

ΟΙΚΟΝΟΜΙΚΟ  
ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΑΘΗΝΩΝ



ATHENS UNIVERSITY  
OF ECONOMICS  
AND BUSINESS

## School of Information Sciences and Technology

Department of Informatics

Information Processing Laboratory

Information Retrieval Research Group

# Fusion of BoW Models for Image Classification

*Valavanis Leonidas, Spyridon Stathopoulos, Theodore Kalamboukis*

# Our framework

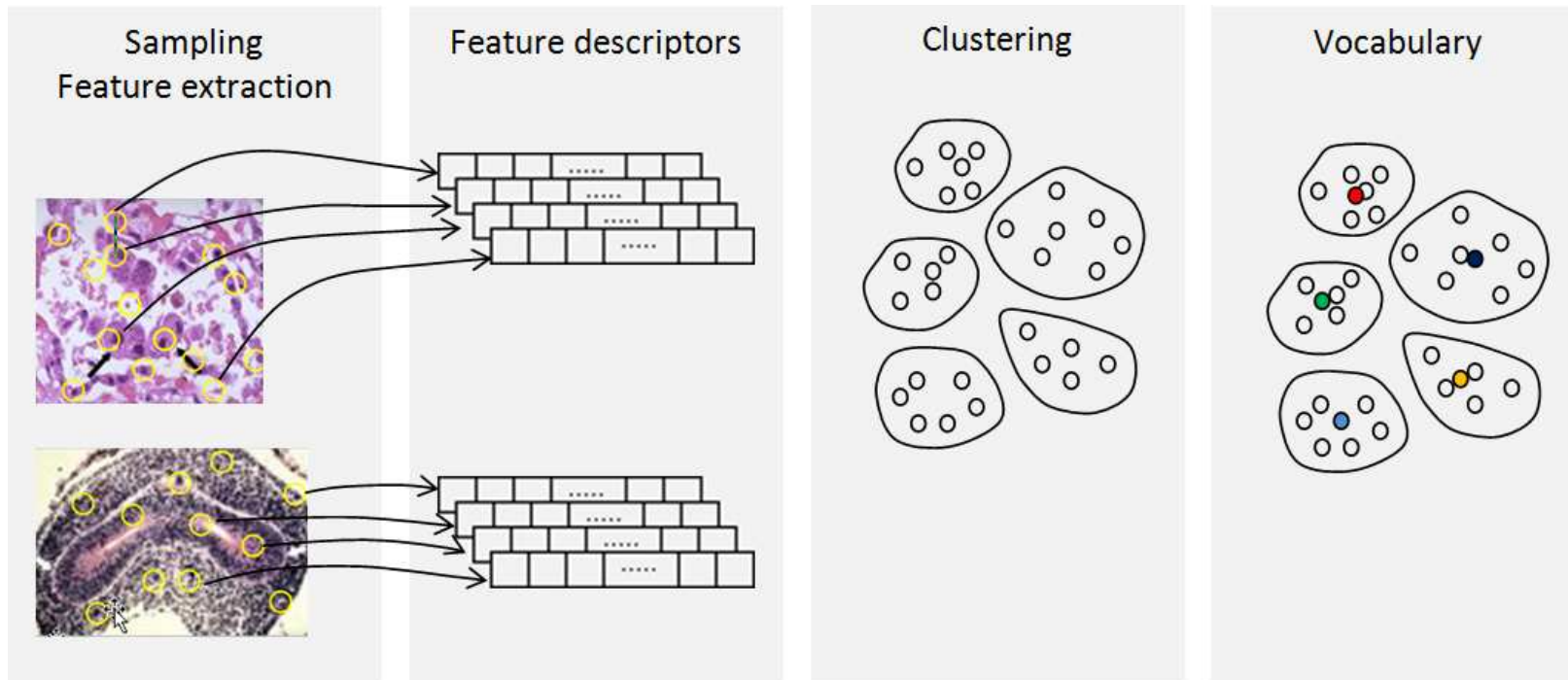
---

Fusion (BoVW, BoC) in an early stage.

**Our Contribution:** a novel implementation and generalization of BoC model that takes into account spatial information of colors based on Quad Tree Decomposition of images (QBoC).

# Bag of Visual Words Model (BoVW)

## Construction of the visual vocabulary



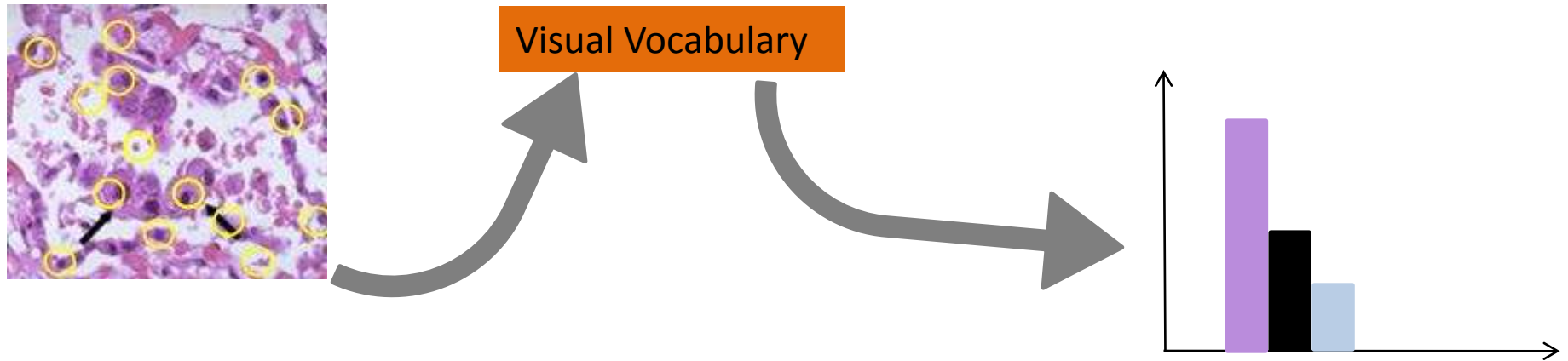
DSIFT

Software : VLFeat ([http://www.vlfeat.org/matlab/vl\\_phow.html](http://www.vlfeat.org/matlab/vl_phow.html))

# BoVW – image representation

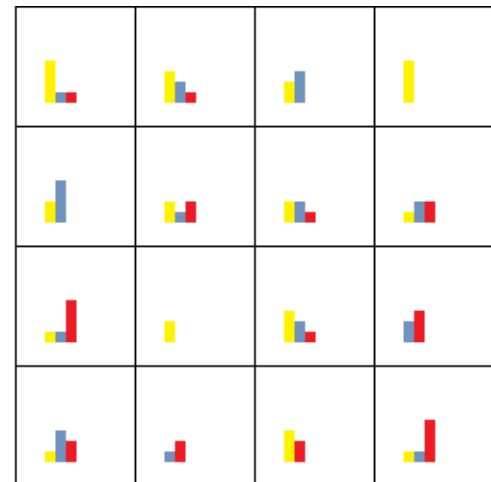
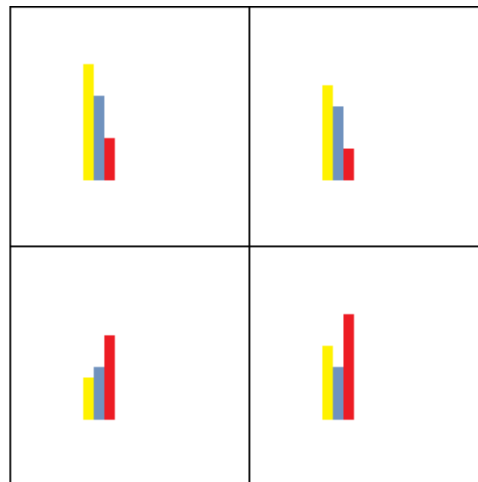
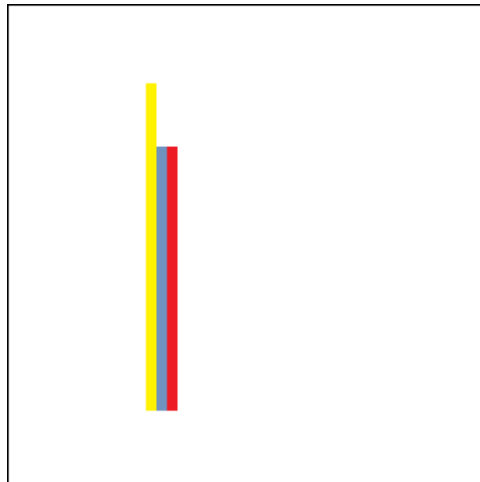
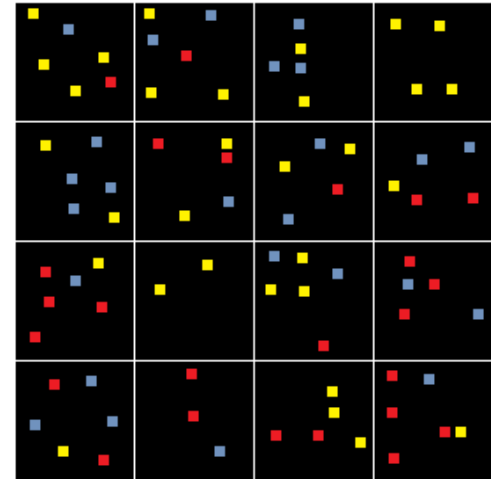
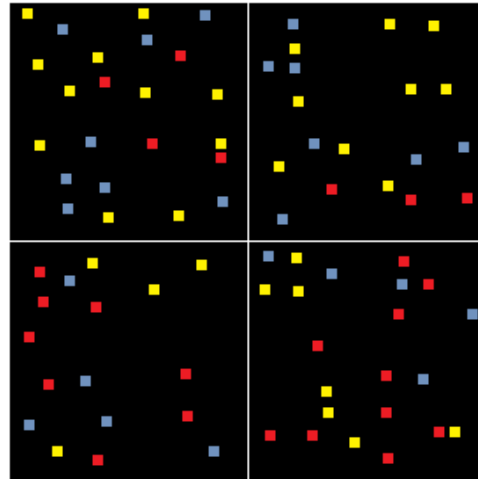
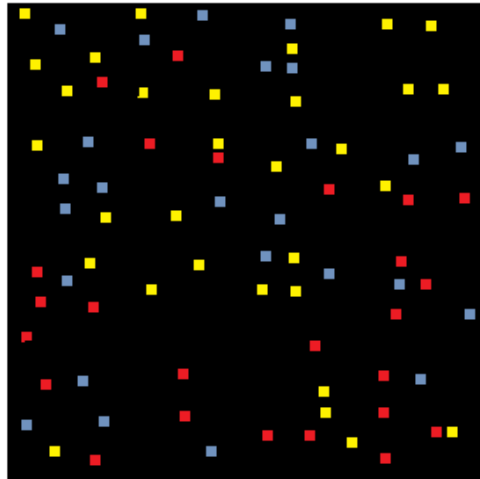
---

for each key-point in the image  
find the nearest visual word



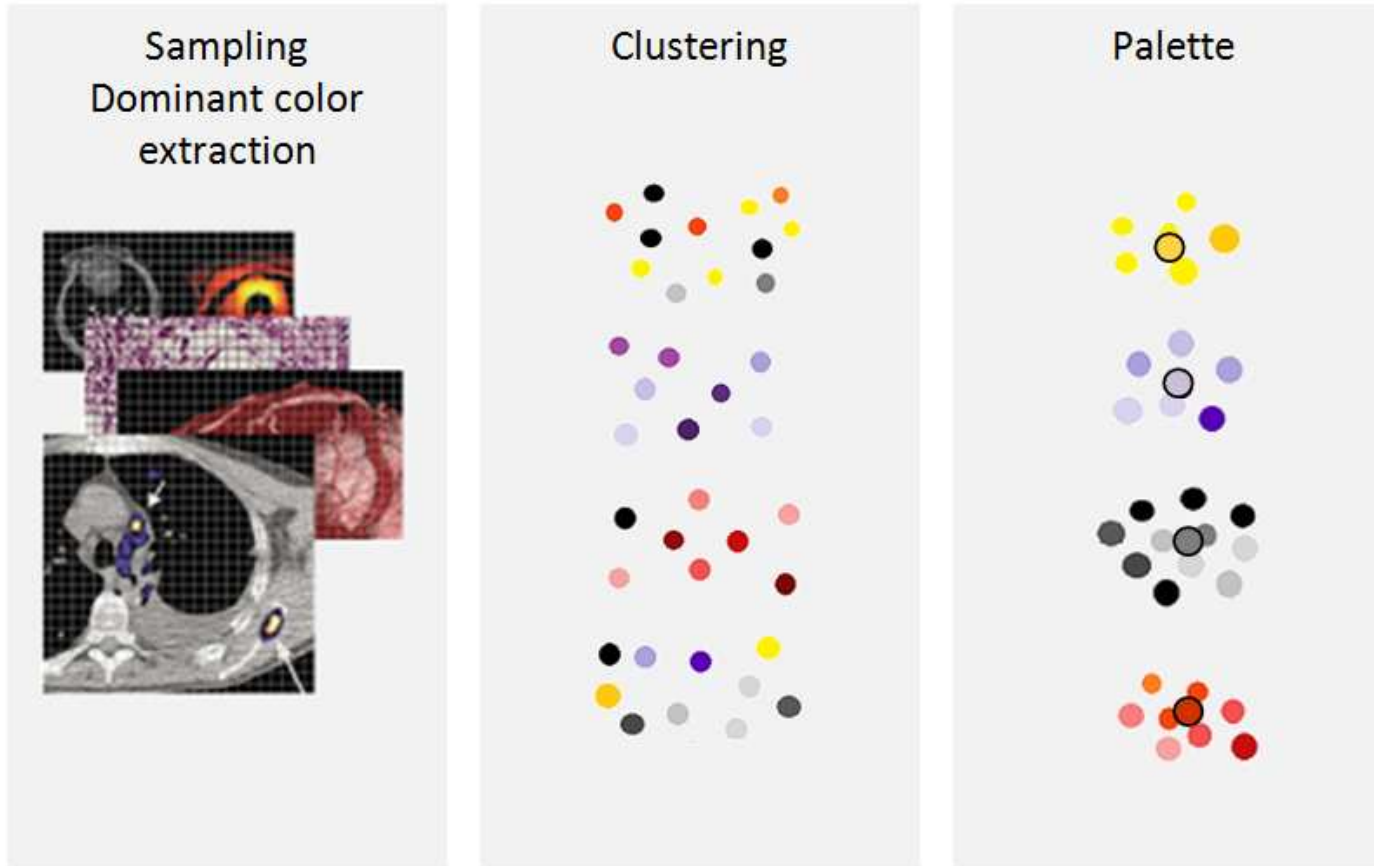
# PHOW

---



# Bag of Colors Model

---

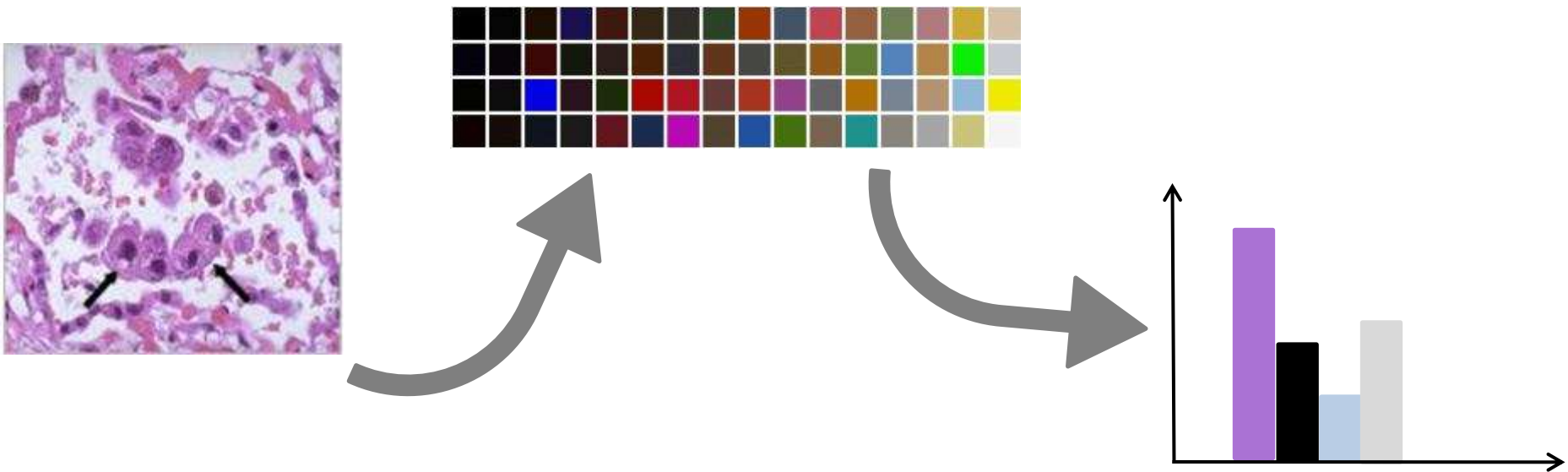


Wengert, C., Douze, M., Jégou, H.: Bag-of-colors for improved image search. In: Proceedings of the 19th International Conference on Multimedia 2011, Scottsdale, AZ, USA, November 28 - December 1, 2011. (2011) 1437–1440

# BoC image representation

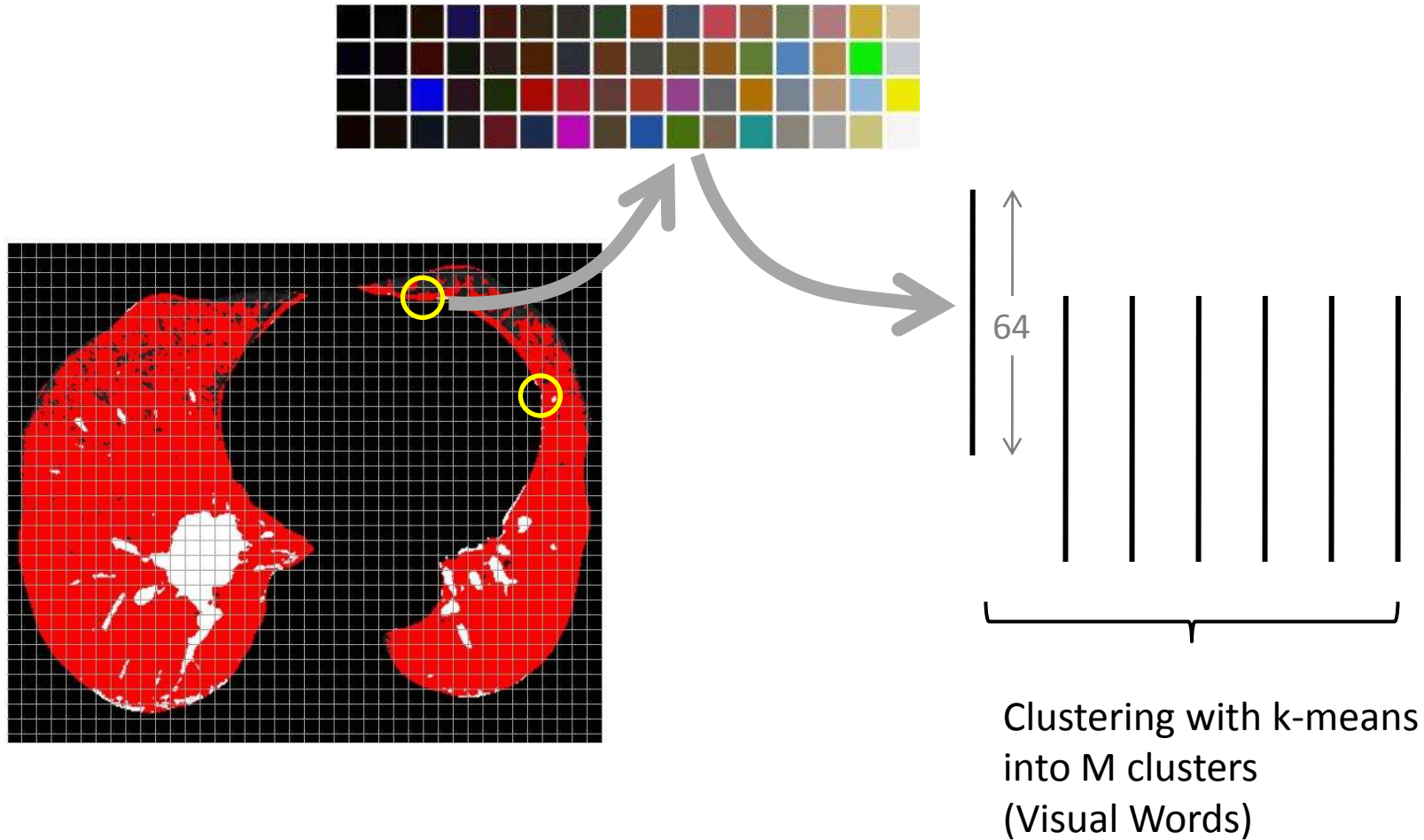
---

for each pixel in the image  
find the nearest color from the palette



# Local Bag of Colors Model

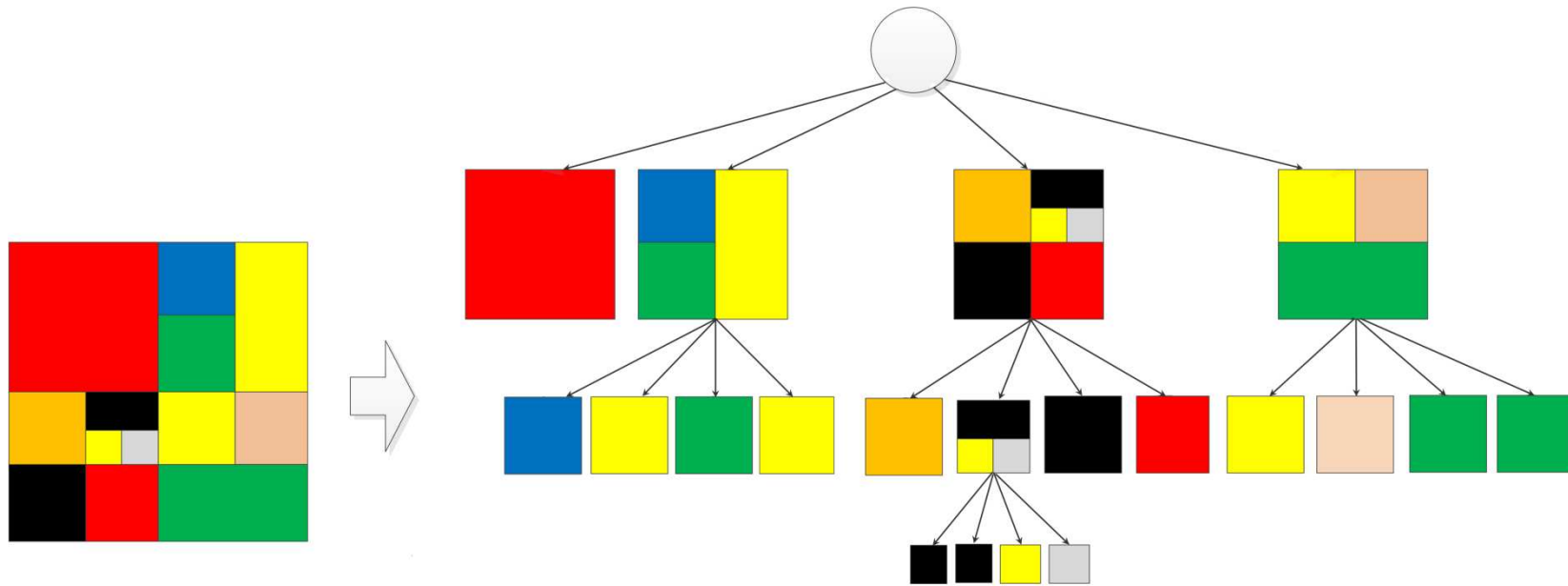
---





# Quad Tree Decomposition

---

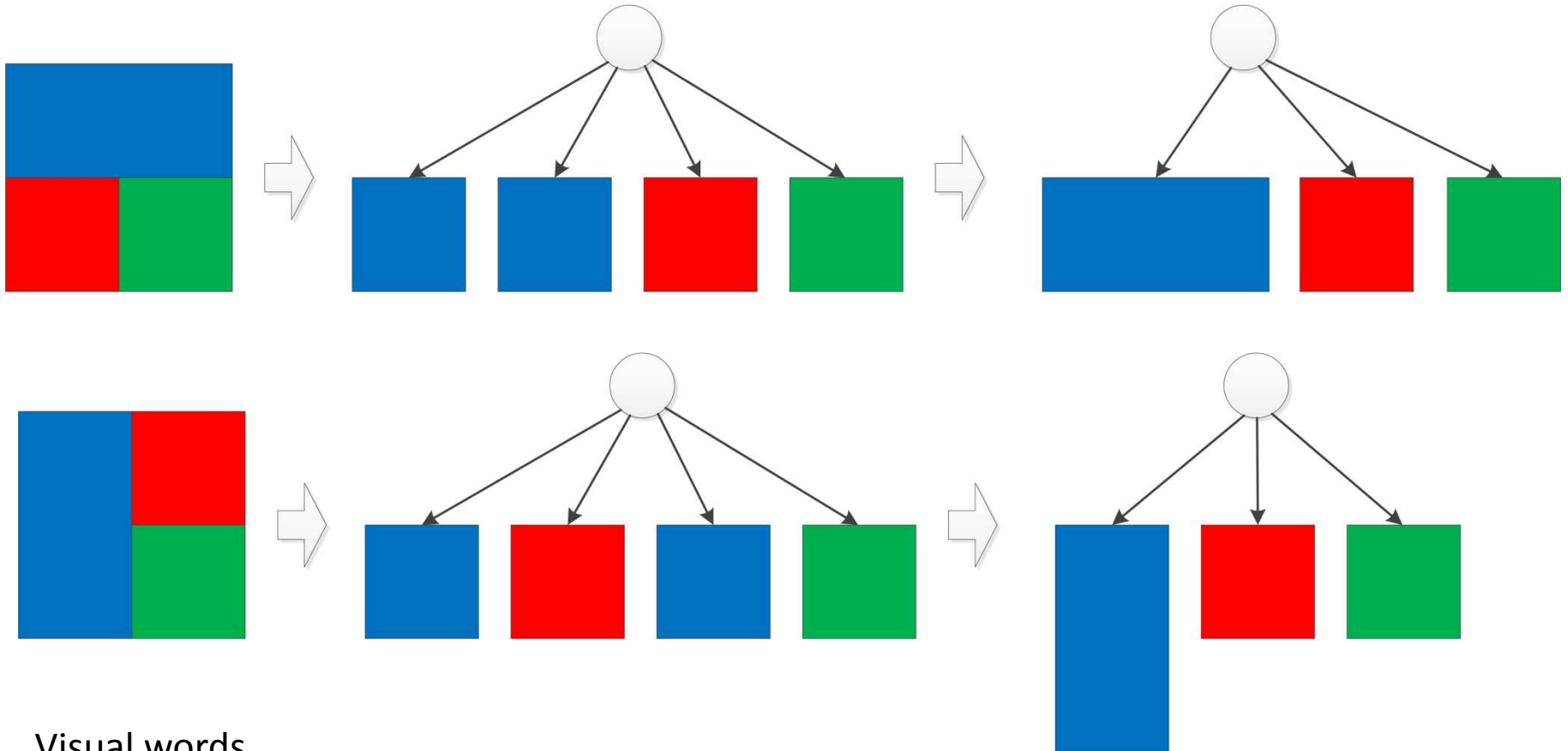


Khan, M., Ohno, Y.: A hybrid image compression technique using quadtree decomposition and parametric line fitting for synthetic images. *Advances in Computer Science and Engineering* 1(3) (2007) 263–283

<https://www.mathworks.com/matlabcentral/fileexchange/9485-quadtree-decomposition--encoding--and-decoding-of-rgb-image>

# Quad Bag of Colors

---



Visual words

`{sqr, hor, ver}_level_color`

(color=the index of the color in the palette)

# Early Fusion

Train set

PHOW
QBoC

Test set

PHOW
QBoC

- I. Original data (distributed by the organizers)
- II. Enriched train set with image from imageCLEF 2013

Late Fusion:  $SCORE(g, C) = \max(w \times P_{PHOW}(g, D) + (1 - w) \times P_{QBoC}(g, C))$

category	CLEF-2016		CLEF-2016
	Train	Test	Train (extended)
D3DR	201	96	201
DMEL	208	88	208
DMFL	906	284	906
DMLI	696	405	696
DMTR	300	96	300
DRAN	17	76	71
DRCO	33	17	55
DRCT	61	71	71
DRMR	139	144	139
DRPE	14	15	30
DRUS	26	129	58
DRXR	51	18	72
DSEC	10	8	39
DSEE	8	3	29
DSEM	5	6	23
DVDM	29	9	80
DVEN	16	8	80
DVOR	55	21	55
GCHE	61	14	61
GFIG	2954	2085	2954
GFLO	20	31	116
GGEL	344	224	344
GEN	179	150	179
GHDR	136	49	136
GMAT	15	3	35
GNCPL	88	20	88
GPLI	1	2	29
GSCR	33	6	33
GSYS	91	75	91
GTAB	79	13	79
<b>TOTAL</b>	<b>6776</b>	<b>4166</b>	<b>7258</b>

# SVM classifier

---

- **LIBLINEAR -- A Library for Large Linear Classification**

<https://www.csie.ntu.edu.tw/~cjlin/liblinear/>

- **Efficient Additive Kernels via the Homogeneous Kernel Map**

<http://www.robots.ox.ac.uk/~vgg/software/homkernmap/#r1>

- **Homogeneous kernel map** <http://vision.princeton.edu/pvt/SiftFu/SiftFu/SIFTransac/vlfeat/doc/api/homkernmap.html>

*R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. [LIBLINEAR: A library for large linear classification](#) *Journal of Machine Learning Research* 9(2008), 1871-1874.*

# Submitted Runs

---

BCSG	SC_BCSG_run8_DeCAF_ResNet-152_PseudoInverse.txt	visual	85.38
BCSG	SC_BCSG_run1_Visual.txt	visual	84.46
IPL	SC_enriched_GBOC_1x1_256_RGB_Phow_Default_1500_EarlyFusion.txt	visual	84.01
IPL	SC_enriched_GBOC_1x1_128_HSV_Phow_RGB_1500_EarlyFusion.txt	visual	83.46
IPL	SC_enriched_GBOC_1x1_128_HSV_Phow_RGB_1500_LateFusion.txt	visual	82.66
IPL	SC_enriched_GBOC_1x1_128_RGB_Phow_Default_1500_LateFusion.txt	visual	82.50
IPL	SC_original_GBOC_1x1_256_RGB_w_0.6_Phow_Default_1500_w_0.4_EarlyFusion.txt	visual	81.73
IPL	SC_original_GBOC_1x1_256_RGB_Phow_Default_1500_EarlyFusion.txt	visual	81.70
IPL	SC_original_GBOC_1x1_128_RGB_Phow_Default_1500_EarlyFusion.txt	visual	81.32
BCSG	SC_BCSG_run7_GoogLeNet-PReLU-Xavier.txt	visual	81.03
IPL	SC_original_GBOC_1x1_256_RGB_Phow_Default_1500_LateFusion.txt	visual	80.17
IPL	SC_original_GBOC_1x1_128_HSV_Phow_RGB_1500_LateFusion.txt	visual	80.14
IPL	SC original GBOC 1x1 128 RGB Phow Default 1500 LateFusion.txt	visual	79.45
BMET	SC-BMET-subfig-test-prob-sum.txt	visual	77.55
BMET	SC-BMET-subfig-test-score-sum-merged.txt	visual	77.53
BMET	SC-BMET-subfig-test-score-sum-crop-scale.txt	visual	77.50
BMET	SC-BMET-subfig-test-majority.txt	visual	77.26
BMET	SC-BMET-subfig-test-prob-max.txt	visual	76.38

# Evaluation of our framework

---

<b>Best Run</b>	<b>Visual</b>	<b>Textual</b> (word2vec)	<b>Mixed</b>
2011	83.59 (XRCE)	70.41 (IPL)	86.91 (XRCE)
<b>BoWs</b>	<b>85.84</b>	<b>77.34</b>	<b>88.57</b>
2012	69.70 (IBM)	41.30 (ITI)	66.20 (medGIFT)
<b>BoWs</b>	<b>71.80</b>	<b>61.40</b>	<b>77.90</b>
2013	80.79 (IBM)	64.17 (IBM)	81.68 (IBM)
<b>BoWs</b>	<b>83.04</b>	<b>68.63</b>	<b>85.71</b>
2016	<b>85.38</b> (BCSG)	72.22 (BCSG)	<b>88.43</b> (BCSG)
<b>BoWs</b>	85.15	<b>73.48</b>	87.78

---

THE END

Thank you for listening