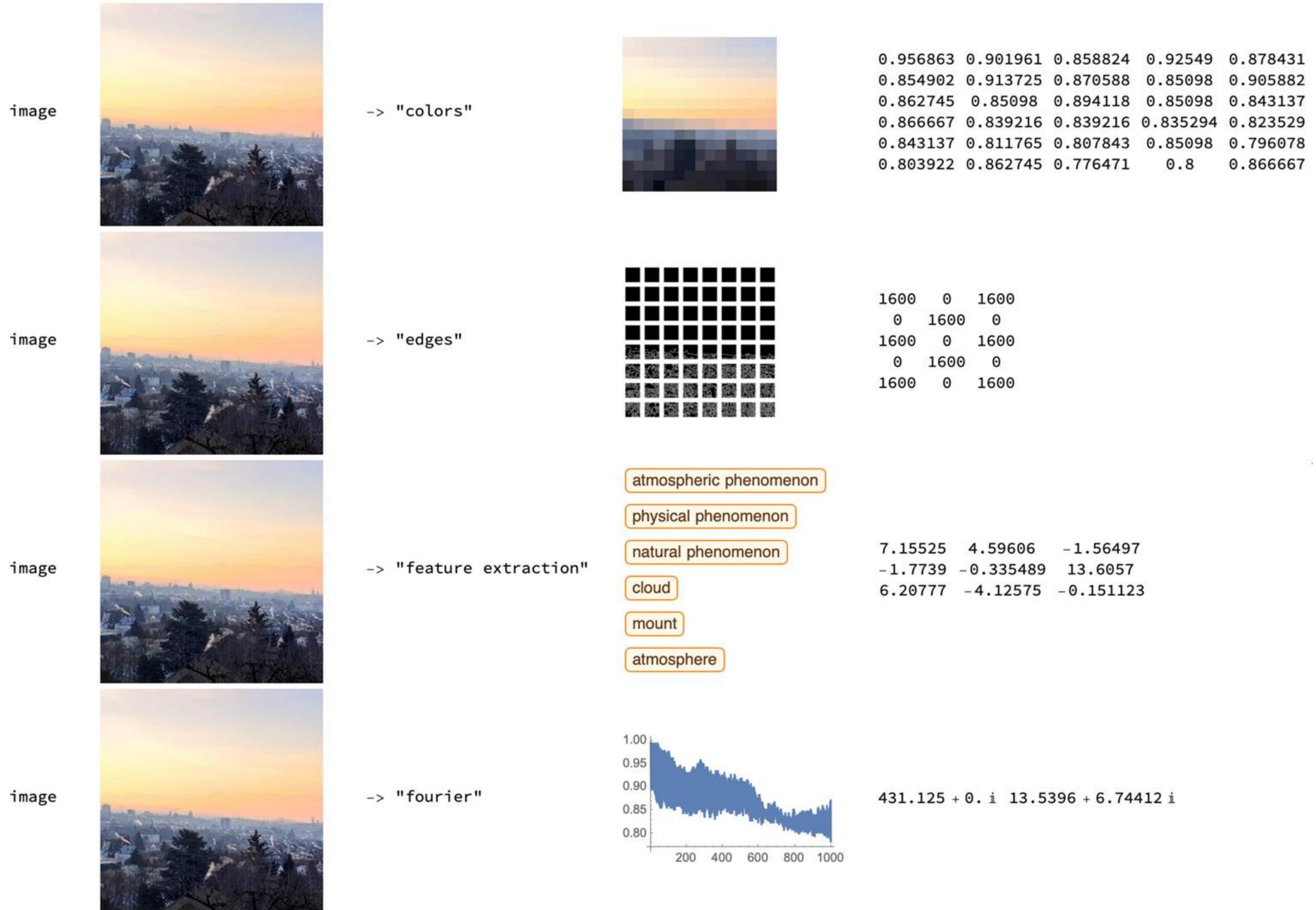
11.9924 5.44124 -2.30834
0.115535 2.4649 4.43479
5.21442 4.58633 -3.29425

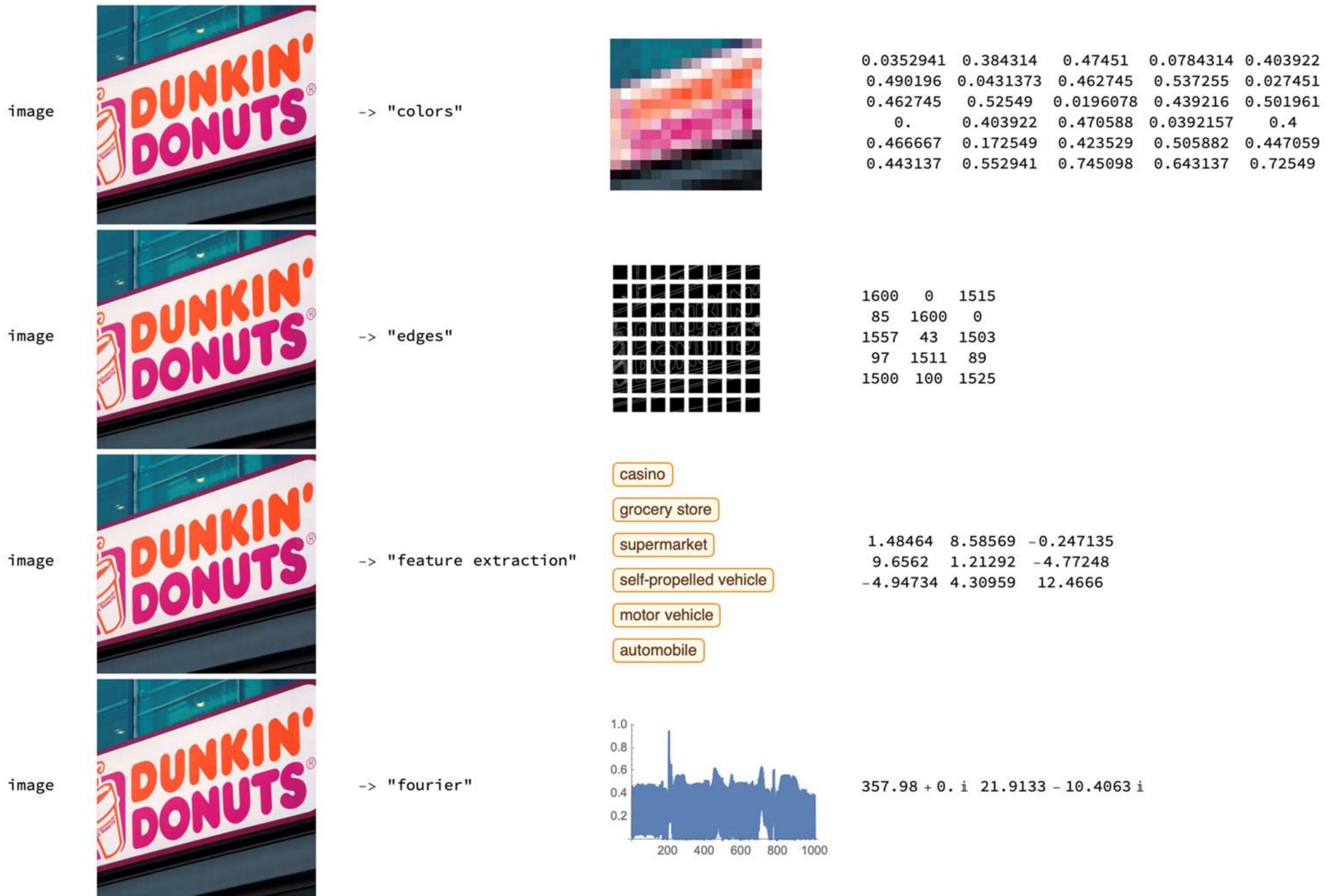


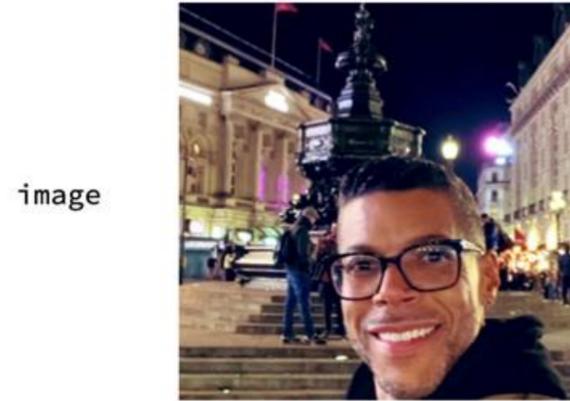
-> "fourier"



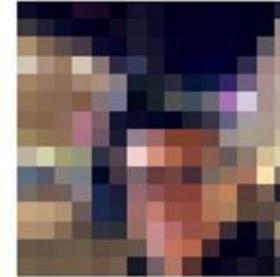
232.95 + 0. i 6.82914 + 15.0475 i



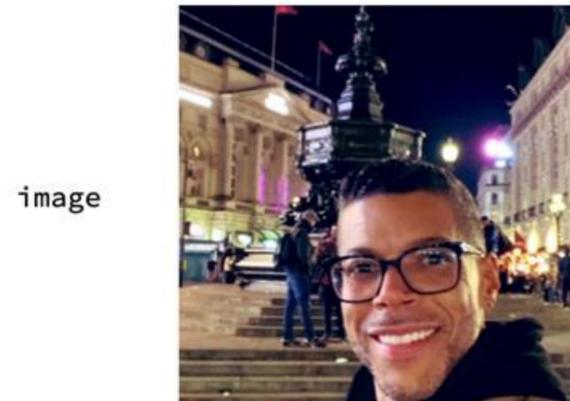




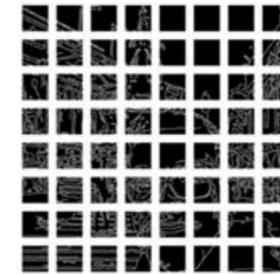
-> "colors"



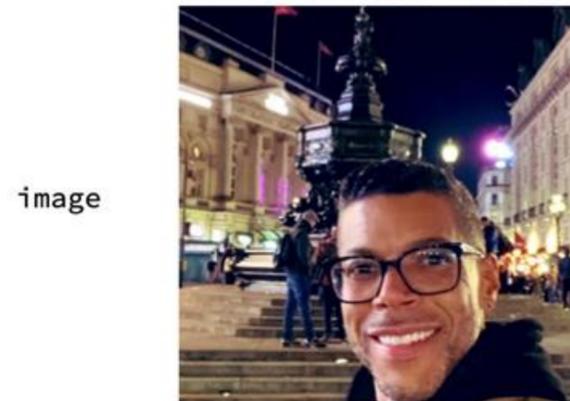
```
0.0588235 0.0627451 0.156863 0. 0.
0.117647 0.0666667 0.0196078 0.133333 0.0235294
0.00784314 0.117647 0.0666667 0.0745098 0.168627
0.0117647 0.0117647 0.12549 0.00392157 0.
0.121569 0.0117647 0.00784314 0.129412 0.
0. 0.0941176 0.160784 0.172549 0.239216
```



-> "edges"



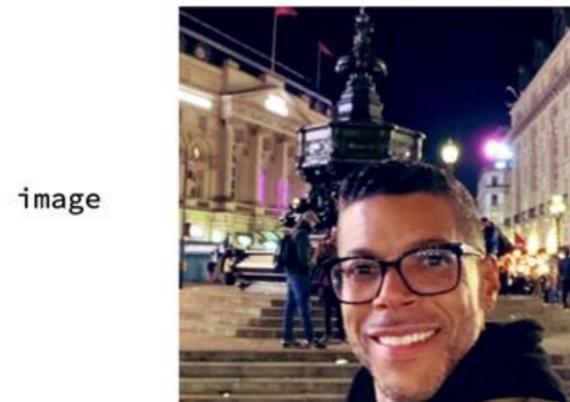
```
1413 187 1525
75 1477 123
1585 15 1432
168 1600 0
1600 0 1600
```



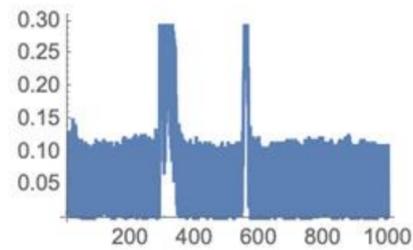
-> "feature extraction"

- person
- hominid
- primate
- mammal
- vertebrate
- anatomy

```
8.05123 -1.32554 -1.67072
-4.65785 3.92969 2.48092
1.45389 -6.90618 2.5591
```

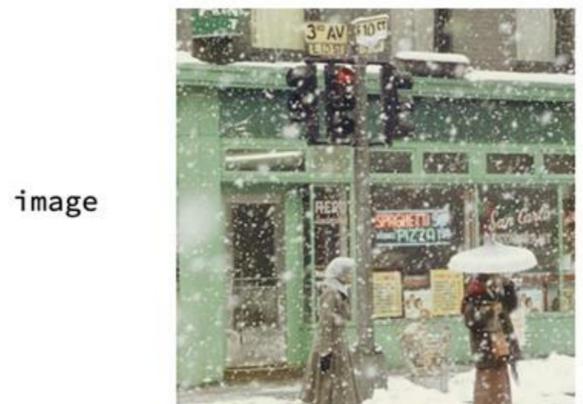


-> "fourier"

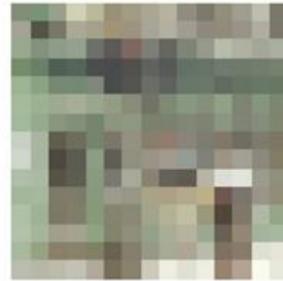


```
234.955 + 0. i 14.517 - 2.76679 i
```

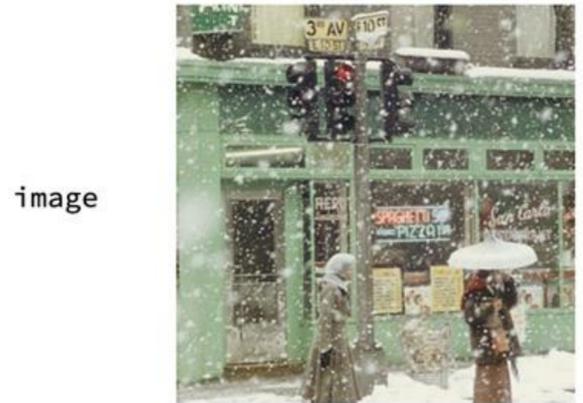
image		-> "colors"		<pre> 0.286275 0.0235294 0.0156863 0.337255 0.0196078 0.0156863 0.384314 0.00784314 0.0117647 0.443137 0.0117647 0.0117647 0.294118 0.0117647 0.0156863 0.407843 0.0156863 0.0117647 0.396078 0.00392157 0.0117647 0.227451 0.00392157 0.0117647 0.152941 0.0117647 0.00784314 0.117647 0.00784314 0.00784314 </pre>
image		-> "edges"		<pre> 1597 3 1538 62 1600 0 1581 19 1594 6 1573 27 1600 0 1600 </pre>
image		-> "feature extraction"	<ul style="list-style-type: none"> laser primate hominid person light source fairy light 	<pre> 10.2982 -1.38993 9.04966 -3.65853 -4.3973 -4.61546 -9.81738 -16.1137 1.08928 </pre>
image		-> "fourier"		<pre> 141.403 + 0. i 74.3727 + 10.9785 i </pre>



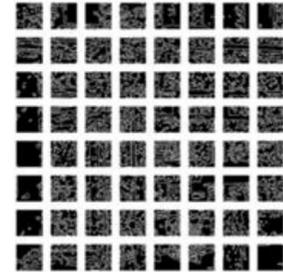
-> "colors"



```
0.513725 0.556863 0.470588 0.52549 0.560784
0.478431 0.784314 0.776471 0.662745 0.65098
0.627451 0.513725 0.639216 0.619608 0.533333
0.721569 0.709804 0.643137 0.568627 0.552941
0.501961 0.533333 0.509804 0.443137 0.682353
0.67451 0.619608 0.54902 0.537255 0.482353
```



-> "edges"



```
1281 319 1216
384 1570 30
1187 413 1254
346 1310 290
1343 257 1562
```



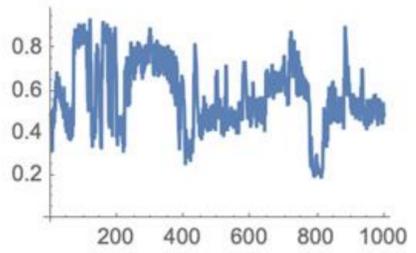
-> "feature extraction"

- sprinkler
- crucifix
- lock
- padlock
- cockpit
- abacus

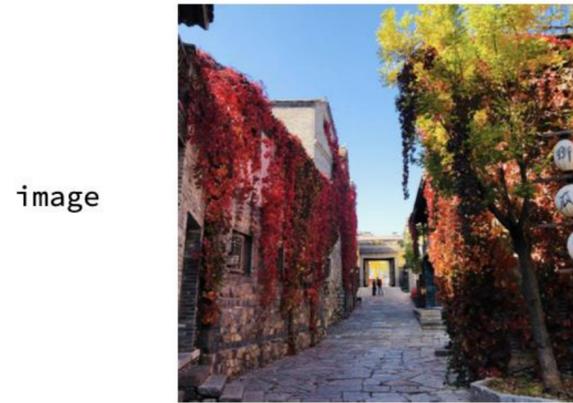
```
10.2982 -1.38993 9.04966
-3.65853 -4.3973 -4.61546
-9.81738 -16.1137 1.08928
```



-> "fourier"



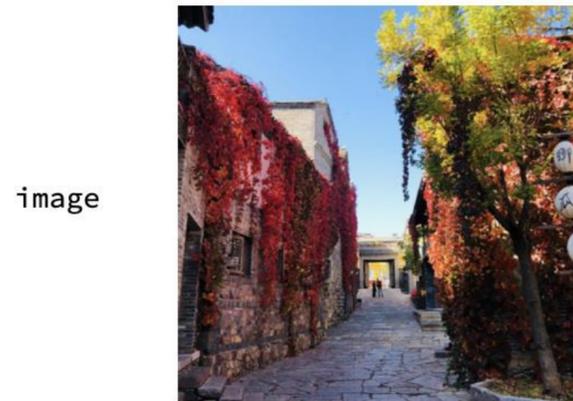
```
381.644 + 0. i 5.46611 + 14.145 i
```



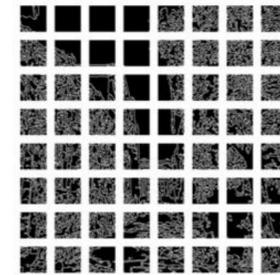
-> "colors"



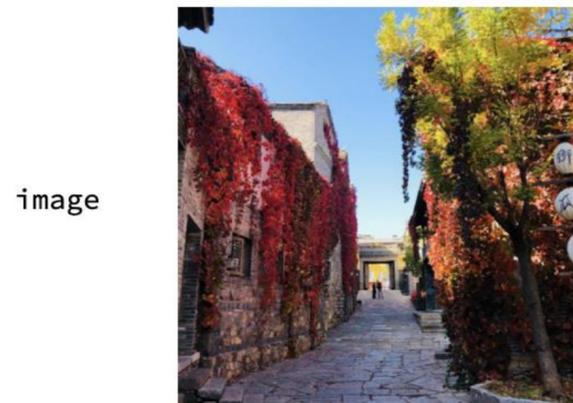
```
0.298039 0.392157 0.494118 0.541176 0.792157
1.      0.498039 0.737255 0.94902 0.454902
0.713725 0.956863 0.45098 0.709804 0.917647
0.666667 0.74902 0.560784 0.737255 0.713725
0.231373 0.713725 0.717647 0.172549 0.690196
0.592157 0.141176 0.592157 0.572549 0.486275
```



-> "edges"



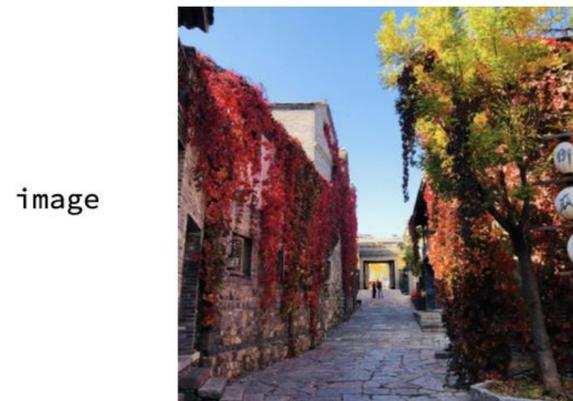
```
1477 123 1600
0 1600 0
1600 0 1596
4 1275 325
1155 445 1080
```



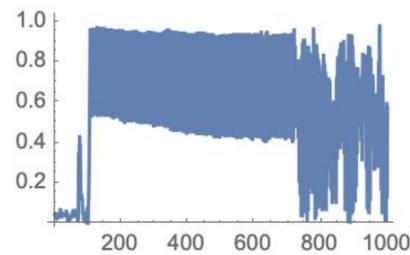
-> "feature extraction"

- Boston ivy
- tree
- flowering tree
- maple tree
- boxelder
- California box elder

```
8.05123 -1.32554 -1.67072
-4.65785 3.92969 2.48092
1.45389 -6.90618 2.5591
```



-> "fourier"



```
258.052 + 0. i 15.2118 + 2.63635 i
```

text	To be, or not to be,--that is the question:-- Whether 'tis nobler in the mind to suffer The slings and arrows of outrageous fortune Or to take arms against a sea of troubles, And by opposing end them?	-> "Sentiment Analysis"	Qualifying text by positive, neutral or negative	0.232578	0.908024	0.0687178		
				0.68047	-0.039263	0.30186	-0.17792	0.42962
				0.032246	-0.41376	0.13228	-0.29847	-0.085253
				0.17118	0.22419	-0.10046	-0.43653	0.33418
				0.67846	0.057204	-0.34448	-0.42785	-0.43275
text	To be, or not to be,--that is the question:-- Whether 'tis nobler in the mind to suffer The slings and arrows of outrageous fortune Or to take arms against a sea of troubles, And by opposing end them?	-> "Word2Vec"	Qualifying text by it's content with word-embeddings	0.55963	0.10032	0.18677	-0.26854	0.037334
				-2.0932	0.22171	-0.39868	0.20912	-0.55725
				3.8826	0.47466	-0.95658	-0.37788	0.20869
				-0.32752	0.12751	0.088359	0.16351	-0.21634
				-0.094375	0.018324	0.21048	-0.03088	-0.19722
				0.082279	-0.09434	-0.073297	-0.064699	-0.26044

text	<p>so long life; For who would bear the whips and scorns of time, The oppressor's wrong, the proud man's contumely, The pangs of despis'd love, the law's delay, The insolence of office, and the spurns That patient merit of the unworthy takes, When he himself might his quietus make With a bare bodkin</p>	-> "Sentiment Analysis"	Qualifying text by positive, neutral or negative	0.232578	0.908024	0.0687178			
				0.68047	-0.039263	0.30186	-0.17792	0.42962	
				0.032246	-0.41376	0.13228	-0.29847	-0.085253	
				0.17118	0.22419	-0.10046	-0.43653	0.33418	
				0.67846	0.057204	-0.34448	-0.42785	-0.43275	
text	<p>so long life; For who would bear the whips and scorns of time, The oppressor's wrong, the proud man's contumely, The pangs of despis'd love, the law's delay, The insolence of office, and the spurns That patient merit of the unworthy takes, When he himself might his quietus make With a bare bodkin</p>	-> "Word2Vec"	Qualifying text by it's content with word-embeddings	0.55963	0.10032	0.18677	-0.26854	0.037334	
				-2.0932	0.22171	-0.39868	0.20912	-0.55725	
				3.8826	0.47466	-0.95658	-0.37788	0.20869	
				-0.32752	0.12751	0.088359	0.16351	-0.21634	
				-0.094375	0.018324	0.21048	-0.03088	-0.19722	
				0.082279	-0.09434	-0.073297	-0.064699	-0.26044	

text	who would these fardels bear, To grunt and sweat under a weary life, But that the dread of something after death,-- The undiscover'd country, from whose bourn No traveller returns,--puzzles the wil	-> "Sentiment Analysis"	Qualifying text by positive, neutral or negative	0.232578	0.908024	0.0687178		
				0.68047	-0.039263	0.30186	-0.17792	0.42962
				0.032246	-0.41376	0.13228	-0.29847	-0.085253
				0.17118	0.22419	-0.10046	-0.43653	0.33418
				0.67846	0.057204	-0.34448	-0.42785	-0.43275
text	who would these fardels bear, To grunt and sweat under a weary life, But that the dread of something after death,-- The undiscover'd country, from whose bourn No traveller returns,--puzzles the wil	-> "Word2Vec"	Qualifying text by it's content with word-embeddings	0.55963	0.10032	0.18677	-0.26854	0.037334
				-2.0932	0.22171	-0.39868	0.20912	-0.55725
				3.8826	0.47466	-0.95658	-0.37788	0.20869
				-0.32752	0.12751	0.088359	0.16351	-0.21634
				-0.094375	0.018324	0.21048	-0.03088	-0.19722
				0.082279	-0.09434	-0.073297	-0.064699	-0.26044