Image retrieval with color co-occurrence matrices

Maja Temerinac

Seminar CBIR
University of Freiburg

November 26th, 2004

Image retrieval with color co-occurrence matrices

1. Introduction

2. Texture: -human description (TEXRET-System)
   -hierarchical (multiscale space)

3. Color: -rgb, hsv, yuv

4. Definition: co-occurrence Matrix

5. The original 14 features

6. Testing results

7. Outlook: -minimal set of features
   -other color spaces
   -motif co-occurrence matrices

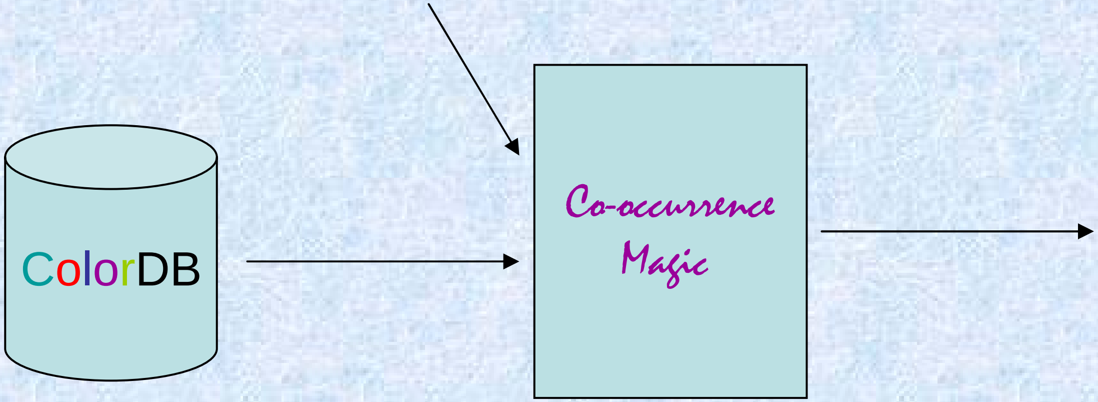
1.Introduction

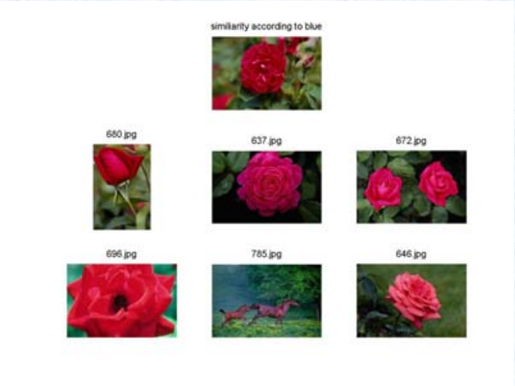




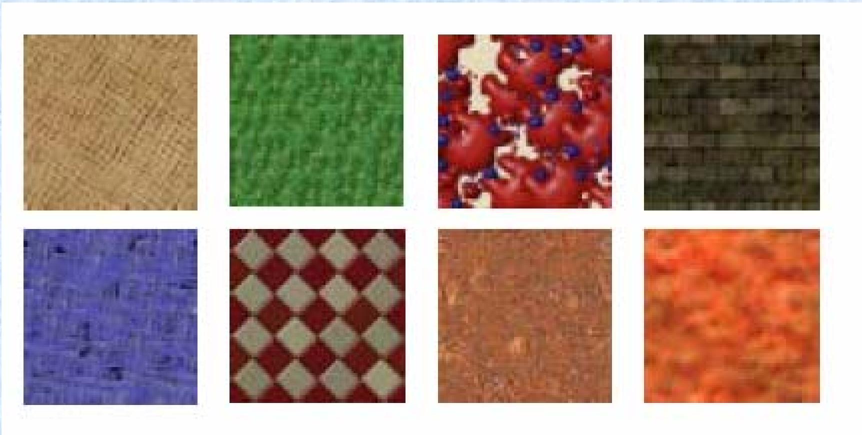






2. Texture: -human description (TEXRET-System)



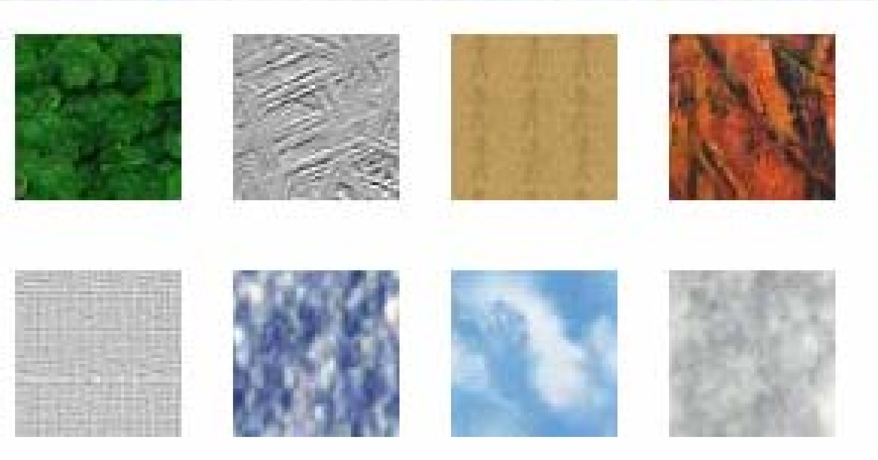


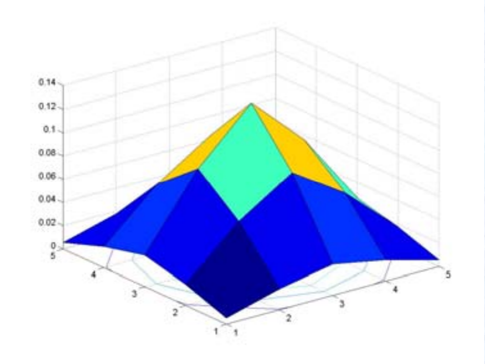
Table 1.Characteristics used for Survey II.

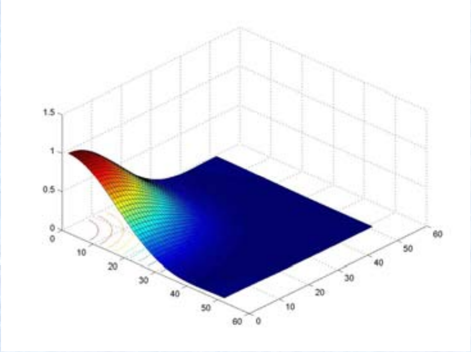
Identifier	Characteristic
1	homogeneous / non homogeneous
2	geometrical / non geometrical
3	pleasant / non pleasant
4	tasty / insipid
5	soft / rough
6	fine / coarse
7	fragile / robust
8	with lines / without lines
9	flat / non flat
10	happy / sad
11	regular / irregular
12	symmetrical / non symmetrical
13	with circles / without circles
14	loud / quiet
15	periodical / non periodical
16	clear / dark
17	simple / complex
18	transparent / non transparent
19	soluble / non soluble
20	natural / artificial
21	defined / diffuse

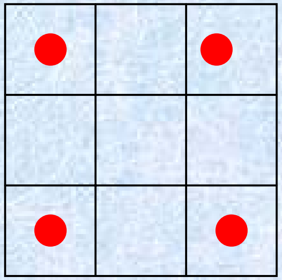
2. Texture: -hierarchical (multiscale space)

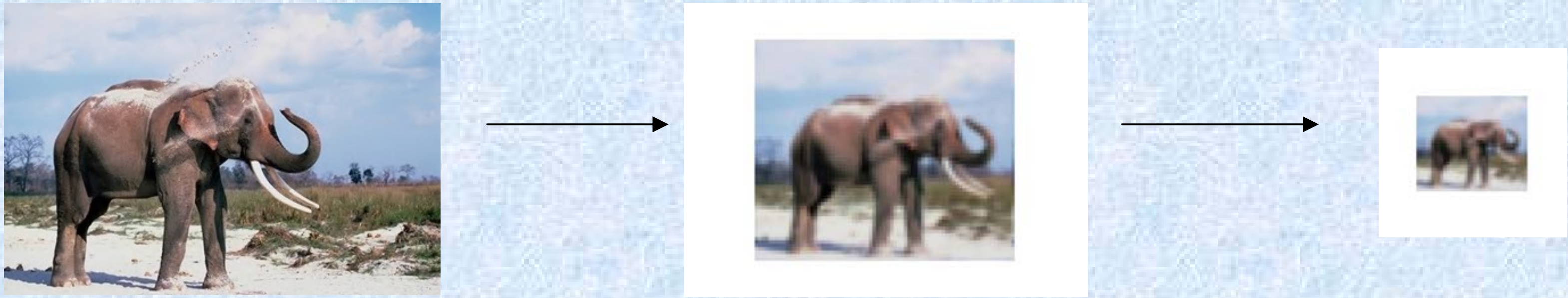
Texture can be regarded as a hierarchical pattern because a characteristic structure can be part of a larger structure that again may be periodic (Metzler,Palm,Lehmann,Aach 2002)

Gauss filter + resize



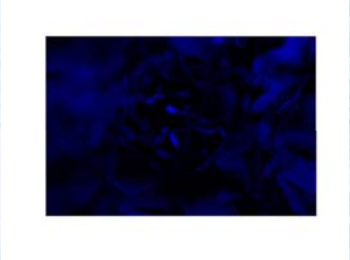




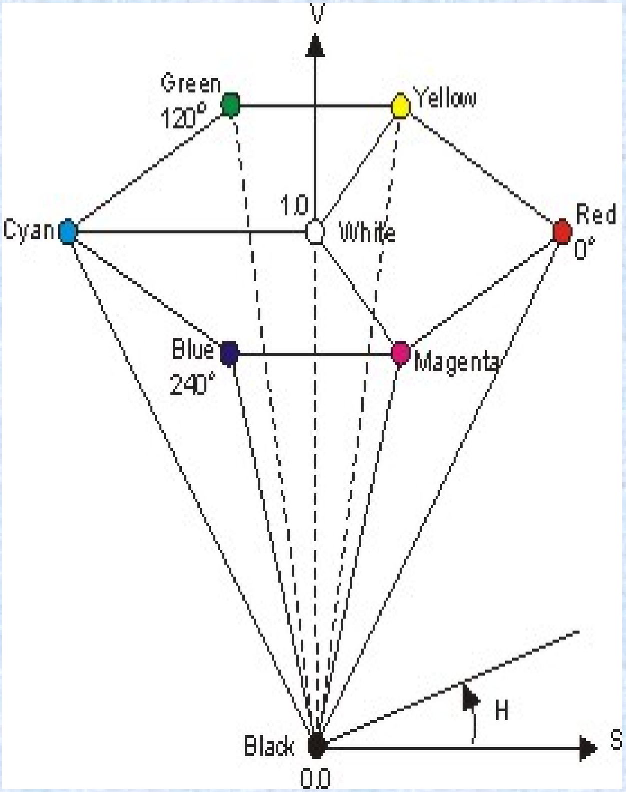


3. Color: -rgb, hsv, yuv

RGB :[0,…,255]



HSV :[0,…,1]



YUV

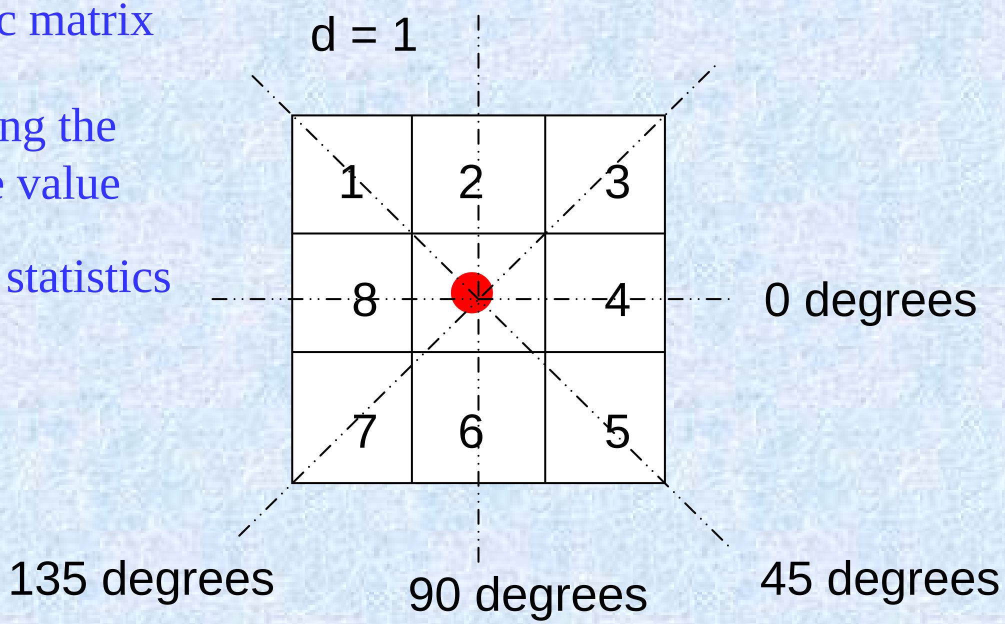
$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.148 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

- creates a black and white image (luma) from the full color image and then subtracts the three primary colors resulting in two additional signals to describe color

- U and V channels subtract the Luminance values from Red (U) and Blue (V) to reduce the color information

4. Definition: co-occurrence Matrix (spatial-dependence matrix)

- Computes the co-occurrence of gray tones at a certain distance
- Normed symmetric matrix
- Size N x N, N being the maximal gray tone value
- Capture 2nd order statistics



4 directions of motion ⇒ 4 co-occurrence matrices

Example: co-occurrence Matrix

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

4x4 image
with four gray tones

d=1

Gray Tone

Gray Tone	0	1	2	3
0	#(0,0)	#(0,1)	#(0,2)	#(0,3)
1	#(1,0)	#(1,1)	#(1,2)	#(1,3)
2	#(2,0)	#(2,1)	#(2,2)	#(2,3)
3	#(3,0)	#(3,1)	#(3,2)	#(3,3)

0 degree

$$\begin{bmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

45 degree

$$\begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

90 degree

$$\begin{bmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

135 degree

$$\begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

5. The original 14 features(proposed by Haralick,Shanmugam ,and Dinstein, 1973)

1. Angular Second Moment
2. Contrast
3. Correlation
4. Sum of Squares: Variance
5. Inverse Difference Moment
6. Sum Average
7. Sum Variance
8. Sum Entropy
9. Entropy
10. Difference Entropy
11. Difference Variance
12. Information Measures of Correlation I
13. Information Measures of Correlation II
14. Maximum Correlation Coefficient

Notation:

p(i,j) (i,j)th entry in a normalized co-occurrence matrix

$p_x(i)$ i-th entry in the marginal-probability matrix obtained by summing the rows of p(i,j), **$p_y(j)$** is defined respectively by summing the columns

N number of distinct gray levels in the quantized image

Function1

Angular Second Moment:

$$f_1 = \sum_i \sum_j \{p(i, j)\}^2$$

-measure of the smoothness of the image

-all pixels have the same gray level I=k

⇨ P(k,k)=1 for i=j and P(i,j)=0, else.

⇨ ASM=1.

-all pixels have different gray level

⇨ P(i,j)=1/R

⇨ ASM=1/R.

Function 2

Contrast:

$$f_2 = \sum_{n=0}^{N-1} n^2 \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} p(i,j) \right\}, |i-j| = n$$

-measure of the image contrast (locally gray-level variations)

-the argument of the first sum is the percentage of pixels whose intensity differs by n

-n^2 weighs the big differences more

-takes large values for large contrast

Function 3-4

Correlation:

$$f_3 = \frac{\sum_i \sum_j (i * j)\, p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

Sum of Squares: Variance

$$f_4 = \sum_i \sum_j (i - \mu)^2\, p(i,j)$$

Function 5 + 9

Inverse Difference Moment:

$$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$$

-takes high values for low-contrast images due to the inverse (i-j)^2

Entropy:

$$f_9 = -\sum_i \sum_j p(i,j) \log(p(i,j))$$

-measure of randomness

-takes low values for smooth images

Function 6-8

Sum Average:

$$f_6 = \sum_{k=2}^{2N} k p_{x+y}(k)$$

Sum Variance:

$$f_7 = \sum_{k=2}^{2N} (k - f_6)^2 p_{x+y}(k)$$

$$p_{x+y}(k) = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i+j=k}}^{N} p(i, j), \quad k = 2,3,\ldots,2N.$$

Sum Entropy:

$$f_8 = -\sum_{k=2}^{2N} p_{x+y}(k) \log\left\{p_{x+y}(k)\right\}$$

Functions 10-11

Difference Entropy:

$$f_{11} = -\sum_{k=0}^{N-1} p_{x-y}(k) \log\left\{p_{x-y}(k)\right\}$$

Difference Variance:

$$f_{10} = \mathrm{var}\,iance(p_{x-y})$$

$$p_{x-y}(k) = \sum_{i=1}^{N} \sum_{\substack{j=1 \\ |i-j|=k}}^{N} p(i, j), \qquad k = 0,1,\ldots,N-1.$$

Functions 12-13

Information Measures of Correlation I+II:

$$f_{12} = \frac{f_9 - HXY1}{\max\{HX, HY\}}$$

$$f_{13} = (1 - \exp[-2.0(HXY2 - f_9)]))^{1/2}$$

$$HXY1 = -\sum_i \sum_j p(i,j) \log\{p_x(i)\, p_y(j)\}$$

$$HXY2 = -\sum_i \sum_j p_x(i)\, p_y(j) \log\{p_x(i)\, p_y(j)\}$$

HX and HY are entropies of p_x and p_y

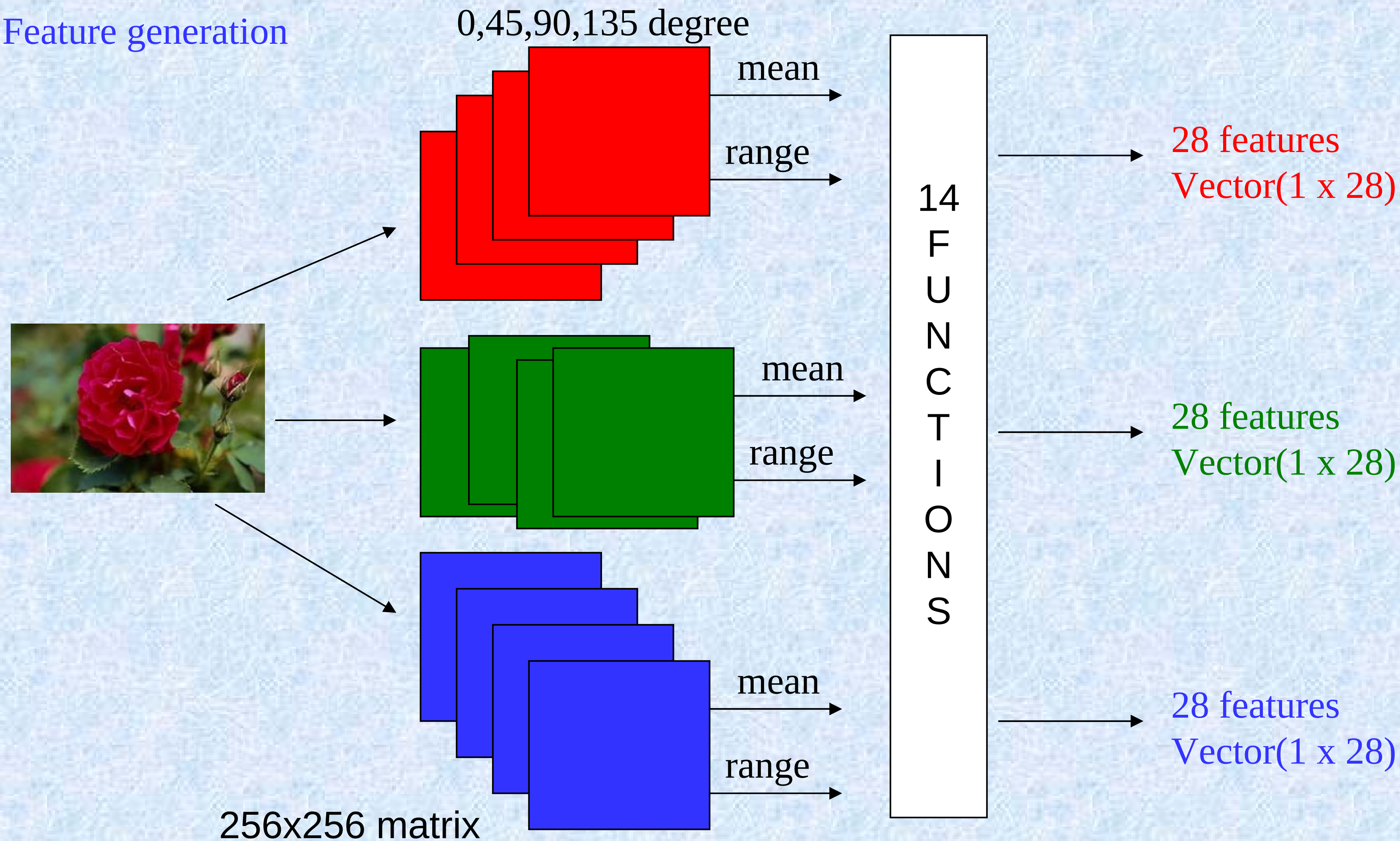
Functions 14

Maximal Correlation Coefficient:

f_{14} = (Second largest eigenvalue of Q)$^{1/2}$

$$Q(i,j) = \sum_k \frac{p(i,k)\,p(j,k)}{p_x(i)\,p_y(k)}$$

Feature generation



6. Testing results

- d=1,3
- Color spaces: RGB,HSV
- For each category was chosen one representative picture ‘1.jpg’,’101.jpg’,..’901.jpg’
- output: 12 pictures

#pictures in the right category out of the first 3 ones

#pictures in the right category

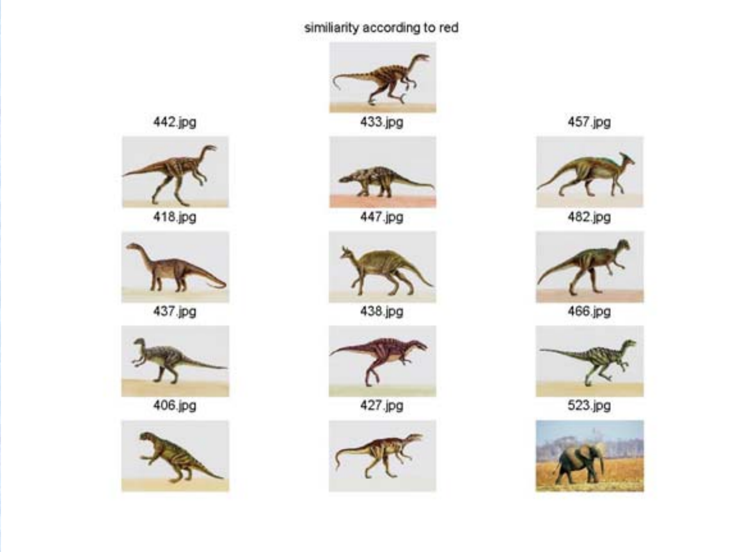
- Euclidian distance as a measure of similarity between the feature vectors d=1 #out of first 3/# out of first 12

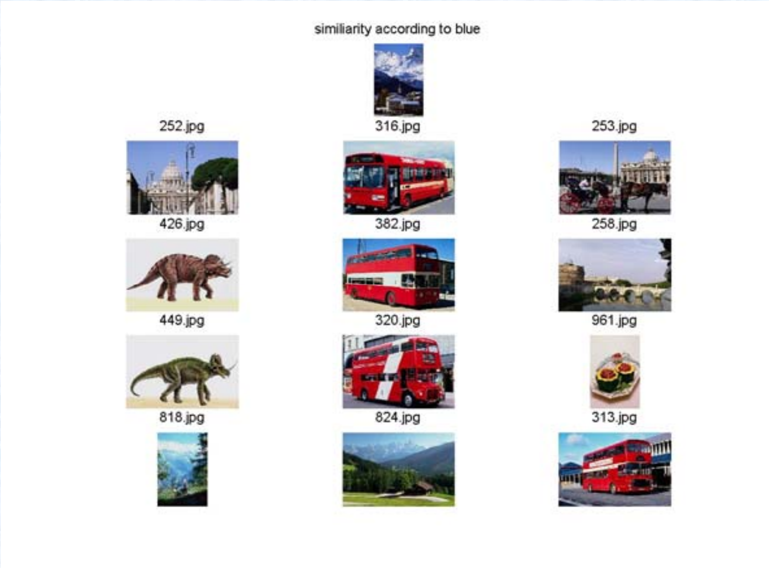
Category	red	green	blue
African People	1/2	0/3	1/7
Beach	0/1	0/2	1/4
Antique building	1/4	0/1	2/5
Bus	0/1	0/1	0/2
Dinosaur	3/11	3/12	3/10
Elephant	0/0	1/5	2/4
Flower	1/3	3/10	3/11
Horse	1/3	1/4	3/8
Mountain	0/0	0/3	0/2
Food	0/2	0/2	0/0

d=3

Category	red	green	blue
African People	0/5	0/1	2/5
Beach	0/1	0/2	0/3
Antique building	1/2	0/1	1/2
Bus	0/0	0/1	0/2
Dinosaur	3/11	3/12	3/10
Elephant	0/1	0/2	1/5
Flower	3/4	3/9	3/11
Horse	1/3	0/3	2/8
Mountain	0/1	0/0	0/0
Food	1/1	1/1	0/0

- category best classified: dinosaur, flower
- category worst classified: mountain, food, bus
- d = 1 better than d=3
- color blue best!






similiarity according to blue
680.jpg
672.jpg
637.jpg
646.jpg
785.jpg
699.jpg
624.jpg
696.jpg
621.jpg
698.jpg
659.jpg
682.jpg

d=1 #out of first 3/# out of first 12 d=3

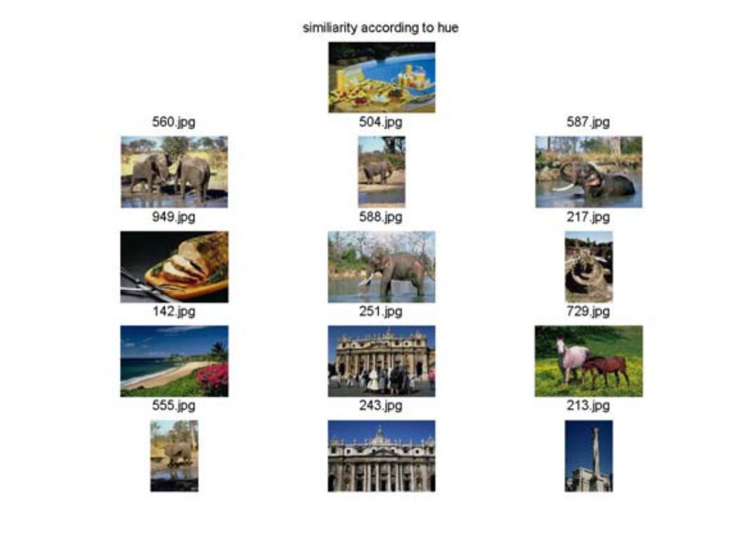
d=1

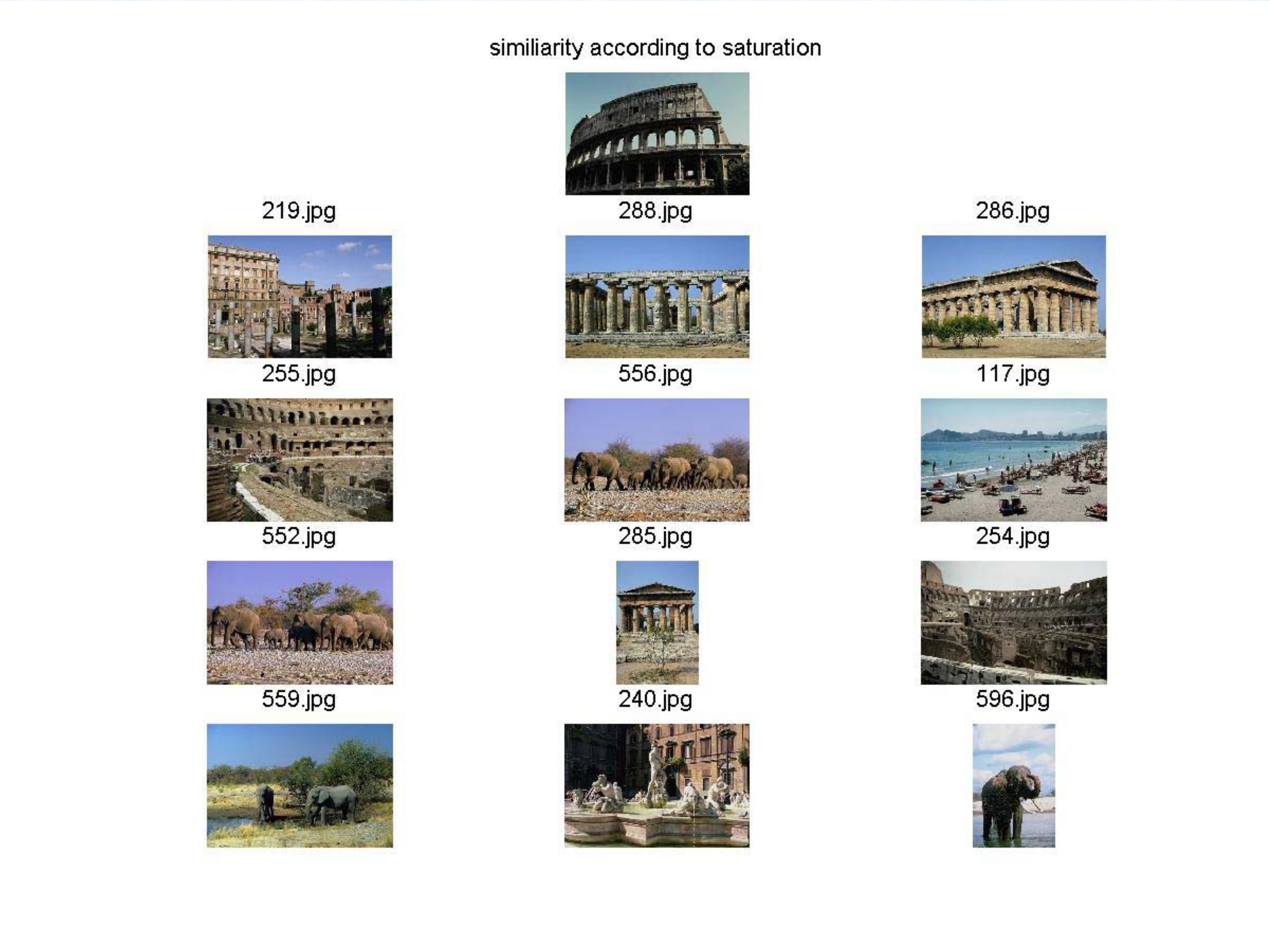
Category	hue	saturation	value
African People	0/1	2/7	0/1
Beach	0/5	1/1	1/2
Antique building	1/2	1/5	2/6
Bus	2/4	2/11	0/2
Dinosaur	3/9	1/6	3/11
Elephant	3/4	0/1	1/4
Flower	1/3	3/6	1/2
Horse	2/3	1/4	2/2
Mountain	2/7	0/0	0/3
Food	1/2	1/2	0/0

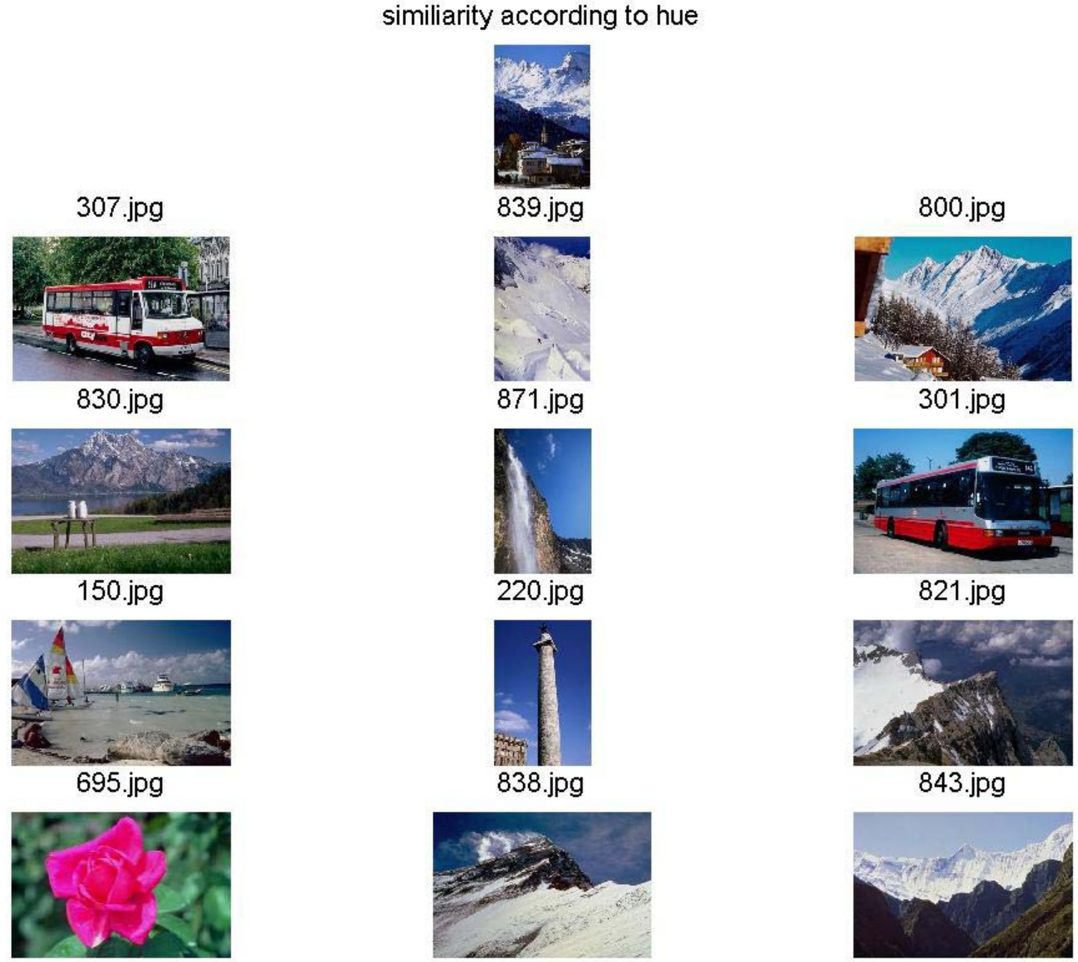
d=3

Category	hue	saturation	value
African People	0/3	2/6	0/2
Beach	1/4	1/2	1/3
Antique building	1/3	2/7	1/4
Bus	0/2	3/11	0/2
Dinosaur	3/10	2/8	3/11
Elephant	1/3	0/0	1/4
Flower	1/5	1/6	2/2
Horse	2/4	1/3	1/4
Mountain	2/6	0/1	0/1
Food	0/1	1/6	0/2

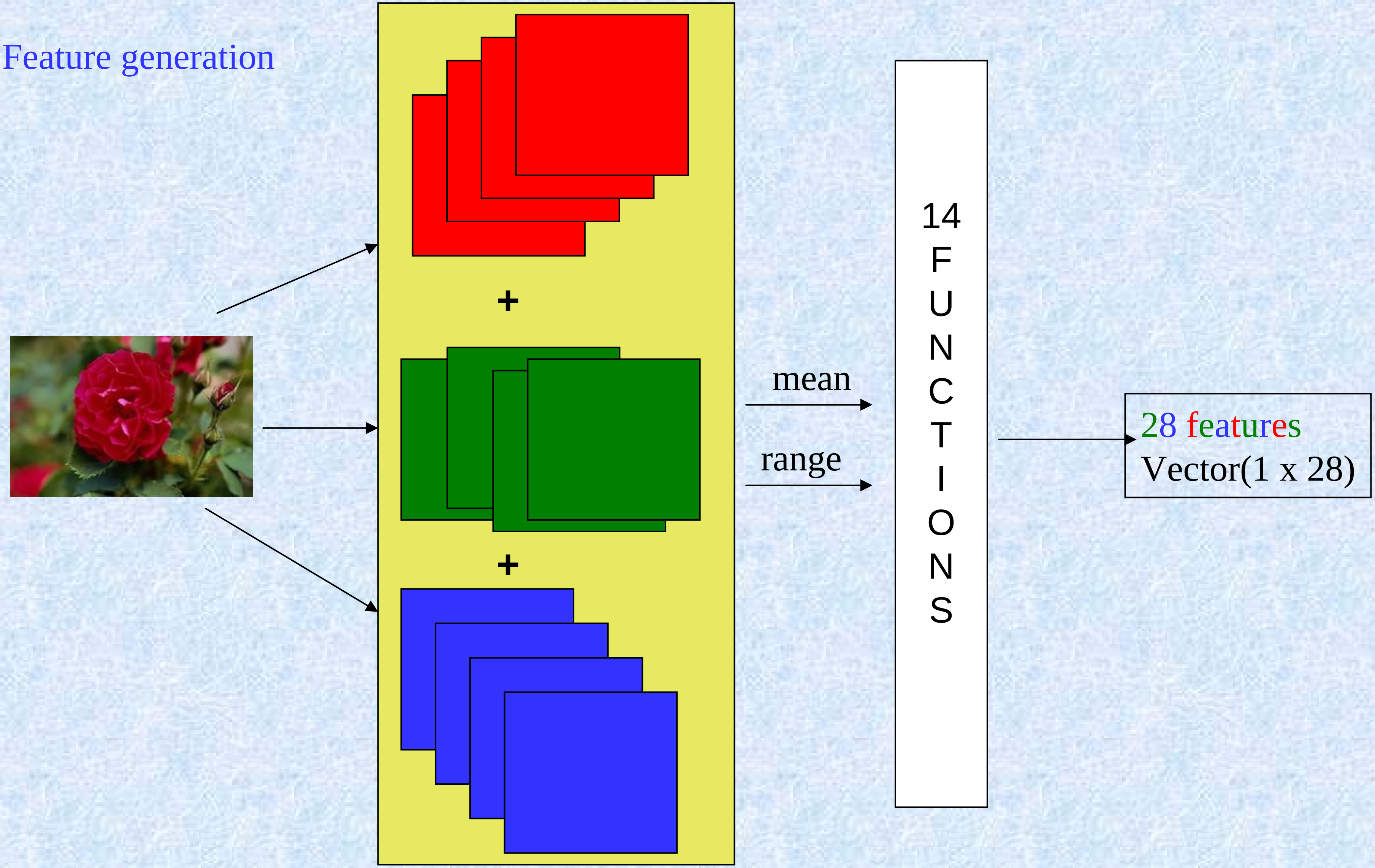
- category best classified: dinosaur, bus
- category worst classified: food
- d = 1 not better than d=3
- hue ,saturation better than value?




similiarity according to saturation
219.jpg
288.jpg
286.jpg
255.jpg
556.jpg
117.jpg
552.jpg
285.jpg
254.jpg
559.jpg
240.jpg
596.jpg


similiarity according to hue
307.jpg
839.jpg
800.jpg
830.jpg
871.jpg
301.jpg
150.jpg
220.jpg
821.jpg
695.jpg
838.jpg
843.jpg

Feature generation



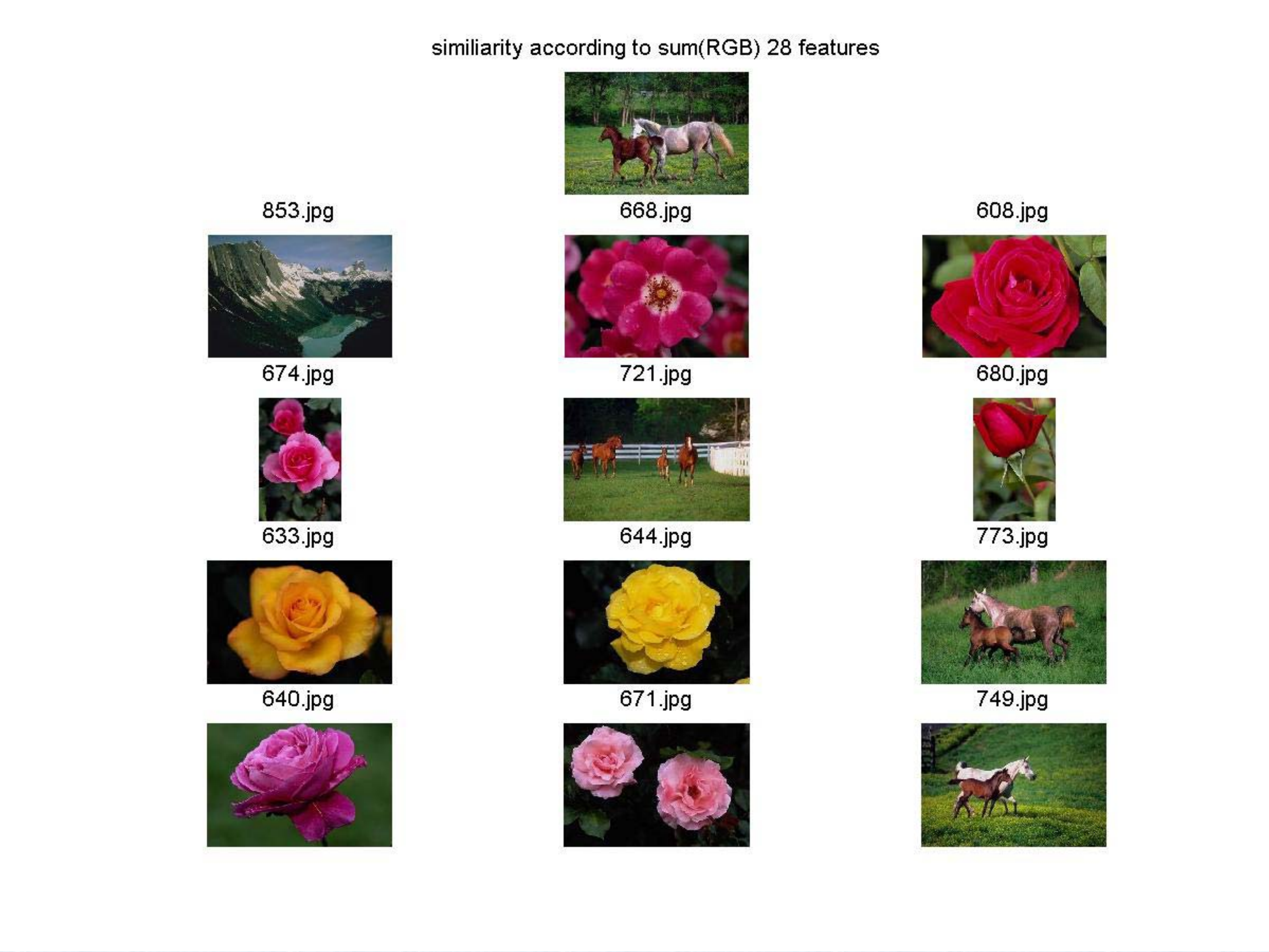
Classification over the sum of RGB/HSV:

Category	sum(rgb) d=1	sum(hsv) d=1
African People	1/2	1/2
Beach	0/0	2/2
Antique building	2/4	3/7
Bus	1/3	2/5
Dinosaur	3/12	1/3
Elephant	2/5	0/3
Flower	1/4	2/7
Horse	0/3	2/5
Mountain	1/1	2/5
Food	0/1	1/4

⇨ category best classified: sum(rgb)->dinosaur
sum(hsv)->bus

⇨ category worst classified: sum(rgb)->beach
sum(hsv)->elephant

⇨ Better to consider the distinct color channels


similiarity according to sum(RGB) 28 features
853.jpg
668.jpg
608.jpg
674.jpg
721.jpg
680.jpg
633.jpg
644.jpg
773.jpg
640.jpg
671.jpg
749.jpg


similiarity according to sum(HSV) 28 features
0.jpg
721.jpg
704.jpg
772.jpg
707.jpg
588.jpg
522.jpg
11.jpg
653.jpg
167.jpg
732.jpg
86.jpg

Different categories?

7. Outlook: -minimal set of features

-What functions 1-14 are really necessary for classification?

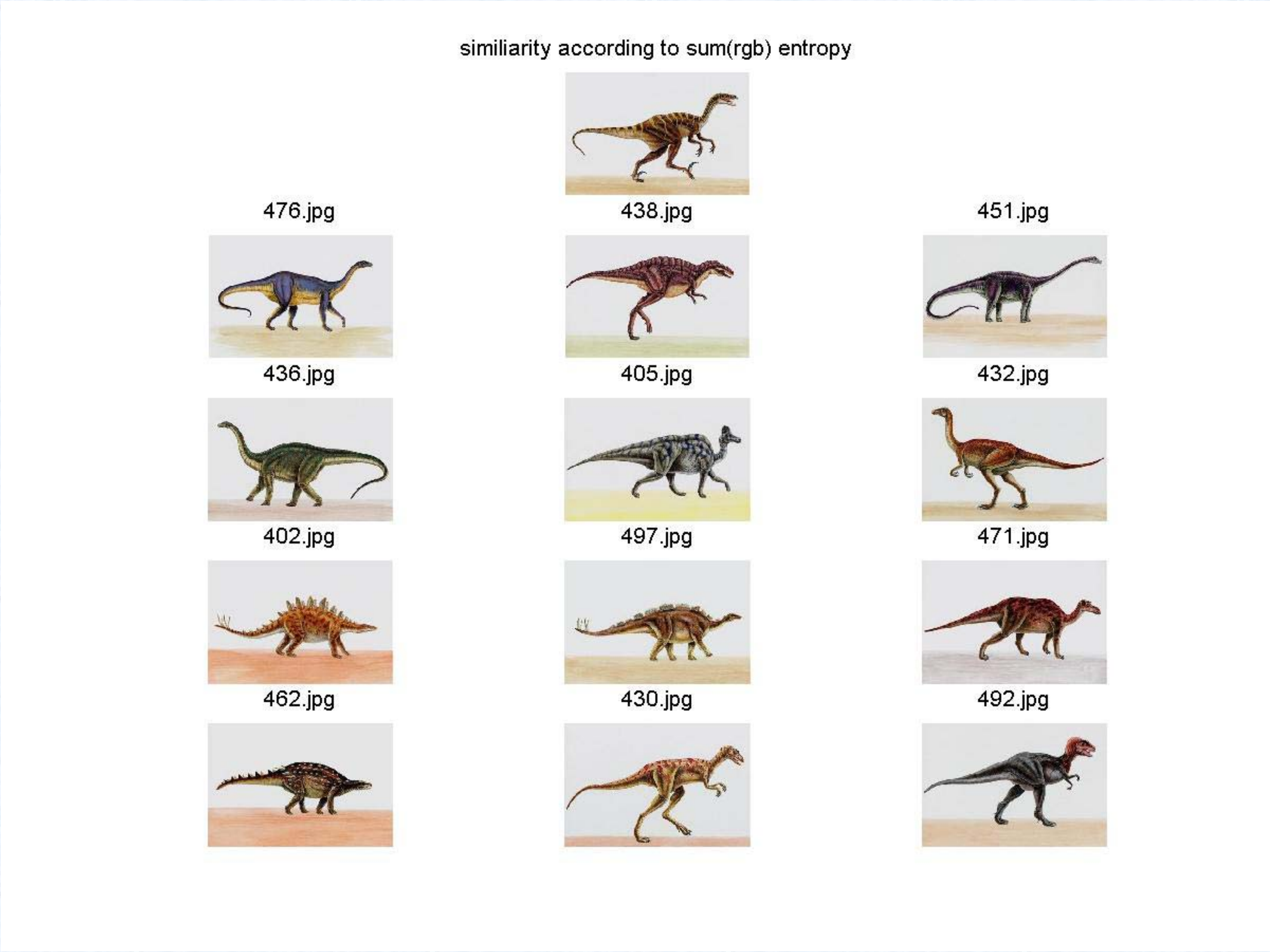
-angular second moment

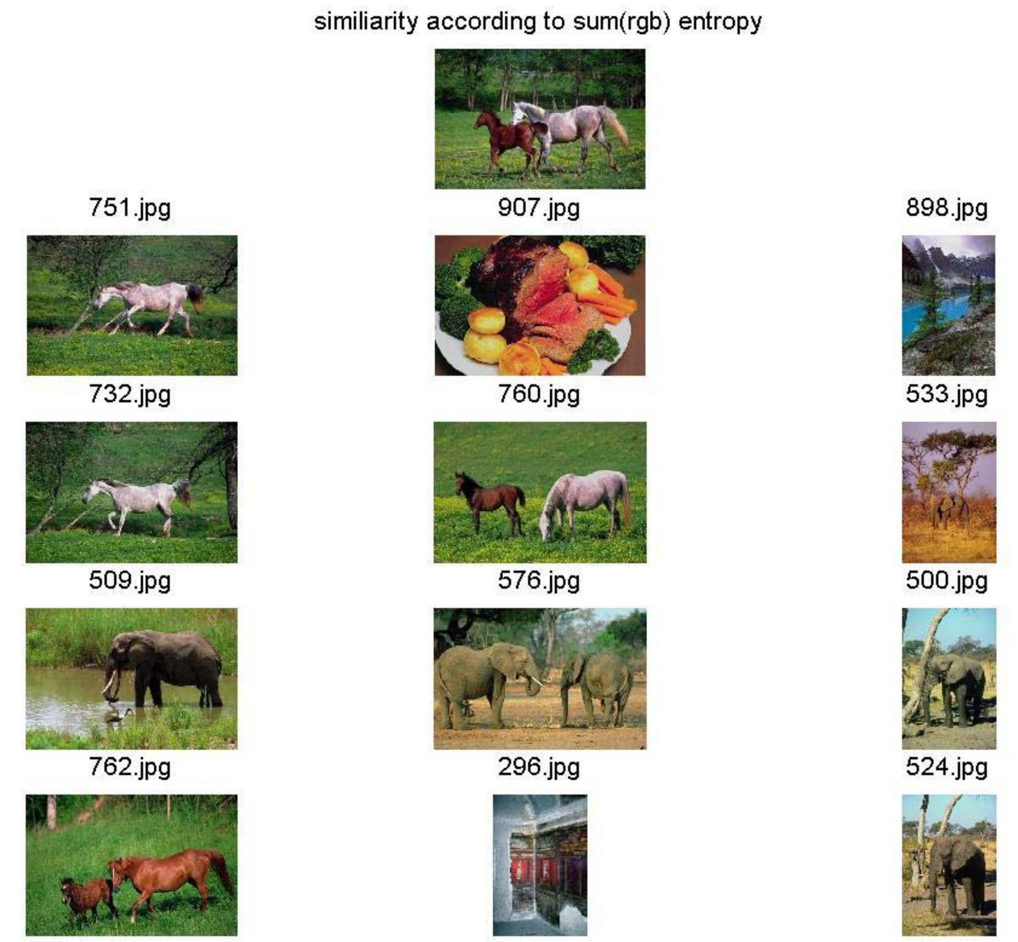
- contrast

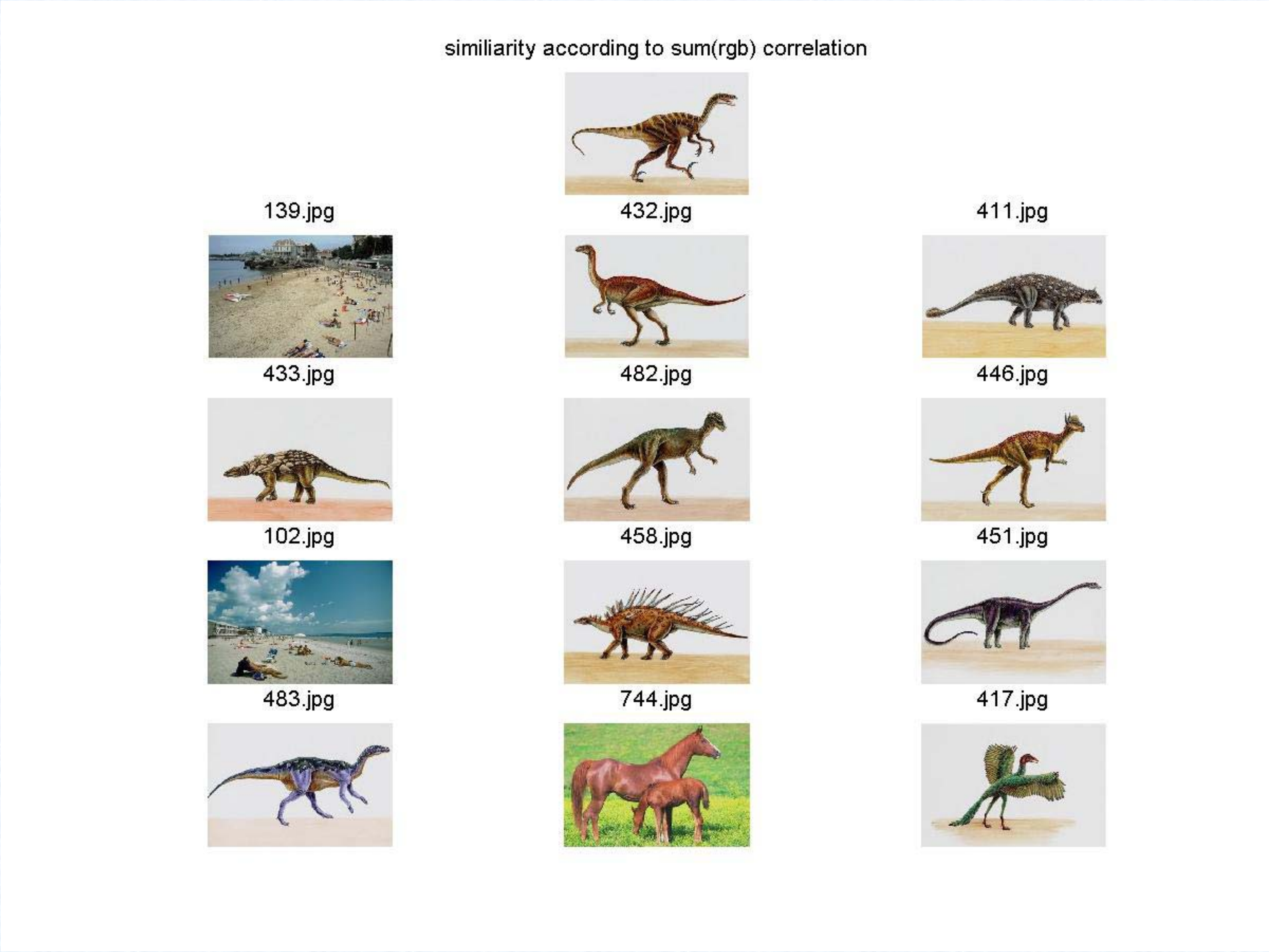
- correlation

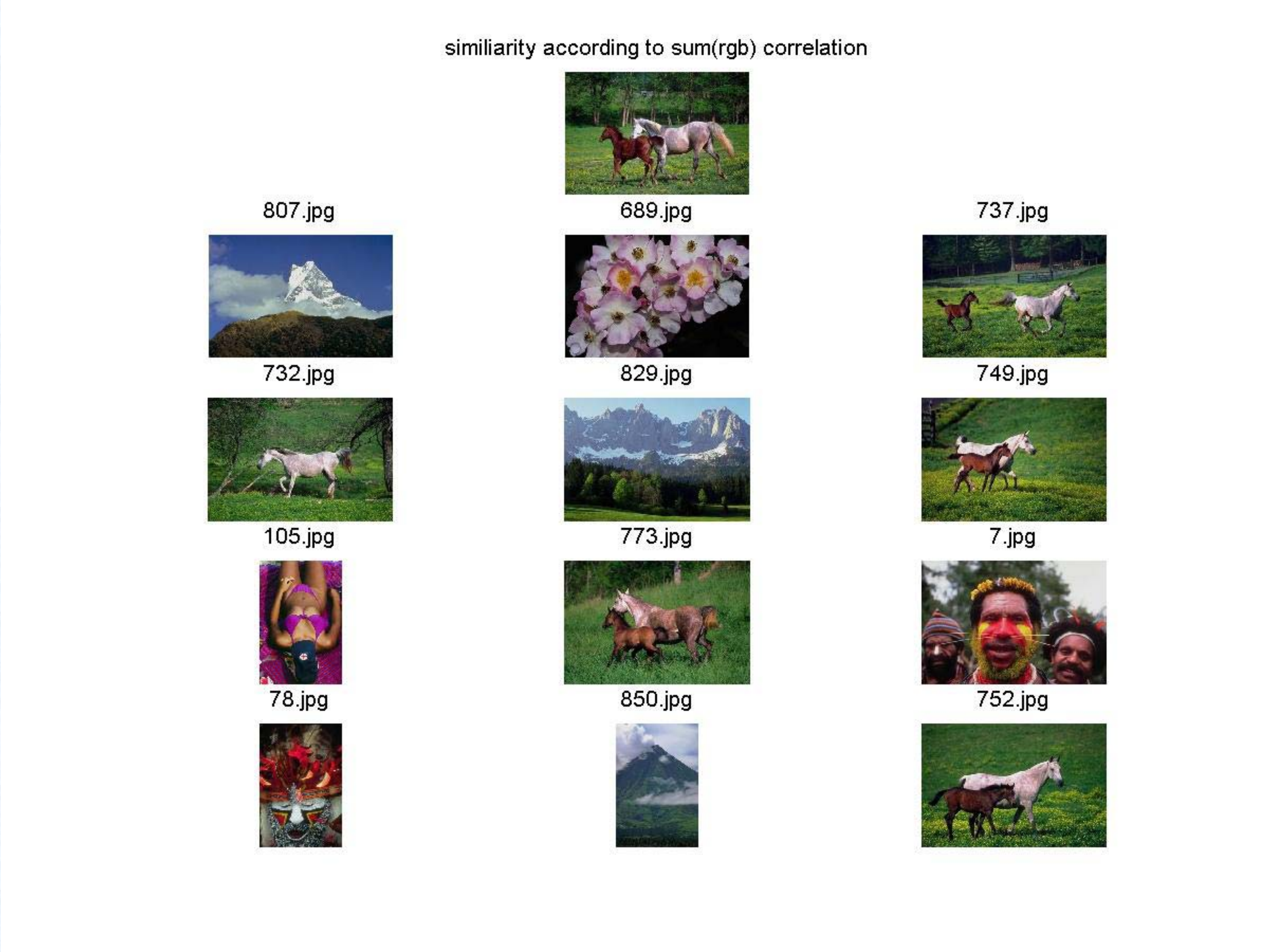
- entropy

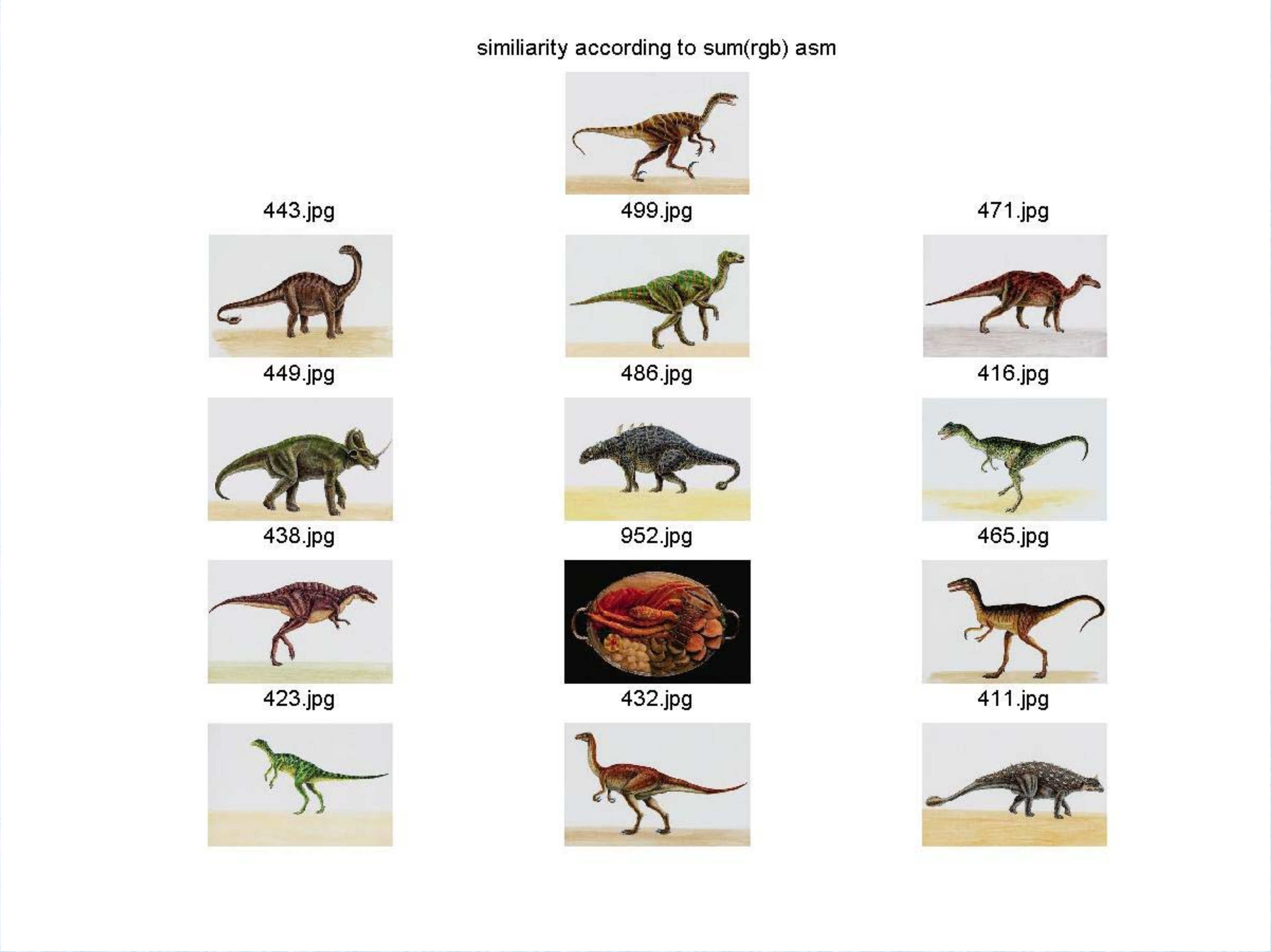
-Are there other interesting functions apart of the 14 proposed ?

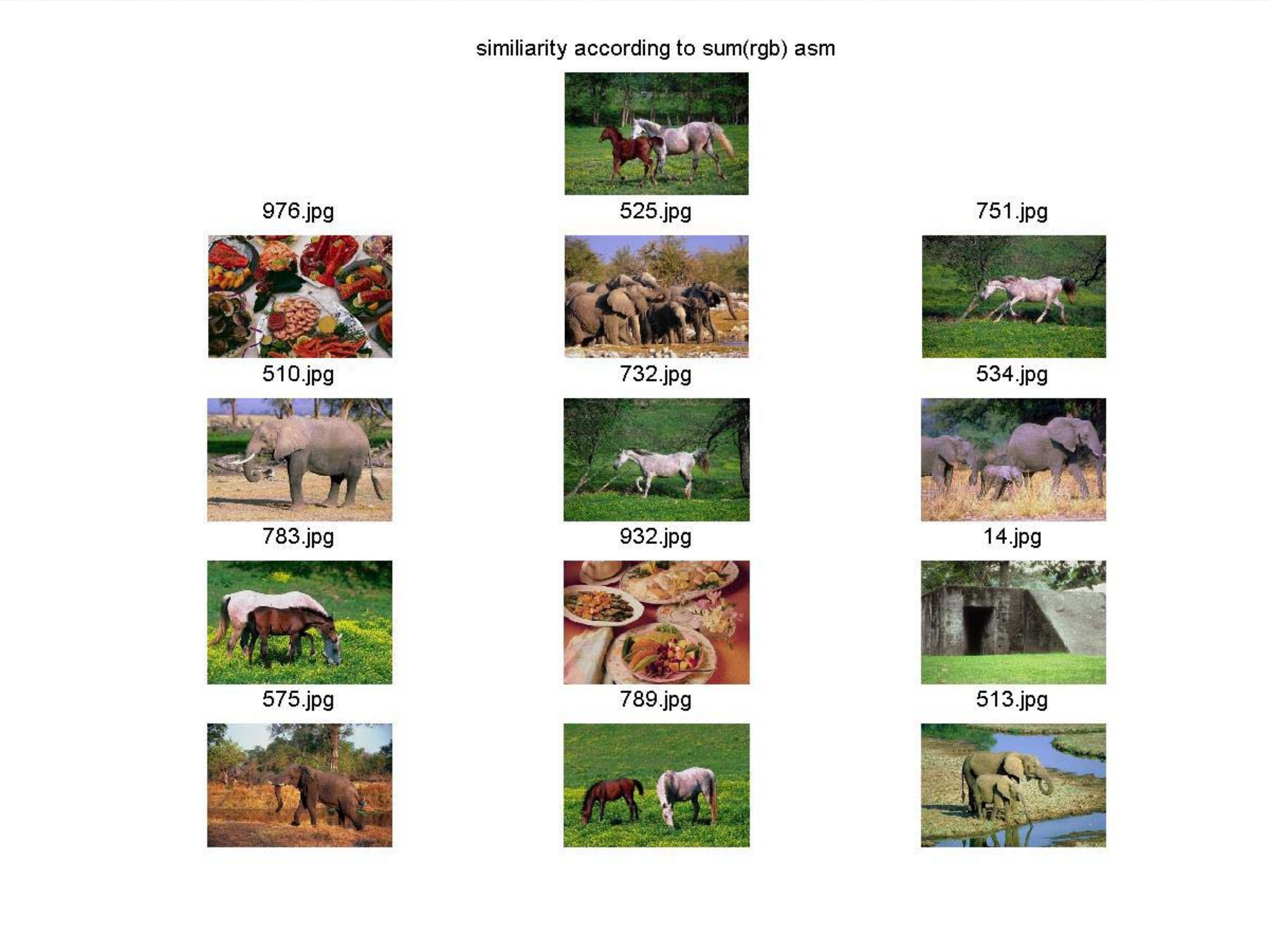

similiarity according to sum(rgb) entropy
476.jpg
438.jpg
451.jpg
436.jpg
405.jpg
432.jpg
402.jpg
497.jpg
471.jpg
462.jpg
430.jpg
492.jpg


similiarity according to sum(rgb) entropy
751.jpg
907.jpg
898.jpg
732.jpg
760.jpg
533.jpg
509.jpg
576.jpg
500.jpg
762.jpg
296.jpg
524.jpg


similiarity according to sum(rgb) correlation
139.jpg
432.jpg
411.jpg
433.jpg
482.jpg
446.jpg
102.jpg
458.jpg
451.jpg
483.jpg
744.jpg
417.jpg


similiarity according to sum(rgb) correlation
807.jpg
689.jpg
737.jpg
732.jpg
829.jpg
749.jpg
105.jpg
773.jpg
7.jpg
78.jpg
850.jpg
752.jpg


similiarity according to sum(rgb) asm
443.jpg
499.jpg
471.jpg
449.jpg
486.jpg
416.jpg
438.jpg
952.jpg
465.jpg
423.jpg
432.jpg
411.jpg

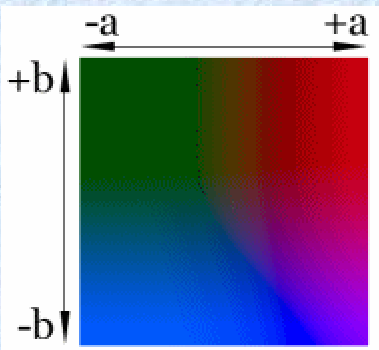

similiarity according to sum(rgb) asm
976.jpg
525.jpg
751.jpg
510.jpg
732.jpg
534.jpg
783.jpg
932.jpg
14.jpg
575.jpg
789.jpg
513.jpg

7. Outlook: -other color spaces

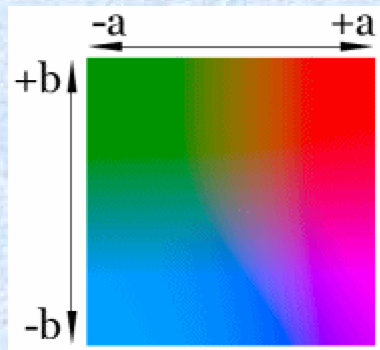
-**YUV** was to difficult to implement because of the negative values

-What about **L*a*b***?

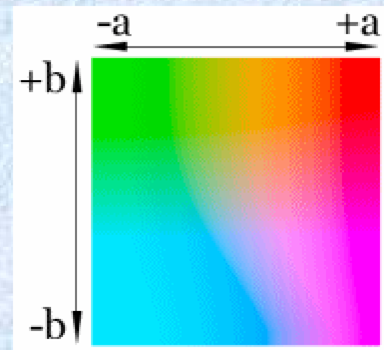
-most complete color model used conventionally to describe all the colors visible to the human eye



Luminance 25%



Luminance 50%



Luminance 75%

-luminance of the color (**L***, the smallest L yields black)

-position between red and green (**a***, the smallest a yields green)

-position between yellow and blue (**b***, the smallest b yields blue)

Advantages of Lab:

-Difference between color can be measured easily

$$d = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}$$

Disadvantages of Lab:

-no linear transformation rgb2lab

L* = 116 (Y/Yn)1/3 - 16 for Y/Yn > 0.008856

L* = 903.3 Y/Yn otherwise

a* = 500 (f(X/Xn) - f(Y/Yn))
b* = 200 (f(Y/Yn) - f(Z/Zn))

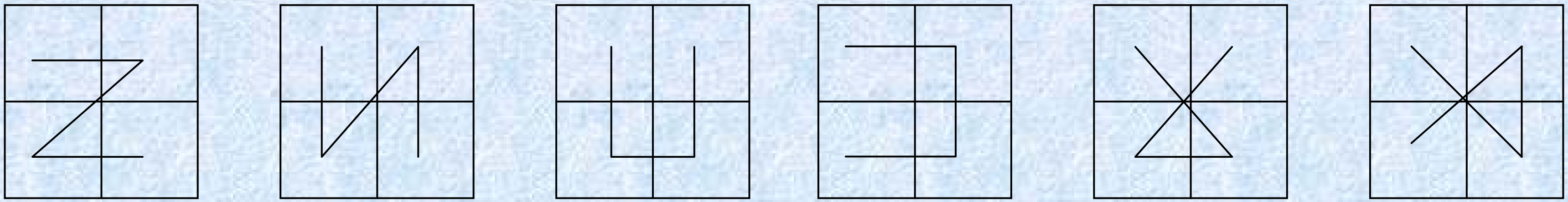
where f(t) = t1/3 for t > 0.008856
f(t) = 7.787 t + 16/116 otherwise

-also negative values

7. Outlook: -motif co-occurenc matrices

(Jhanwar, Chaudhuri, Seetharaman, Zavidovique, 2002)

Primitive scans (called motifs)

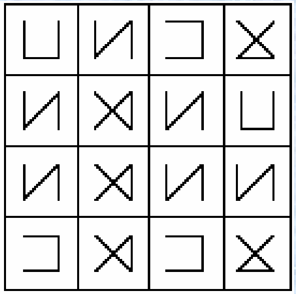


z n u c gamma alpha

18	23
23	46

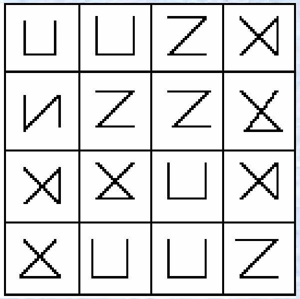
Problem 1: z or n ?

202	53	149	54	255	255	255	124
78	55	84	52	57	190	186	250
129	68	35	128	160	38	36	255
183	29	140	68	54	31	144	182
176	52	47	43	47	53	145	156
145	38	61	45	40	62	140	176
150	186	95	188	220	211	87	167
99	196	189	174	155	159	151	106



original image

53	149	54	255	255	255	124	202
55	84	52	57	190	186	250	78
68	35	128	160	38	36	255	129
29	140	68	54	31	144	182	183
52	47	43	47	53	145	156	176
38	61	45	40	62	140	176	145
186	95	188	220	211	87	167	150
196	189	174	155	159	151	106	99



horizontally shifted image

Problem 2: shifted image → capture 3rd order statistics

Literature:

[1] R.M. Haralick, K. Shanmugam, I. Dinstein, “Textural Features for Image Classification”in SMC(3), No.6, Nov 1973, pp. 610-621. Co-occurrence Matrix. Classic co-occurrence Matrix computation and use.

[2] J.Ruiz-del-Solar, M.Jochmann, “On determining human description of textures” in SCIA 2001, pp. 288-294, June 11-14, Bergen, Norway.

[3] V. Metzler, C. Palm, T. Lehmann, T.Aach ”Texture Classification of Graylevel Images by Multiscale Cross-Cooccurrence Matrices” in ICPR 2000, Barcelona, pp. 549-552.

[4] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique, ”Content Based Image Retrieval Using Motif Cooccurrence Matrix” in *IVC(22)*, No. 14, 1 December 2004, pp. 1211-1220.

Thank you for your attention!