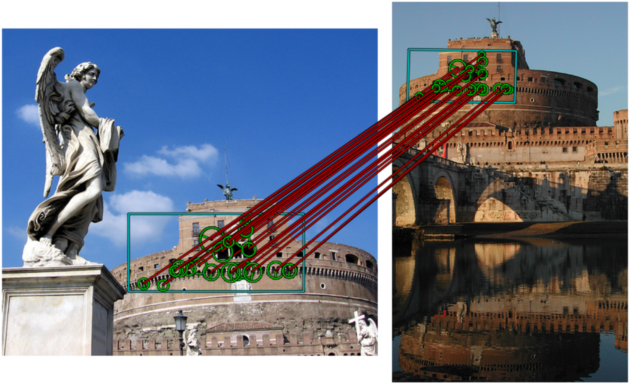
Geometry in feature detection, matching, search, and clustering

Yannis Avrithis

Philadelphia, June 2015

motivation: visual search

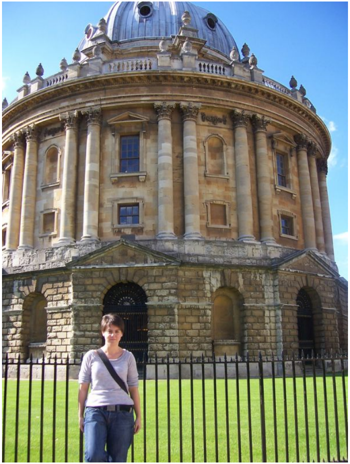
challenges



- viewpoint
- lighting
- occlusion
- large scale

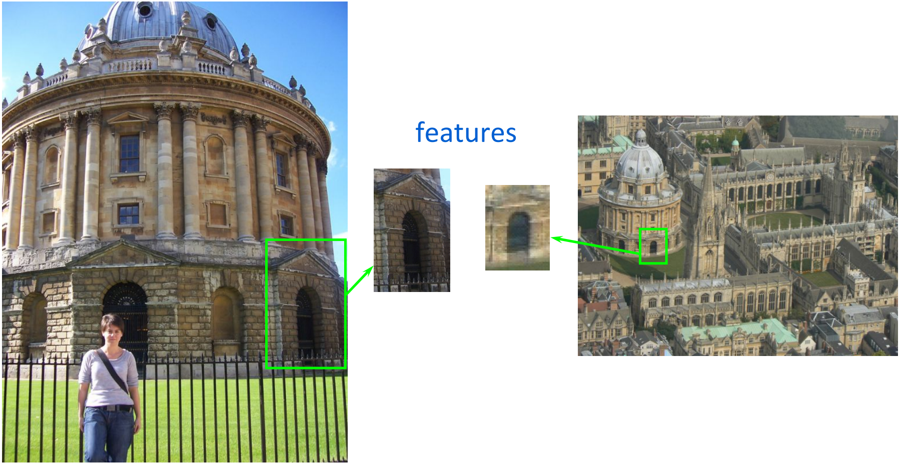
discriminative local features

[Lowe, ICCV 1999]



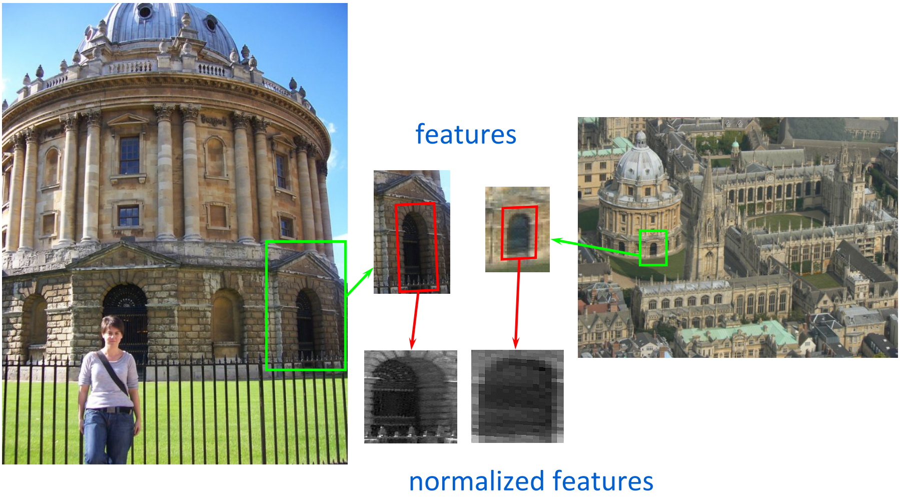
discriminative local features

[Lowe, ICCV 1999]

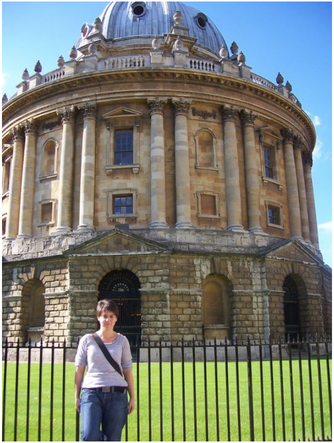


discriminative local features

[Lowe, ICCV 1999]



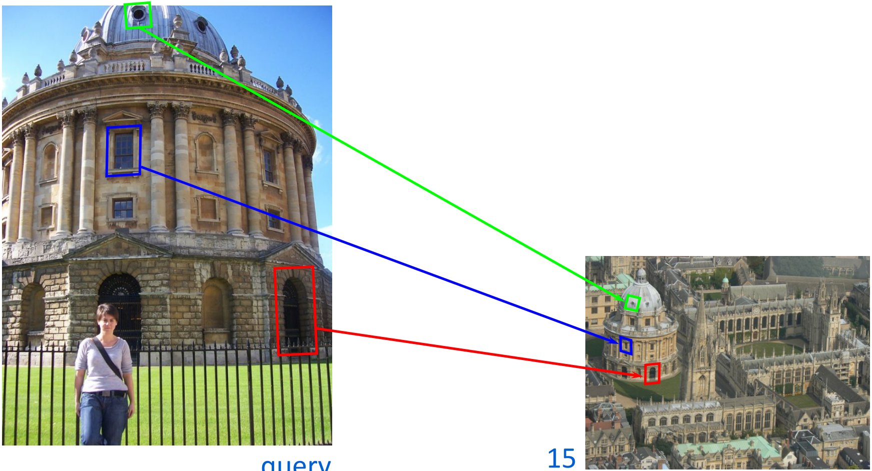
descriptor matching



query

15

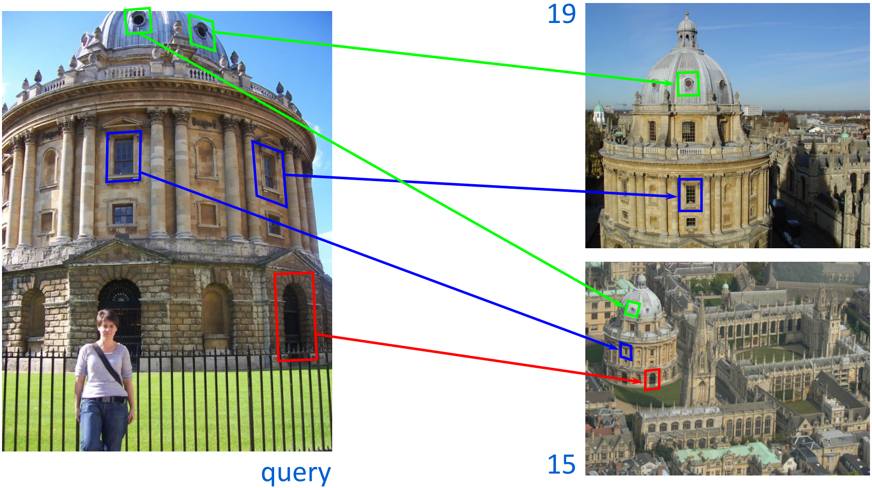
descriptor matching



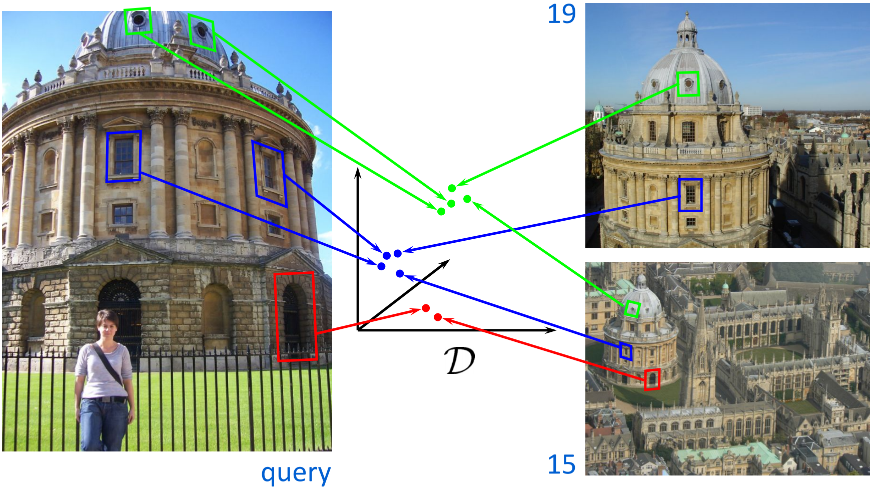
query

15

descriptor matching

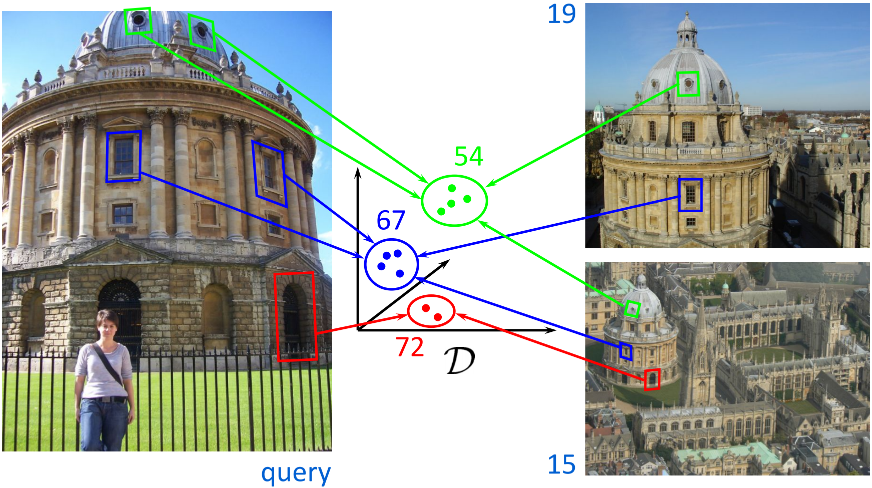


descriptor matching



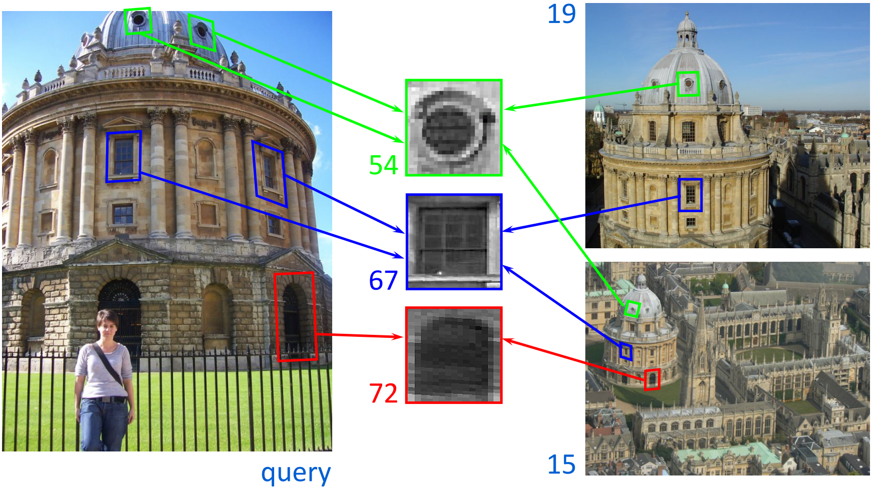
vector quantization → visual words

[Sivic and Zisserman, ICCV 2003]

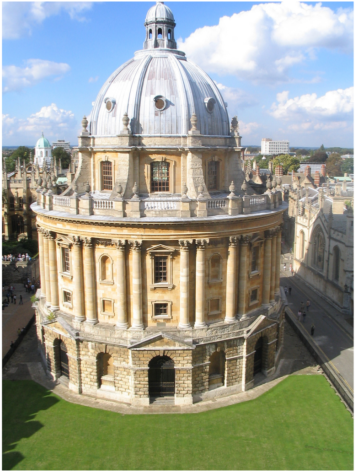


vector quantization → visual words

[Sivic and Zisserman, ICCV 2003]

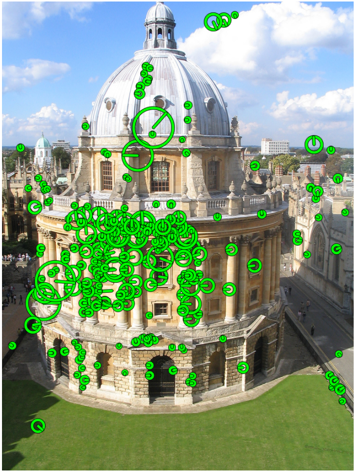


spatial matching



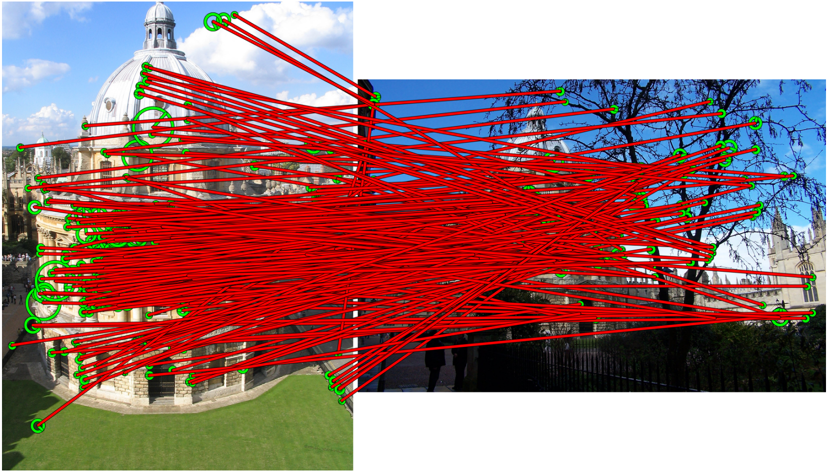
original images

spatial matching



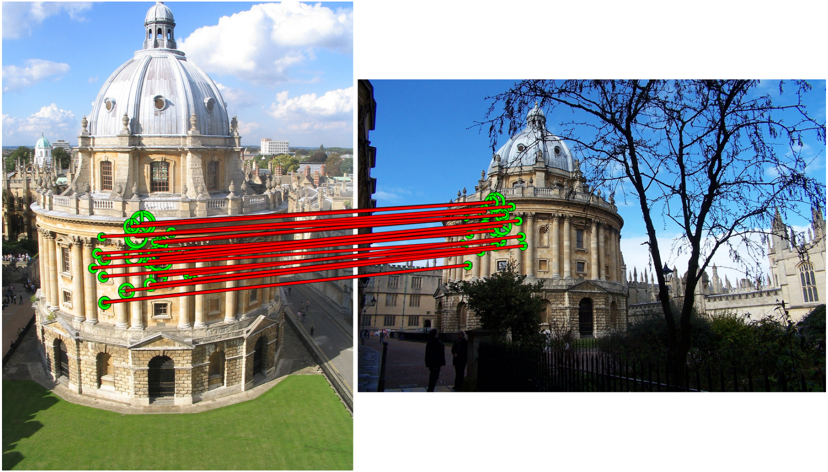
local features

spatial matching



tentative correspondences

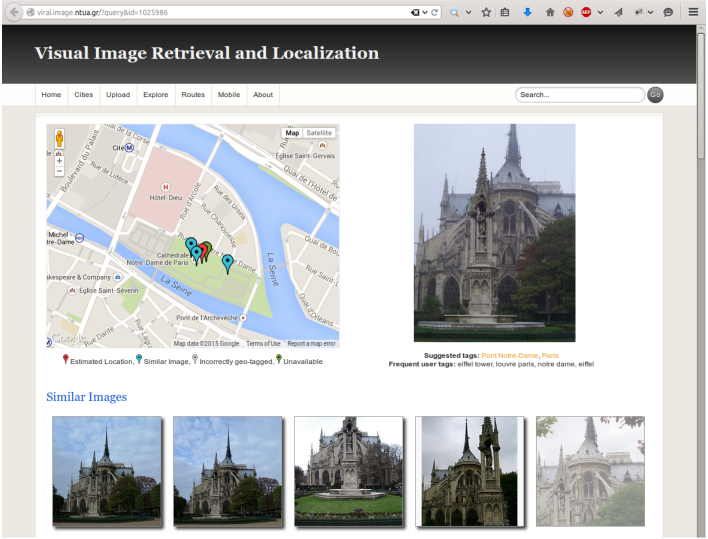
spatial matching



inliers

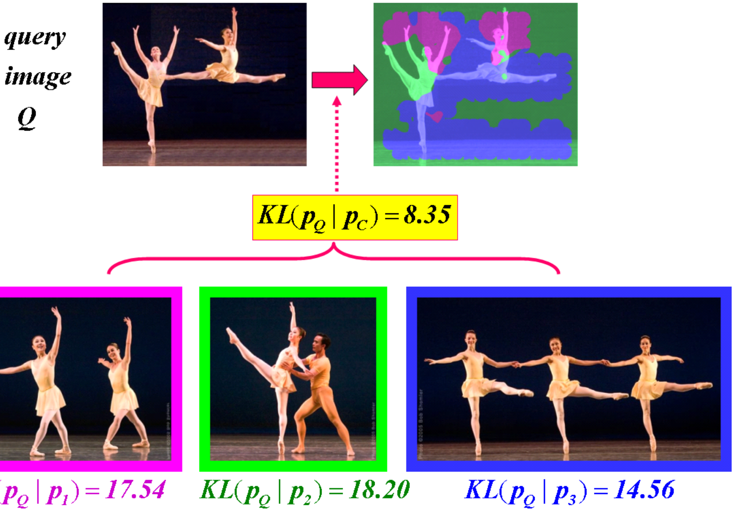
applications

instance recognition [Kalantidis et al. 2011]



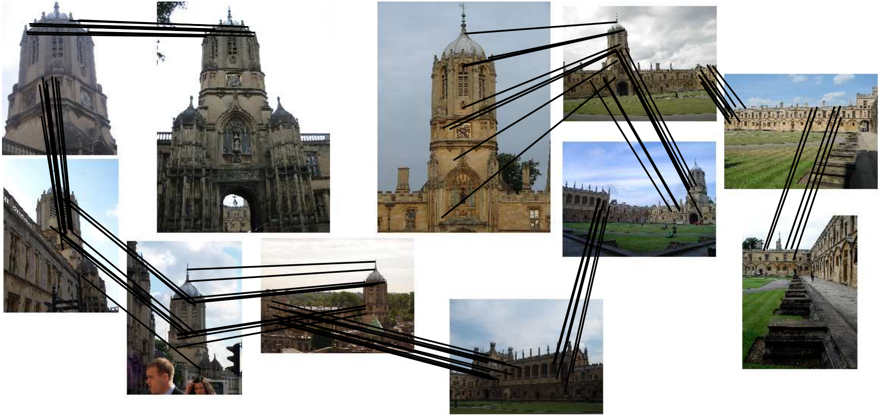
applications

class recognition [Boiman et al. 2008]



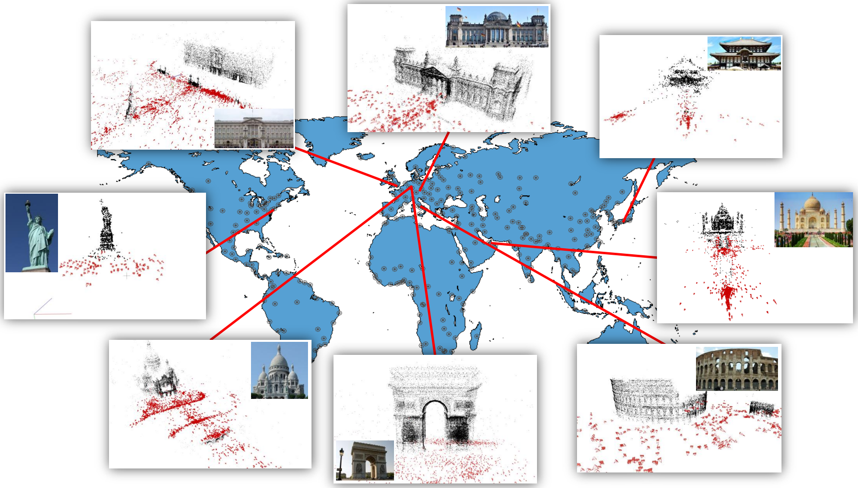
applications

object mining [Chum & Matas 2008]



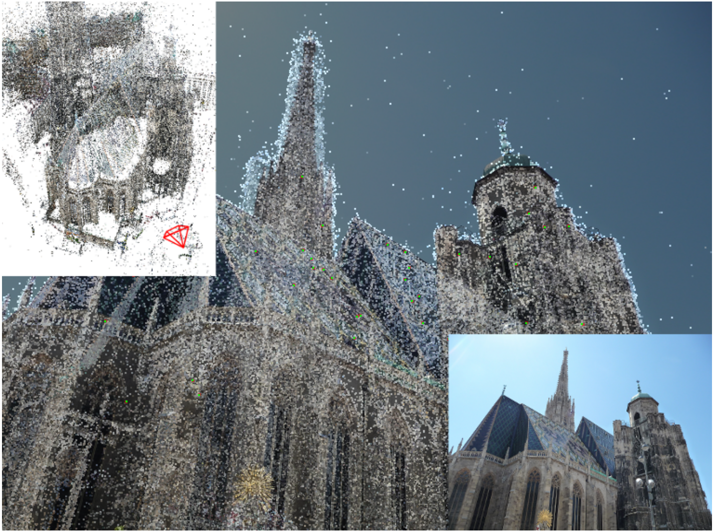
applications

reconstruction [Heinly et al. 2015]



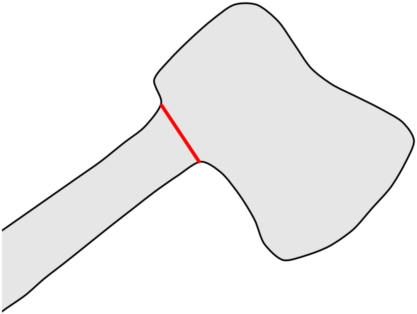
applications

pose estimation [Sattler et al. 2012]



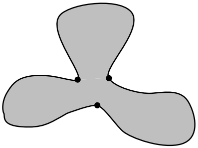
overview

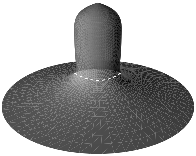
- planar shape decomposition
- local feature detection
- feature geometry & spatial matching
- descriptors, kernels & embeddings
- nearest neighbor search
- clustering
- mining, location & instance recognition



planar shape decomposition

psychophysical studies



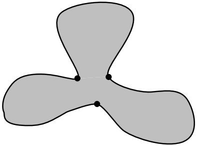


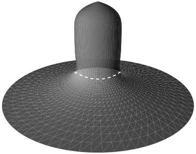
minima rule

[Hoffman & Richards 1983]

“divide a silhouette into parts at concave cusps and negative minima of curvature”

psychophysical studies

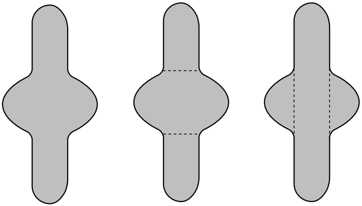




minima rule

[Hoffman & Richards 1983]

"divide a silhouette into parts at concave cusps and negative minima of curvature"

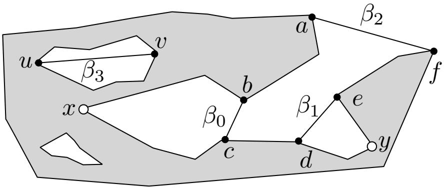


short-cut rule

[Singh *et al.* 1999]

"divide a silhouette into parts using the shortest possible cuts"

computational models



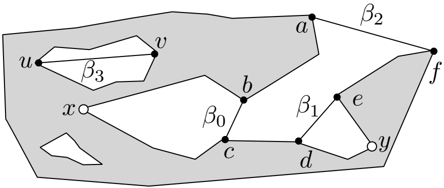
current work

e.g. dual space decomposition

[Liu *et al.* 2014]

- mostly based on convexity
- requires optimization
- rules applied indirectly

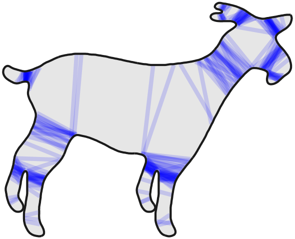
computational models



current work

e.g. dual space decomposition
[Liu *et al.* 2014]

- mostly based on convexity
- requires optimization
- rules applied indirectly



quantitative evaluation

practically non-existent until
[De Winter & Wagemans 2006]

medial axis

planar shape

- a set $X \subset \mathbb{R}^2$ whose boundary ∂X is a finite union of disjoint simple closed curves, such that for each curve there is a parametrization $\alpha : [0,1] \to \partial X$ by arc length that is piecewise smooth

distance map

- maps each point $x \in X$ to its minimal distance to boundary ∂X

$$\mathcal{D}(X)(x) = \inf_{y \in \partial X} d(x, y)$$

projection

- the set of points on ∂X at minimal distance to x

$$\pi(x) = \{y \in \partial X : d(x, y) = \mathcal{D}(X)(x)\}$$

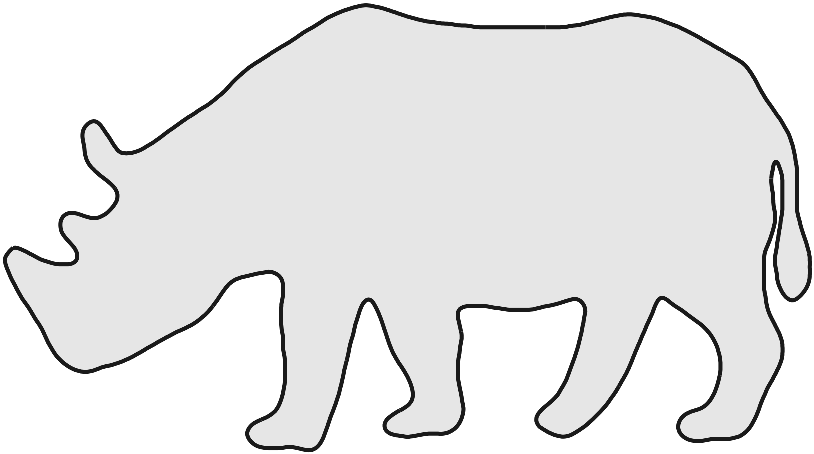
medial axis

- the set of points with more than one projection points

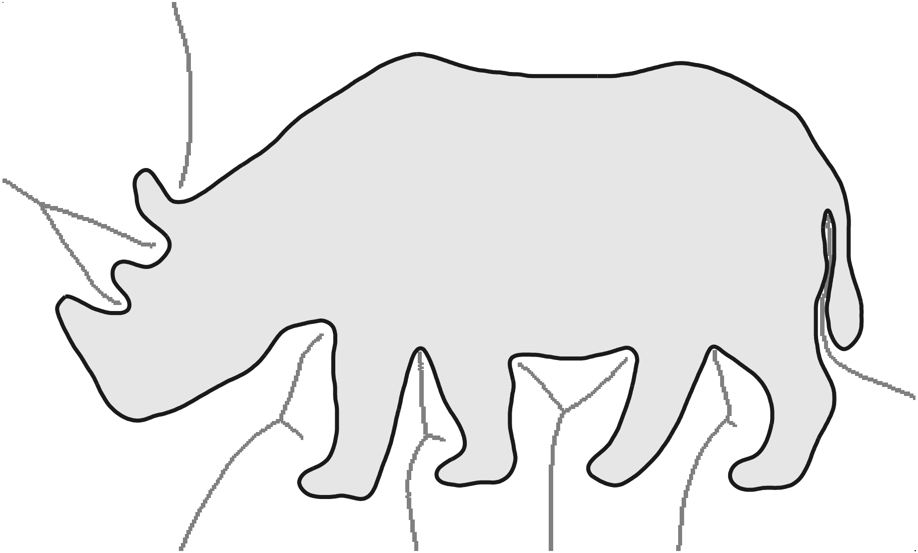
$$\mathcal{M}(X) = \{x \in \mathbb{R}^2 : |\pi(x)| > 1\}$$

medial axis decomposition

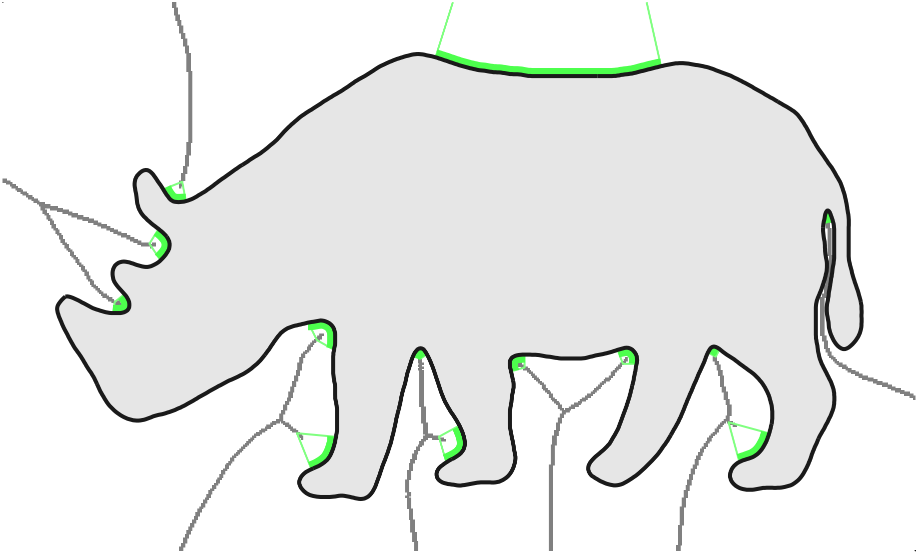
[Papanelopoulos & Avrithis, ongoing]



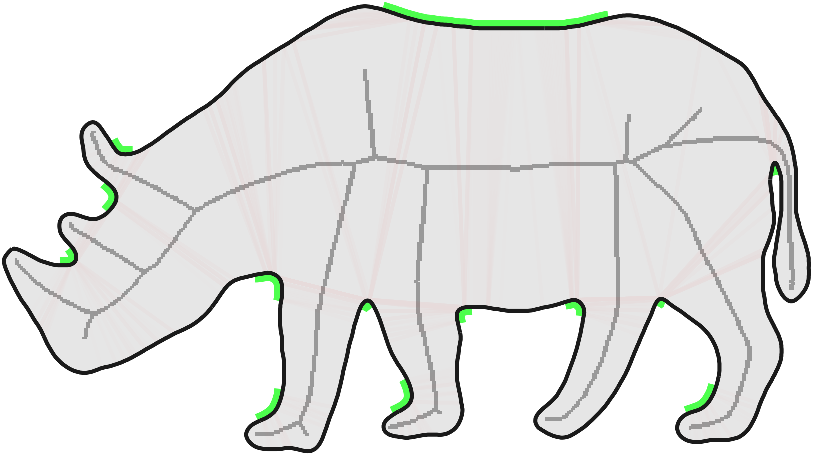
exterior medial axis



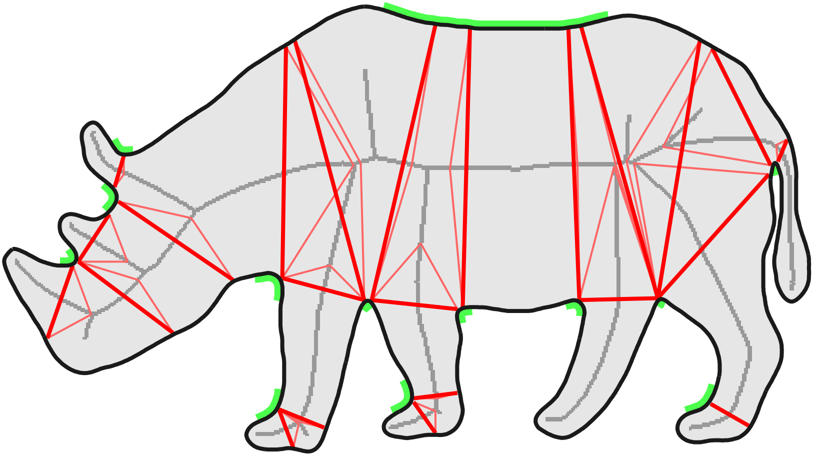
concave corners and “locale”



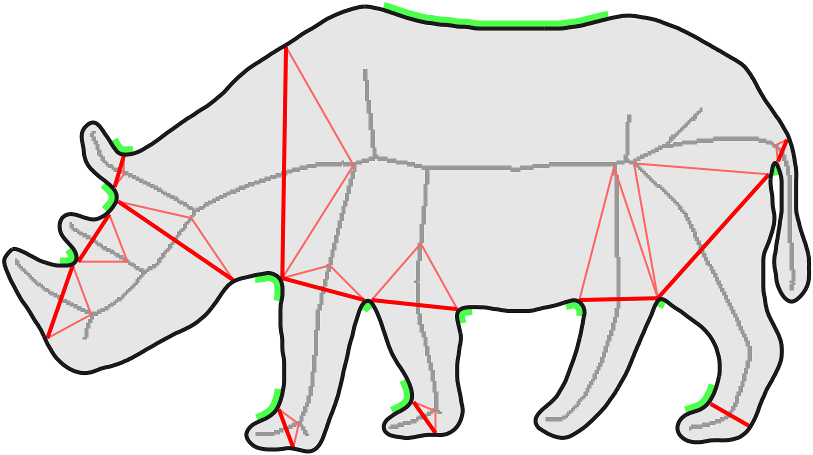
interior medial axis and raw cuts



cut equivalence on corners and branches



local convexity and short-cut rule



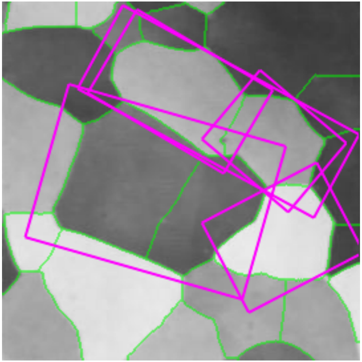
quantitative evaluation

	average		majority	
	H	R	H	R
DCE	0.208	0.497	0.188	0.466
SB	0.163	0.402	0.131	0.335
MD	0.151	0.371	0.126	0.328
FD	0.145	0.350	0.112	0.267
ACD	0.128	0.323	0.092	0.251
MAD	0.157	**0.193**	0.118	**0.154**
CBE	**0.111**	0.288	**0.069**	0.186
Human	–	–	0.104	0.137

H = Hamming distance; R = Rand index

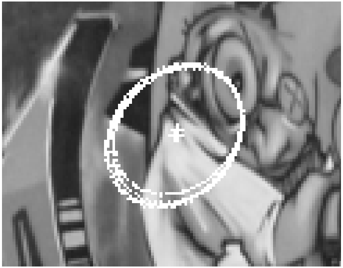
medial axis decomposition...

- practically “reads off” all information from the medial axis
- requires no differentiation
- requires no optimization
- is based on local decisions only
- can use arbitrary salience measures



local feature detection

feature detectors

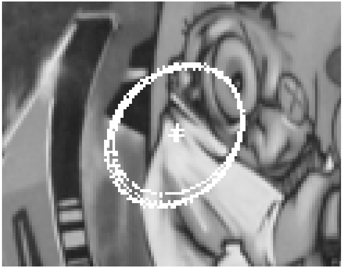


Hessian affine

[Mikolajczyk & Schmid 2004]

- de facto standard in visual search
- too many responses

feature detectors



Hessian affine

[Mikolajczyk & Schmid 2004]

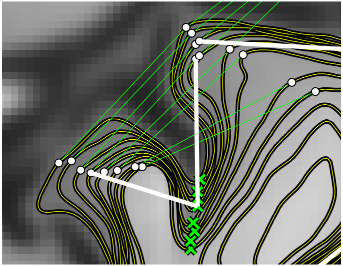
- de facto standard in visual search
- too many responses

maximally stable extremal regions

[Matas *et al.* 2002]

- arbitrary shape
- too constrained

feature detectors

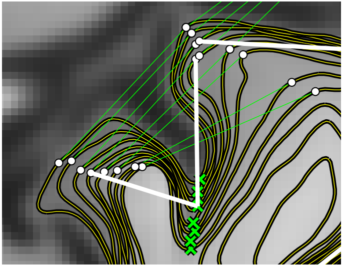


affine frames on isophotes

[Perdoch *et al.* 2007]

- only local stability
- based on bitangents

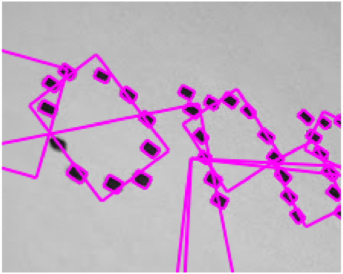
feature detectors



affine frames on isophotes

[Perdoch *et al.* 2007]

- only local stability
- based on bitangents



medial features

[Avrithis & Rapantzikos 2011]

medial features

[Avrithis & Rapantzikos, ICCV 2011]

additively weighted distance map

- given a non-increasing function $f : X \to \mathbb{R}$ of gradient strength, where X is the image plane,

$$\mathcal{D}(f)(x) = \min_{y \in X}\{d(x,y) + f(y)\}$$

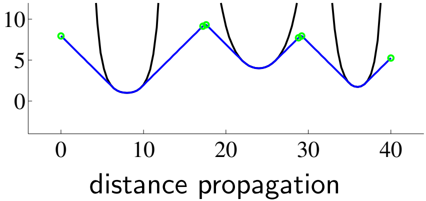
for $x \in X$

weighted medial

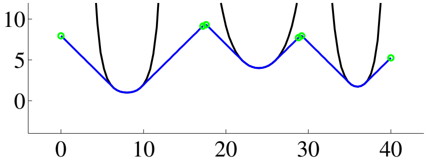
- similarly to unweighted case

$$\mathcal{M}(f) = \{x \in \mathbb{R}^2 : |\pi(x)| > 1\}$$

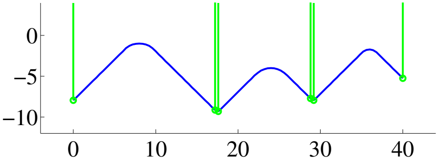
region/boundary duality



region/boundary duality

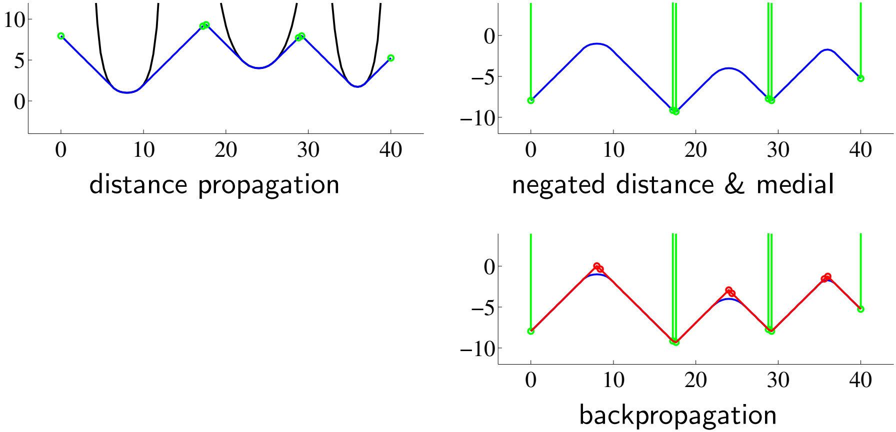


distance propagation

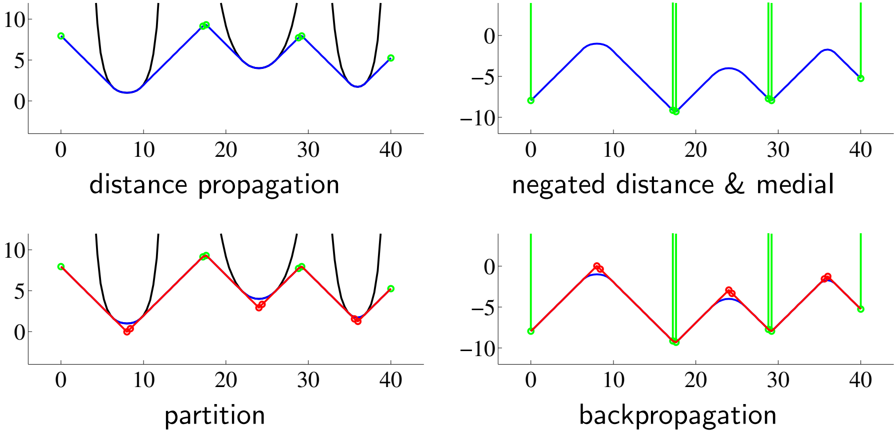


negated distance & medial

region/boundary duality

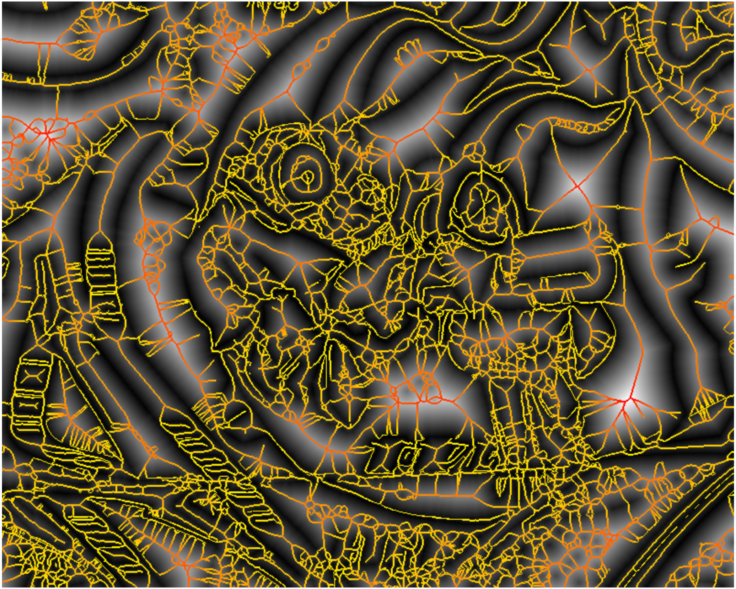


region/boundary duality

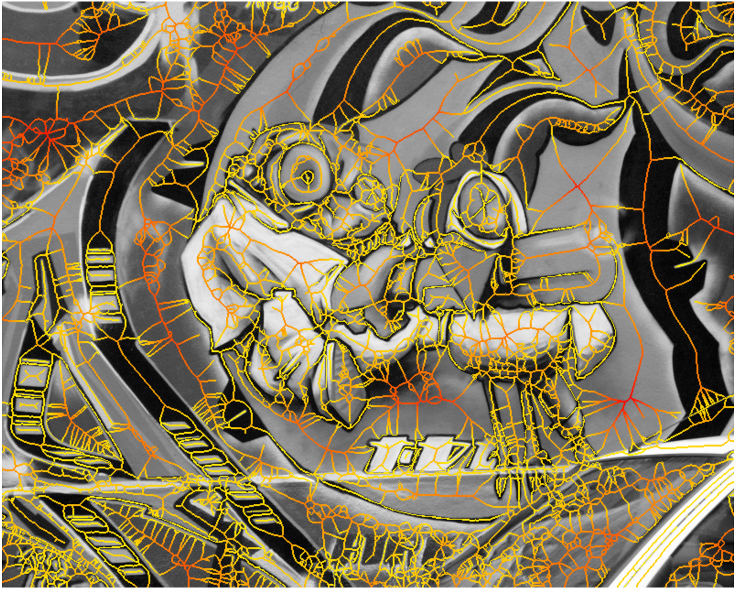


original image

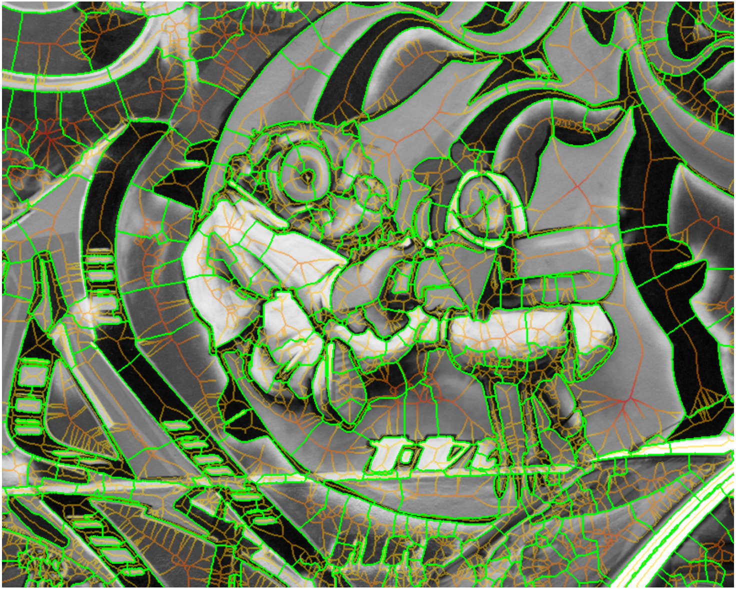
weighted distance map + medial



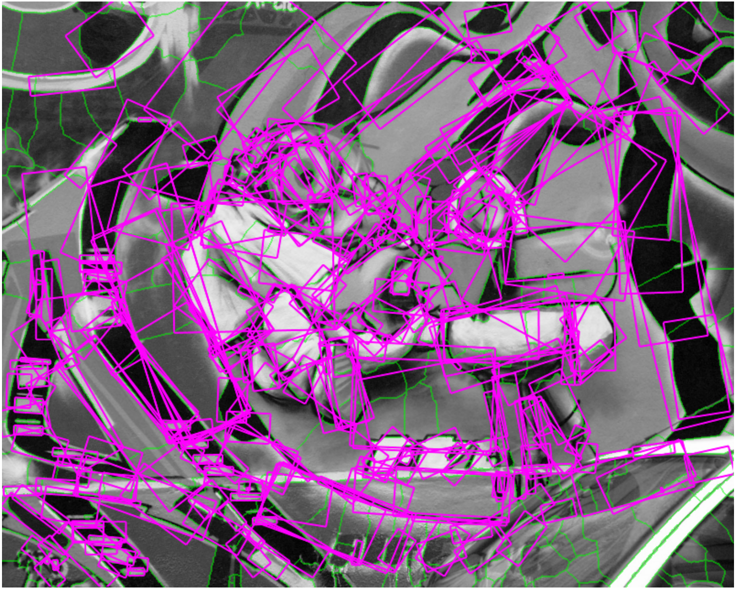
original image + weighted medial



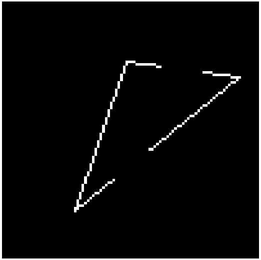
region/boundary duality & partition

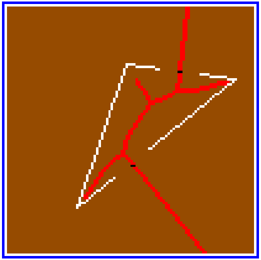


original image + features



fragmentation factor


binary input


point labels

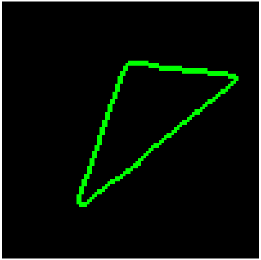


image partition

$$\phi(\kappa) = \frac{1}{a(\kappa)} \sum_{e \in E(\kappa)} w^2(x(e))$$

- selection criterion: is a region well-enclosed by boundaries?

law of closure & perceptual grouping

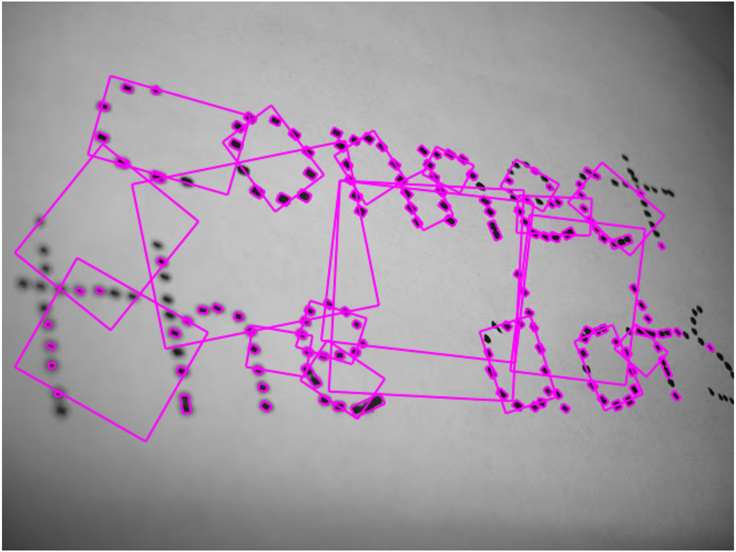


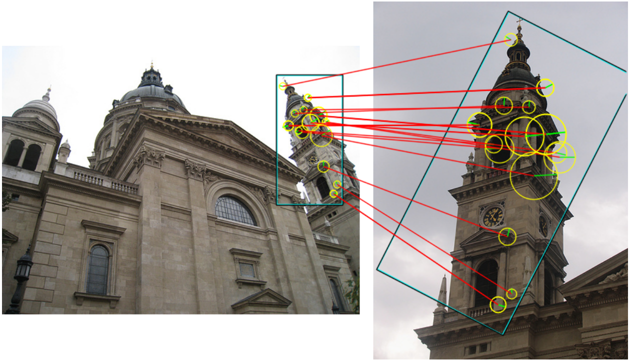
image search experiment

mAP on Oxford 5k

mAP	Inv. index		Re-ranking	
Detector	50k	200k	50k	200k
MFD	**0.515**	**0.580**	**0.568**	**0.617**
Hessian-affine	0.488	0.573	0.537	0.614
MSER	0.473	0.544	0.537	0.589
SURF	0.488	0.531	0.497	0.536
SIFT	0.395	0.457	0.434	0.495

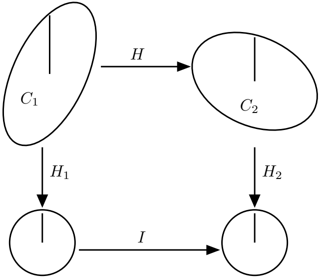
medial features...

- have arbitrary scale and shape
- are not contrained to extremal regions
- decompose shapes into parts
- capture law of closure



feature geometry & spatial matching

spatial matching for instance recognition

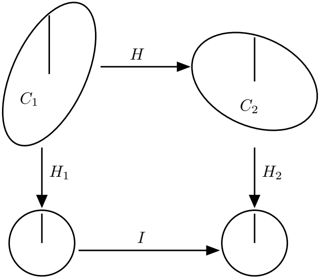


fast spatial matching

[Philbin *et al.* 2007]

- RANSAC variant
- single-correspondence hypotheses
- enumerate them all—$O(n^2)$

spatial matching for instance recognition



fast spatial matching

[Philbin *et al.* 2007]

- RANSAC variant
- single-correspondence hypotheses
- enumerate them all—$O(n^2)$

scale-invariant features

[Lowe 1999]

- Hough voting in 4d transformation space
- verification needed—still $O(n^2)$

spatial matching for class recognition

$$x^\star = \arg \max_{x \in \{0,1\}^n} x^\top A x$$

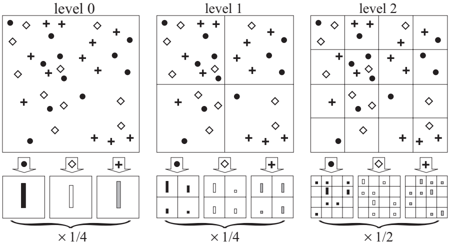
spectral matching

[Leordeanu & Hebert *et al.* 2005]

- based on pairwise affinity
- mapping constraints
- relaxed to an eigenvalue problem

spatial matching for class recognition

$$x^\star = \arg \max_{x \in \{0,1\}^n} x^\top A x$$



spectral matching

[Leordeanu & Hebert *et al.* 2005]

- based on pairwise affinity
- mapping constraints
- relaxed to an eigenvalue problem

spatial pyramid matching

[Lazebnik *et al.* 2006]

- flexible matching
- non-invariant

Hough pyramid matching

[Tolias & Avrithis, ICCV 2011]

- do not seek for inliers
- rather, look for hypotheses that agree with each other
- Hough voting in the 4d transformation space

$$F(c) = F(q)F(p)^{-1} = \begin{bmatrix} M(c) & \mathbf{t}(c) \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

$$f(c) = (x(c), y(c), \sigma(c), \theta(c))$$

- pyramid matching in the transformation space

$$s(c) = g(b_0) + \sum_{k=1}^{L-1} 2^{-k}\{g(b_k) - g(b_{k-1})\}$$

$$s(C) = \sum_{c \in C \setminus X} w(c)s(c)$$

Hough pyramid matching

[Tolias & Avrithis, ICCV 2011]

- do not seek for inliers
- rather, look for hypotheses that agree with each other
- Hough voting in the 4d transformation space

$$F(c) = F(q)F(p)^{-1} = \begin{bmatrix} M(c) & \mathbf{t}(c) \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

$$f(c) = (x(c), y(c), \sigma(c), \theta(c))$$

- pyramid matching in the transformation space

$$s(c) = g(b_0) + \sum_{k=1}^{L-1} 2^{-k}\{g(b_k) - g(b_{k-1})\}$$

$$s(C) = \sum_{c \in C \setminus X} w(c)s(c)$$

Hough pyramid matching

[Tolias & Avrithis, ICCV 2011]

- do not seek for inliers
- rather, look for hypotheses that agree with each other
- Hough voting in the 4d transformation space

$$F(c) = F(q)F(p)^{-1} = \begin{bmatrix} M(c) & \mathbf{t}(c) \\ \mathbf{0}^\top & 1 \end{bmatrix}$$

$$f(c) = (x(c), y(c), \sigma(c), \theta(c))$$

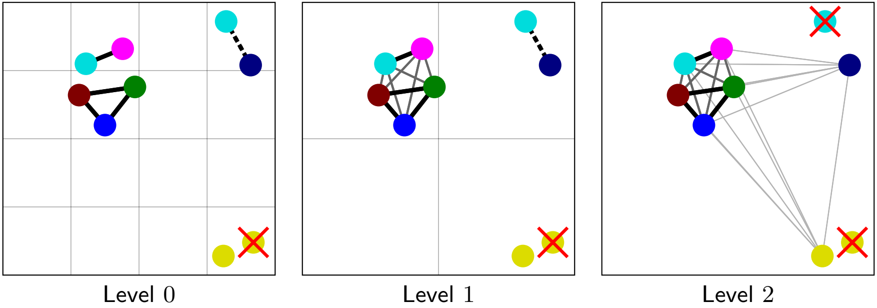
- pyramid matching in the transformation space

$$s(c) = g(b_0) + \sum_{k=1}^{L-1} 2^{-k}\{g(b_k) - g(b_{k-1})\}$$

$$s(C) = \sum_{c \in C \setminus X} w(c)s(c)$$

toy example

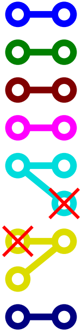
Hough pyramid



toy example

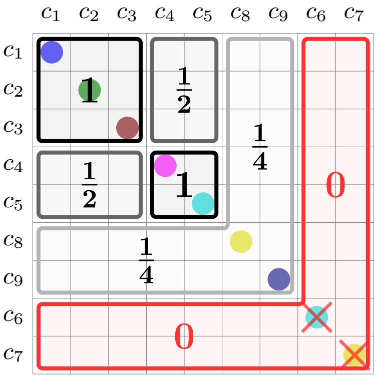
correspondences, strengths

	p q	strength
c_1		$(2 + \frac{1}{2}2 + \frac{1}{4}2)w(c_1)$
c_2		$(2 + \frac{1}{2}2 + \frac{1}{4}2)w(c_2)$
c_3		$(2 + \frac{1}{2}2 + \frac{1}{4}2)w(c_3)$
c_4		$(1 + \frac{1}{2}3 + \frac{1}{4}2)w(c_4)$
c_5		$(1 + \frac{1}{2}3 + \frac{1}{4}2)w(c_5)$
c_6		0
c_7		0
c_8		$\frac{1}{4}6w(c_8)$
c_9		$\frac{1}{4}6w(c_9)$



toy example

affinity matrix

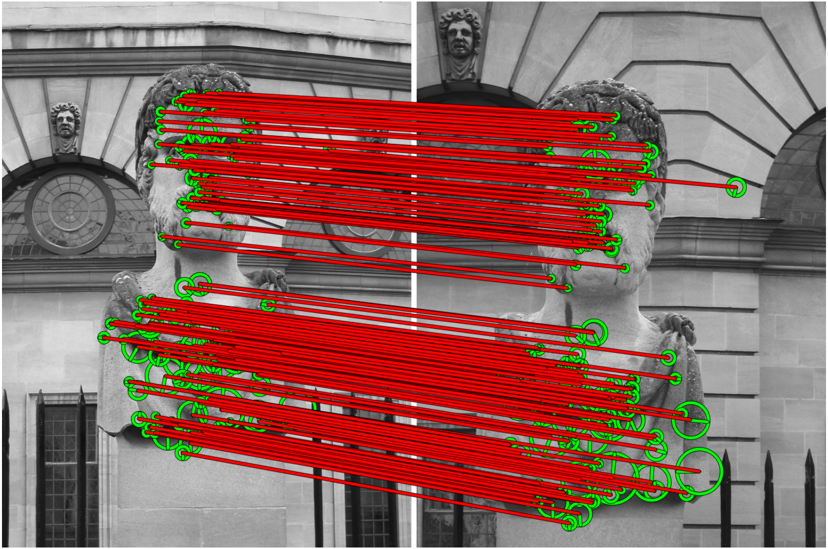


Hough pyramid matching ...

- is invariant to similarity transformations
- is flexible, allowing non-rigid motion and multiple matching surfaces or objects
- imposes one-to-one mapping

examples

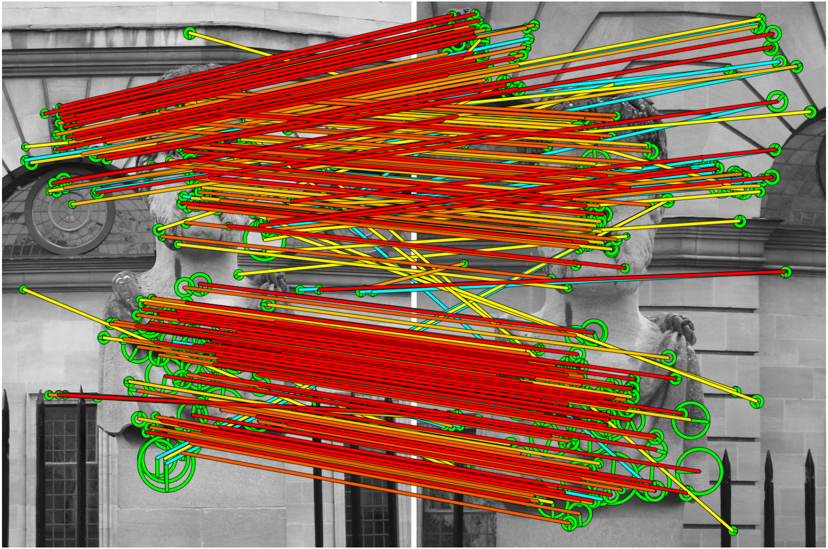
HPM vs FSM [Philbin et al. 2007]



fast spatial matching

examples

HPM vs FSM [Philbin et al. 2007]



Hough pyramid matching

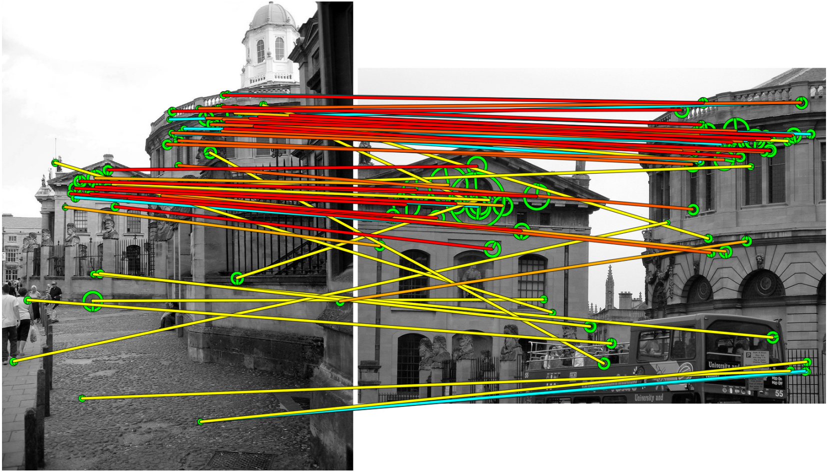
examples

HPM vs FSM [Philbin et al. 2007]

fast spatial matching

examples

HPM vs FSM [Philbin et al. 2007]



Hough pyramid matching

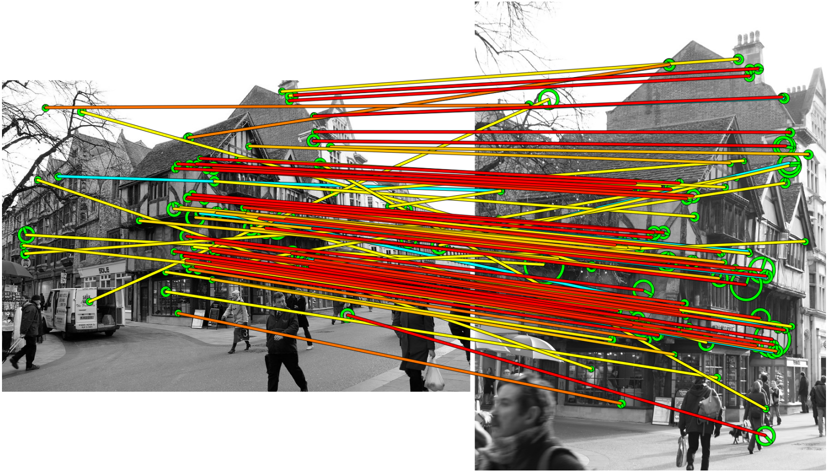
examples

HPM vs FSM [Philbin et al. 2007]

fast spatial matching

examples

HPM vs FSM [Philbin et al. 2007]



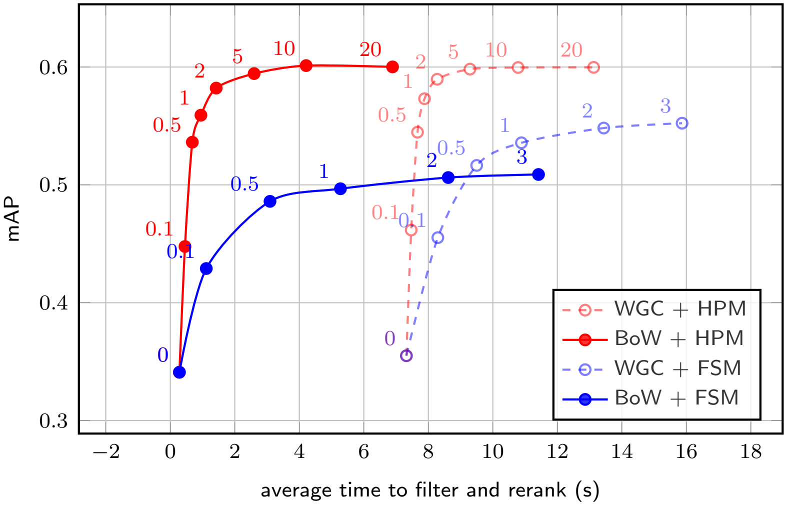
Hough pyramid matching

Hough pyramid matching . . .

- is non-iterative, and linear in the number of correspondences
- in a given query time, can re-rank one order of magnitude more images than the state of the art
- typically needs less than one millisecond to match a pair of images, on average

performance vs time

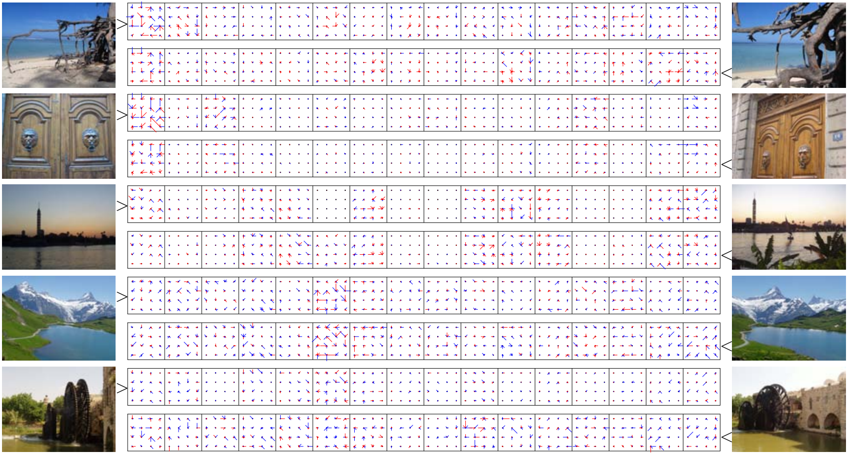
on World Cities 2M



comparison to state of the art

[Avrithis & Tolias, IJCV 2014]

method	Ox5K	Ox105K	Paris	Holidays
HPM (this work)	**0.789**	**0.730**	0.725	**0.790**
[Shen *et al.* 2012]	0.752	0.729	**0.741**	0.762
GVP [Zhang *et al.* 2011]	0.696	-	-	-
SBoF [Cao *et al.* 2010]	0.656	-	0.632	-
[Perdoch *et al.* 2009]	**0.789**	0.726	-	0.715
FSM [Philbin *et al.* 2007]	0.647	0.541	-	-



descriptors, kernels & embeddings

set kernels & embeddings

normalized sum set kernel [Bo & Sminchisescu 2009]

- given kernel function k, define (finite) set kernel

$$K(X,Y) = \frac{1}{|X||Y|} \sum_{x \in X} \sum_{y \in Y} k(x,y)$$

example: Gaussian mixtures [Liu & Perronnin 2008]

- model set X by finite mixture distribution

$$f_X(z) = \frac{1}{|X|} \sum_{x \in X} \mathcal{N}(z|x,\Sigma), \quad z \in \mathbb{R}^d$$

- then,

$$\langle f_X, f_Y \rangle = \frac{1}{|X||Y|} \sum_{x \in X} \sum_{y \in Y} \mathcal{N}(x|y, 2\Sigma)$$

set kernels & embeddings

normalized sum set kernel [Bo & Sminchisescu 2009]

- given kernel function k, define (finite) set kernel

$$K(X,Y) = \frac{1}{|X||Y|} \sum_{x \in X} \sum_{y \in Y} k(x,y)$$

example: Gaussian mixtures [Liu & Perronnin 2008]

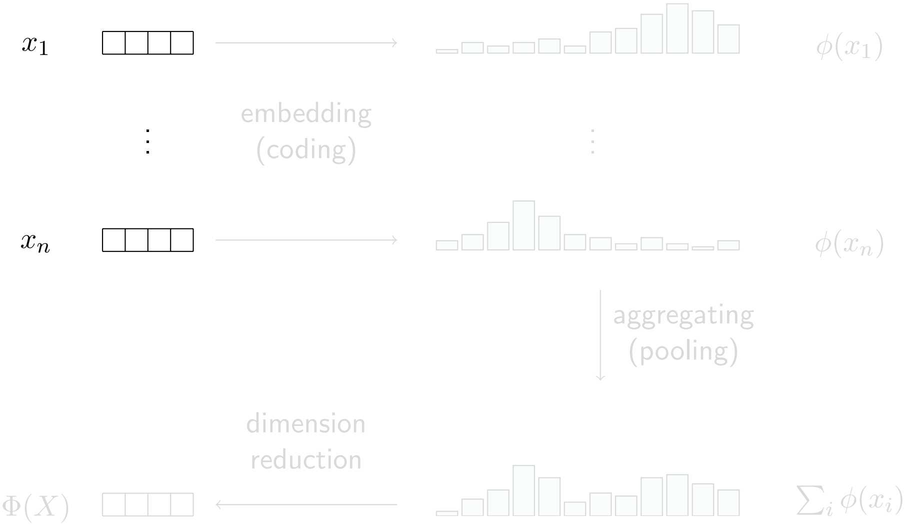
- model set X by finite mixture distribution

$$f_X(z) = \frac{1}{|X|} \sum_{x \in X} \mathcal{N}(z|x,\Sigma), \quad z \in \mathbb{R}^d$$

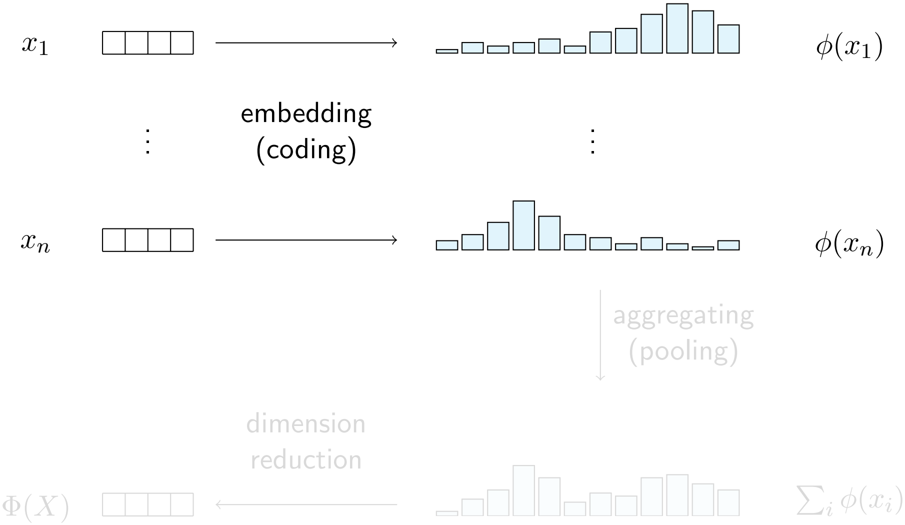
- then,

$$\langle f_X, f_Y \rangle = \frac{1}{|X||Y|} \sum_{x \in X} \sum_{y \in Y} \mathcal{N}(x|y, 2\Sigma)$$

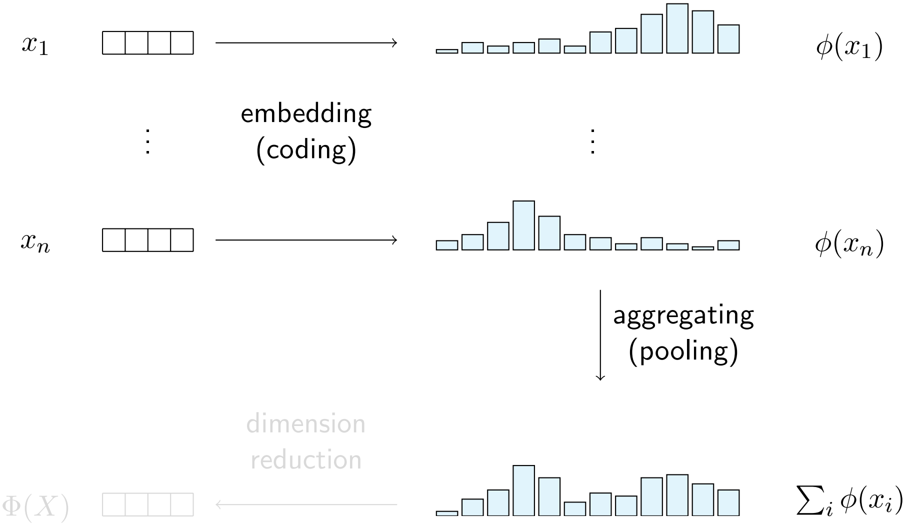
explicit feature maps



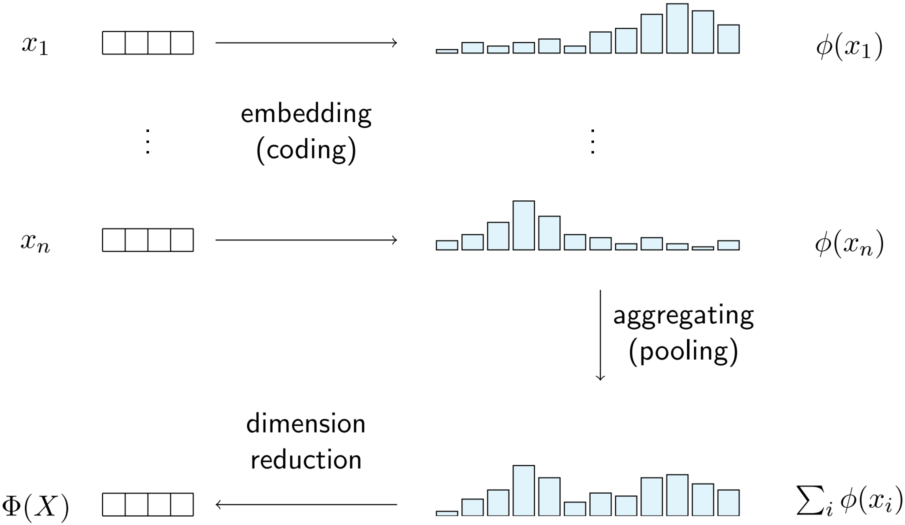
explicit feature maps



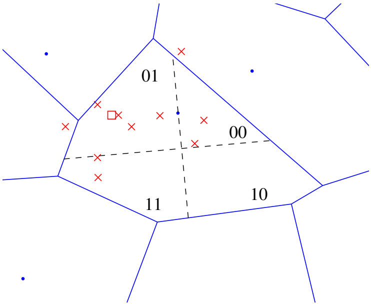
explicit feature maps



explicit feature maps



two different perspectives



Hamming embedding

[Jégou *et al.* 2008]

- large vocabulary
- binary signature & descriptor voting
- not aggregated
- selective: discard weak votes

VLAD

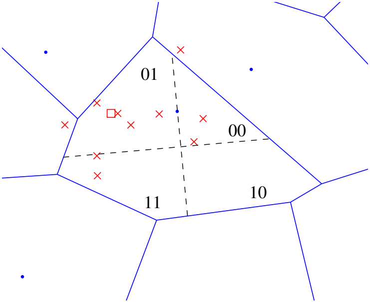
[Jégou *et al.* 2010]

- small vocabulary
- one aggregated vector per cell
- linear operation
- not selective

$$V(X_c) = \sum_{x \in X_c} x - q(x)$$

$$X_c = \{x \in X : q(x) = c\}$$

two different perspectives



$$V(X_c) = \sum_{x \in X_c} x - q(x)$$

$$X_c = \{x \in X : q(x) = c\}$$

Hamming embedding

[Jégou *et al.* 2008]

- large vocabulary
- binary signature & descriptor voting
- not aggregated
- selective: discard weak votes

VLAD

[Jégou *et al.* 2010]

- small vocabulary
- one aggregated vector per cell
- linear operation
- not selective

common model: image similarity

$$K(X,Y) = \gamma(X)\,\gamma(Y) \sum_{c \in C} w_c \kappa\left(X_c, Y_c\right)$$

normalization factor

cell weighting

cell similarity

common model: image similarity

$$K(X,Y) = \gamma(X)\,\gamma(Y) \sum_{c \in C} w_c \kappa\left(X_c, Y_c\right)$$

normalization factor

cell weighting

cell similarity

common model: cell similarity

non aggregated

$$\kappa_n(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} \sigma \left(\phi(x)^\top \phi(y) \right)$$

selectivity function

descriptor representation (residual, binary, scalar)

aggregated

$$\kappa_a(X_c, Y_c) = \sigma \left\{ \psi \left(\sum_{x \in X_c} \phi(x) \right)^\top \psi \left(\sum_{y \in Y_c} \phi(y) \right) \right\} = \sigma \left(\Phi(X_c)^\top \Phi(Y_c) \right)$$

normalization (ℓ_2, power-law)

cell representation

common model: cell similarity

non aggregated

$$\kappa_n(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} \sigma \left(\phi(x)^\top \phi(y) \right)$$

selectivity function

descriptor representation (residual, binary, scalar)

aggregated

$$\kappa_a(X_c, Y_c) = \sigma \left\{ \psi \left(\sum_{x \in X_c} \phi(x) \right)^\top \psi \left(\sum_{y \in Y_c} \phi(y) \right) \right\} = \sigma \left(\Phi(X_c)^\top \Phi(Y_c) \right)$$

normalization (ℓ_2, power-law)

cell representation

common model: cell similarity

non aggregated

$$\kappa_n(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} \sigma \left(\phi(x)^\top \phi(y) \right)$$

selectivity function

descriptor representation (residual, binary, scalar)

aggregated

$$\kappa_a(X_c, Y_c) = \sigma \left\{ \psi \left(\sum_{x \in X_c} \phi(x) \right)^\top \psi \left(\sum_{y \in Y_c} \phi(y) \right) \right\} = \sigma \left(\Phi(X_c)^\top \Phi(Y_c) \right)$$

normalization (ℓ_2, power-law)

cell representation

common model: cell similarity

non aggregated

$$\kappa_n(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} \sigma \left(\phi(x)^\top \phi(y) \right)$$

selectivity function

descriptor representation (residual, binary, scalar)

aggregated

$$\kappa_a(X_c, Y_c) = \sigma \left\{ \psi \left(\sum_{x \in X_c} \phi(x) \right)^\top \psi \left(\sum_{y \in Y_c} \phi(y) \right) \right\} = \sigma \left(\Phi(X_c)^\top \Phi(Y_c) \right)$$

normalization (ℓ_2, power-law)

cell representation

BoW, HE and VLAD in the common model

model	$\kappa(X_c, Y_c)$	$\phi(x)$	$\sigma(u)$	$\psi(z)$	$\Phi(X_c)$
BoW	κ_n or κ_a	1	u	z	$\lvert X_c \rvert$
HE	κ_n only	$\hat{b}_x$	$w\left(\frac{B}{2}(1-u)\right)$	—	—
VLAD	κ_n or κ_a	$r(x)$	u	z	$V(X_c)$

BoW $$\kappa(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} 1 = |X_c| \times |Y_c|$$

HE $$\kappa(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} w(\mathrm{h}(b_x, b_y))$$

VLAD $$\kappa(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} r(x)^\top r(y) = V(X_c)^\top V(Y_c)$$

$$\kappa_n(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} \sigma\left(\phi(x)^\top \phi(y)\right)$$

$$\kappa_a(X_c, Y_c) = \sigma\left\{\psi\left(\sum_{x \in X_c} \phi(x)\right)^\top \psi\left(\sum_{y \in Y_c} \phi(y)\right)\right\} = \sigma\left(\Phi(X_c)^\top \Phi(Y_c)\right)$$

BoW, HE and VLAD in the common model

model	$\kappa(X_c, Y_c)$	$\phi(x)$	$\sigma(u)$	$\psi(z)$	$\Phi(X_c)$
BoW	κ_n or κ_a	1	u	z	$\|X_c\|$
HE	κ_n only	$\hat{b}_x$	$w\left(\frac{B}{2}(1-u)\right)$	—	—
VLAD	κ_n or κ_a	$r(x)$	u	z	$V(X_c)$

BoW $$\kappa(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} 1 = |X_c| \times |Y_c|$$

HE $$\kappa\left(X_c, Y_c\right) = \sum_{x \in X_c} \sum_{y \in Y_c} w\left(\mathrm{h}\left(b_x, b_y\right)\right)$$

VLAD $$\kappa\left(X_c, Y_c\right) = \sum_{x \in X_c} \sum_{y \in Y_c} r(x)^\top r(y) = V(X_c)^\top V(Y_c)$$

$$\kappa_n(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} \sigma\left(\phi(x)^\top \phi(y)\right)$$

$$\kappa_a(X_c, Y_c) = \sigma\left\{\psi\left(\sum_{x \in X_c} \phi(x)\right)^\top \psi\left(\sum_{y \in Y_c} \phi(y)\right)\right\} = \sigma\left(\Phi(X_c)^\top \Phi(Y_c)\right)$$

BoW, HE and VLAD in the common model

model	$\kappa(X_c, Y_c)$	$\phi(x)$	$\sigma(u)$	$\psi(z)$	$\Phi(X_c)$
BoW	κ_n or κ_a	1	u	z	$\lvert X_c \rvert$
HE	κ_n only	$\hat{b}_x$	$w\left(\frac{B}{2}(1-u)\right)$	—	—
VLAD	κ_n or κ_a	$r(x)$	u	z	$V(X_c)$

BoW $$\kappa(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} 1 = |X_c| \times |Y_c|$$

HE $$\kappa\left(X_c, Y_c\right) = \sum_{x \in X_c} \sum_{y \in Y_c} w\left(\mathrm{h}\left(b_x, b_y\right)\right)$$

VLAD $$\kappa\left(X_c, Y_c\right) = \sum_{x \in X_c} \sum_{y \in Y_c} r(x)^\top r(y) = V(X_c)^\top V(Y_c)$$

$$\kappa_n(X_c, Y_c) = \sum_{x \in X_c} \sum_{y \in Y_c} \sigma\left(\phi(x)^\top \phi(y)\right)$$

$$\kappa_a(X_c, Y_c) = \sigma\left\{\psi\left(\sum_{x \in X_c} \phi(x)\right)^\top \psi\left(\sum_{y \in Y_c} \phi(y)\right)\right\} = \sigma\left(\Phi(X_c)^\top \Phi(Y_c)\right)$$

aggregated selective match kernel

[Tolias et al. ICCV 2013]

- cell similarity

$$\text{ASMK}(X_c, Y_c) = \sigma_\alpha\left(\hat{V}(X_c)^\top \hat{V}(Y_c)\right)$$

- cell representation: ℓ_2-normalized aggregated residual

$$\Phi(X_c) = \hat{V}(X_c) = V(X_c)/\|V(X_c)\|$$

- selectivity function

$$\sigma_\alpha(u) = \begin{cases} \text{sgn}(u)|u|^\alpha, & u > \tau \\ 0, & \text{otherwise} \end{cases}$$

aggregated selective match kernel

[Tolias et al. ICCV 2013]

- cell similarity

$$\text{ASMK}(X_c, Y_c) = \sigma_\alpha \left(\hat{V}(X_c)^\top \hat{V}(Y_c) \right)$$

- cell representation: ℓ_2-normalized aggregated residual

$$\Phi(X_c) = \hat{V}(X_c) = V(X_c)/\|V(X_c)\|$$

- selectivity function

$$\sigma_\alpha(u) = \begin{cases} \text{sgn}(u)|u|^\alpha, & u > \tau \\ 0, & \text{otherwise} \end{cases}$$

aggregated selective match kernel

[Tolias et al. ICCV 2013]

- cell similarity

$$\text{ASMK}(X_c, Y_c) = \sigma_\alpha\left(\hat{V}(X_c)^\top \hat{V}(Y_c)\right)$$

- cell representation: ℓ_2-normalized aggregated residual

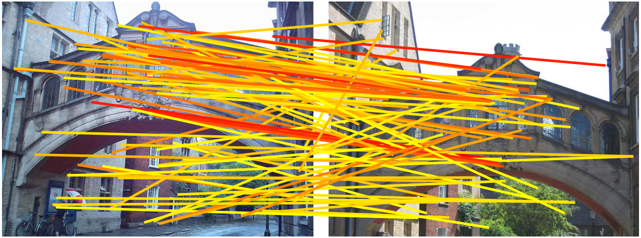
$$\Phi(X_c) = \hat{V}(X_c) = V(X_c)/\|V(X_c)\|$$

- selectivity function

$$\sigma_\alpha(u) = \begin{cases} \operatorname{sgn}(u)|u|^\alpha, & u > \tau \\ 0, & \text{otherwise} \end{cases}$$

impact of selectivity

$$\alpha = 1,\ \tau = 0.0$$

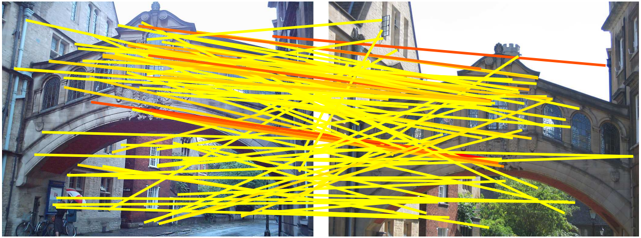


$$\alpha = 1,\ \tau = 0.25$$

thresholding removes false correspondences

impact of selectivity

$\alpha = 3, \ \tau = 0.0$



$\alpha = 3, \ \tau = 0.25$

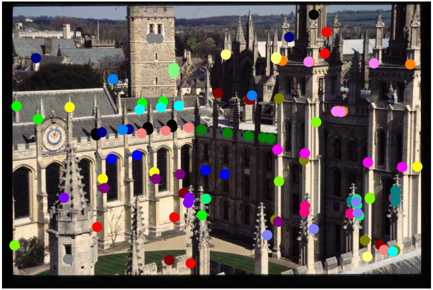
correspondences weighed based on confidence

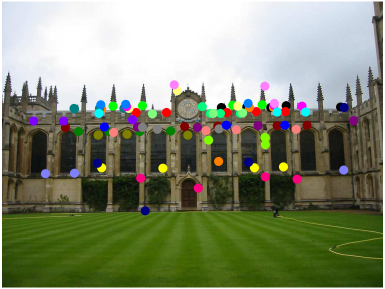
impact of aggregation & burstiness

$k = 128$ as in VLAD

impact of aggregation & burstiness

$k = 65\text{k}$ as in HE

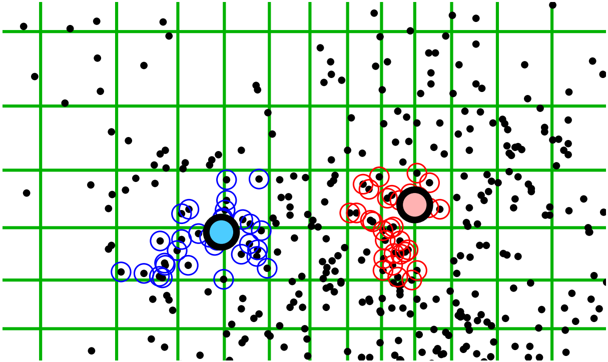





comparison to state of the art

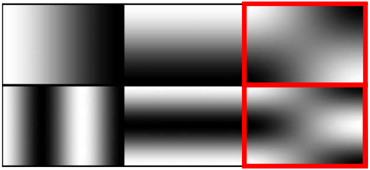
[Tolias et al. IJCV 2015]

Dataset	MA	Oxf5k	Oxf105k	Par6k	Holiday
ASMK*		76.4	69.2	74.4	80.0
ASMK*	×	80.4	75.0	77.0	81.0
ASMK		78.1	-	76.0	81.2
ASMK	×	81.7	-	78.2	82.2
HE [Jégou *et al.* '10]		51.7	-	-	74.5
HE [Jégou *et al.* '10]	×	56.1	-	-	77.5
HE-BURST [Jain *et al.* '10]		64.5	-	-	78.0
HE-BURST [Jain *et al.* '10]	×	67.4	-	-	79.6
Fine vocab. [Mikulík *et al.* '10]	×	74.2	67.4	74.9	74.9
AHE-BURST [Jain *et al.* '10]		66.6	-	-	79.4
AHE-BURST [Jain *et al.* '10]	×	69.8	-	-	81.9
Rep. structures [Torri *et al.* '13]	×	65.6	-	-	74.9
Locality [Tao *et al.* '14]	×	77.0	-	-	78.7



nearest neighbor search

binary codes

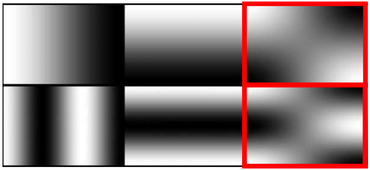


spectral hashing

[Weiss *et al.* 2008]

- similarity preserving, balanced, uncorrelated
- spectral relaxation
- out of sample extension: uniform assumption

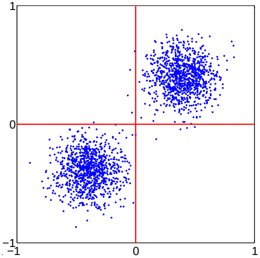
binary codes



spectral hashing

[Weiss *et al.* 2008]

- similarity preserving, balanced, uncorrelated
- spectral relaxation
- out of sample extension: uniform assumption



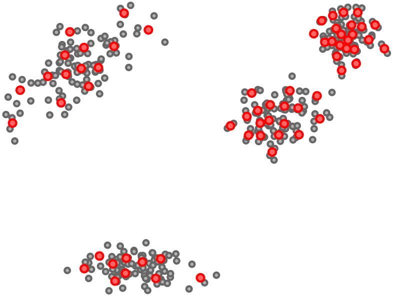
iterative quantization

[Gong & Lazebnik 2011]

- quantize to closest vertex of binary cube
- PCA followed by interleaved rotation and quantization

vector quantization

[Gray 1984]

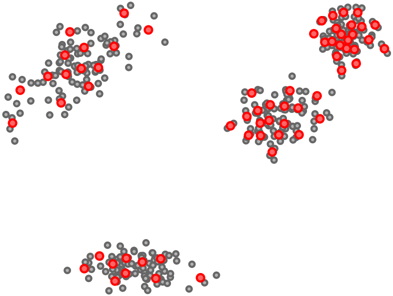


$$\text{minimize } E(C) = \sum_{\mathbf{x} \in X} \min_{\mathbf{c} \in C} \|\mathbf{x} - \mathbf{c}\|^2 = \sum_{\mathbf{x} \in X} \|\mathbf{x} - q(\mathbf{x})\|^2$$

distortion — dataset — codebook — quantizer

vector quantization

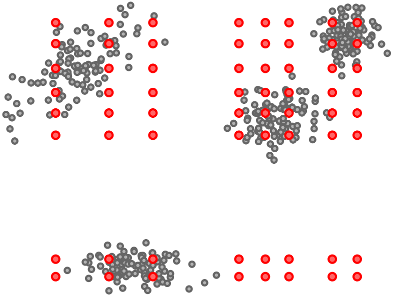
[Gray 1984]



- For small distortion → large $k = |C|$:
  - hard to train
  - too large to store
  - too slow to search

product quantization

[Jégou et al. 2011]

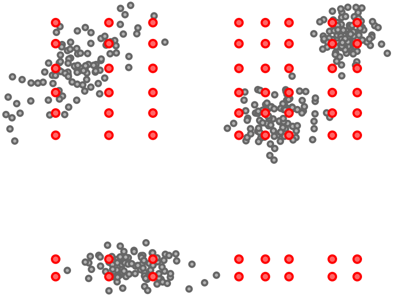


$$\text{minimize} \quad \sum_{\mathbf{x} \in X} \min_{\mathbf{c} \in C} \|\mathbf{x} - \mathbf{c}\|^2$$

$$\text{subject to} \quad C = C^1 \times \cdots \times C^m$$

product quantization

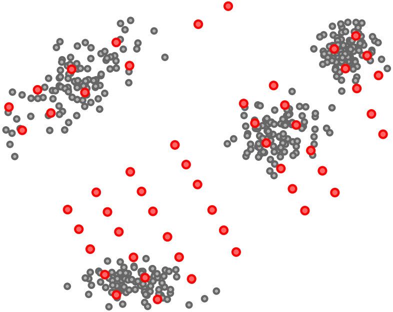
[Jégou et al. 2011]



- train: $q = (q^1, \dots, q^m)$ where $q^1, \dots, q^m$ obtained by VQ
- store: $|C| = k^m$ with $|C^1| = \dots = |C^m| = k$
- search: $\|\mathbf{y} - q(\mathbf{x})\|^2 = \sum_{j=1}^{m} \|\mathbf{y}^j - q^j(\mathbf{x}^j)\|^2$ where $q^j(\mathbf{x}^j) \in C^j$

optimized product quantization

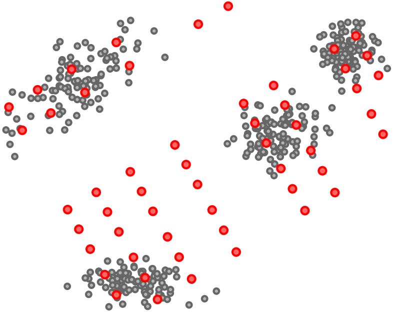
[Ge et al. 2013]



$$
\begin{aligned}
\text{minimize} \quad & \sum_{\mathbf{x} \in X} \min_{\hat{\mathbf{c}} \in \hat{C}} \|\mathbf{x} - R^\top \hat{\mathbf{c}}\|^2 \\
\text{subject to} \quad & \hat{C} = C^1 \times \cdots \times C^m \\
& R^\top R = I
\end{aligned}
$$

optimized product quantization

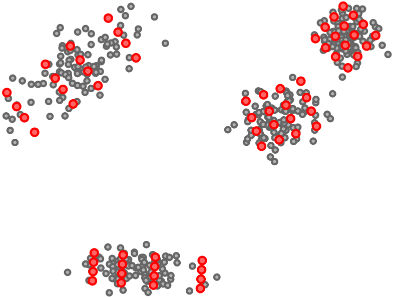
Parametric solution for $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \Sigma)$



- independence: PCA-align by diagonalizing Σ as $U\Lambda U^\top$
- balanced variance: permute Λ by π such that $\prod_i \lambda_i$ is constant in each subspace; $R \leftarrow UP_\pi^\top$
- find $\hat{C}$ by PQ on rotated data $\hat{X} = RX$

locally optimized product quantization

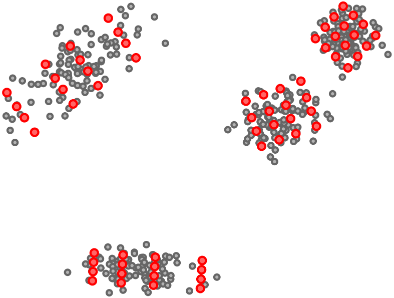
[Kalantidis & Avrithis, CVPR 2014]



- compute residuals $r(\mathbf{x}) = \mathbf{x} - q(\mathbf{x})$ on coarse quantizer q
- collect residuals $Z_{\mathbf{c}} = \{r(\mathbf{x}) : q(\mathbf{x}) = \mathbf{c}\}$ per cell
- train $(R_{\mathbf{c}}, q_{\mathbf{c}}) \leftarrow \mathrm{OPQ}(Z_{\mathbf{c}})$ per cell

locally optimized product quantization

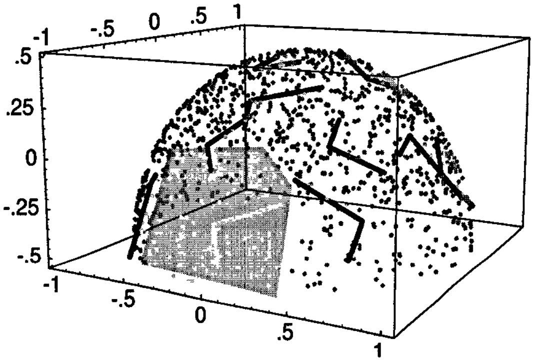
[Kalantidis & Avrithis, CVPR 2014]



- residual distributions closer to Gaussian assumption
- better captures the support of data distribution, like local PCA
  - multimodal (e.g. mixture) distributions
  - distributions on nonlinear manifolds

local principal component analysis

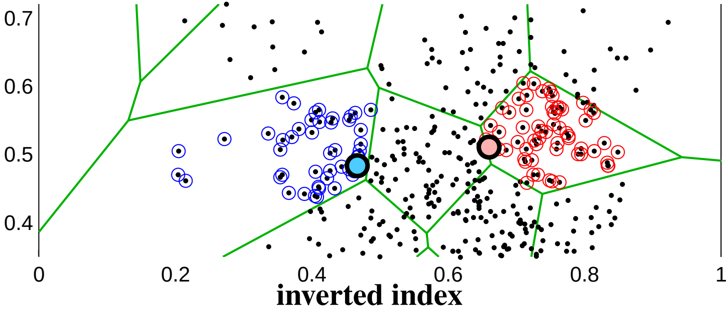
[Kambhatla & Leen 1997]



but, we are not doing dimensionality reduction!

inverted multi-index

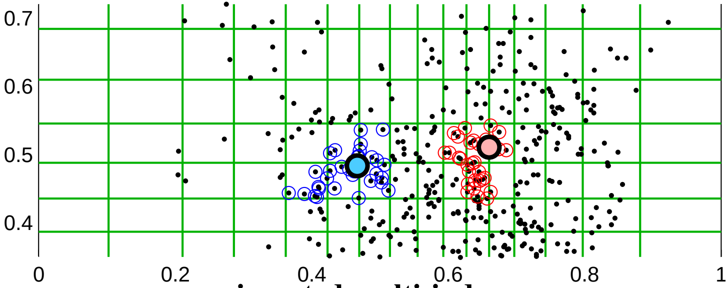
[Babenko & Lempitsky 2012]



- train codebook C from dataset $\{\mathbf{x}_n\}$
- this codebook provides a coarse partition of the space

inverted multi-index

[Babenko & Lempitsky 2012]



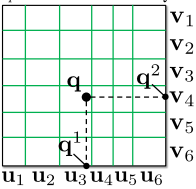
inverted multi-index

- decompose vectors as $\mathbf{x} = (\mathbf{x}^1, \mathbf{x}^2)$
- train codebooks C^1, C^2 from datasets $\{\mathbf{x}^1_n\}, \{\mathbf{x}^2_n\}$
- induced codebook $C^1 \times C^2$ gives a finer partition
- given query $\mathbf{q}$, visit cells $(\mathbf{c}^1, \mathbf{c}^2) \in C^1 \times C^2$ in ascending order of distance to $\mathbf{q}$, by first computing distances to $\mathbf{q}^1, \mathbf{q}^2$

inverted multi-index

multi-sequence algorithm

space subdivision via PQ



product quantization →

$\mathbf{q}^1$ vs. $\mathcal{U}$

i	$\mathbf{u}_{\alpha(i)}$	r
1	$\mathbf{u_3}$	0.5
2	$\mathbf{u_4}$	0.7
3	$\mathbf{u_5}$	4
4	$\mathbf{u_2}$	6
5	$\mathbf{u_1}$	8
6	$\mathbf{u_6}$	9

$\mathbf{q}^2$ vs. $\mathcal{V}$

j	$\mathbf{v}_{\beta(j)}$	s
1	$\mathbf{v_4}$	0.1
2	$\mathbf{v_3}$	2
3	$\mathbf{v_5}$	3
4	$\mathbf{v_2}$	6
5	$\mathbf{v_6}$	7
6	$\mathbf{v_1}$	11

multi-sequence algorithm →

$[\mathbf{u}_{\alpha(i)}\ \mathbf{v}_{\beta(j)}]$	(i,j)	$r(i)+s(j)$
$[\mathbf{u_3}\ \mathbf{v_4}]$	(1,1)	0.6 (0.5+0.1)
$[\mathbf{u_4}\ \mathbf{v_4}]$	(2,1)	0.8 (0.7+0.1)
$[\mathbf{u_3}\ \mathbf{v_3}]$	(1,2)	2.5 (0.5+2)
$[\mathbf{u_4}\ \mathbf{v_3}]$	(2,2)	2.7 (0.7+2)
$[\mathbf{u_3}\ \mathbf{v_5}]$	(1,3)	3.5 (0.5+3)
$[\mathbf{u_4}\ \mathbf{v_5}]$	(2,3)	3.7 (0.7+3)
$[\mathbf{u_5}\ \mathbf{v_4}]$	(3,1)	4.1 (4+0.1)
$[\mathbf{u_5}\ \mathbf{v_3}]$	(3,2)	6 (4+2)
$[\mathbf{u_3}\ \mathbf{v_2}]$	(1,4)	6.5 (0.5+6)
…		

	1	2	3	4	5	6	
1	0.6	0.8	4.1	6.1	8.1	9.1	
2	2.5	2.7	6	8	10	11	
3	3.5	3.7	7	9	11	12	
4	6.5	6.7	10	12	14	15	
5	7.5	7.7	11	13	15	16	
6	11.5	11.7	15	17	19	20	
	$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_2$	$\mathbf{u}_1$	$\mathbf{u}_6$	

OUTPUT:

1	2	3	4	5	6
0.6	0.8	4.1	6.1	8.1	9.1
2.5	2.7	6	8	10	11
3.5	3.7	7	9	11	12
6.5	6.7	10	12	14	15
7.5	7.7	11	13	15	16
11.5	11.7	15	17	19	20
$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_2$	$\mathbf{u}_1$	$\mathbf{u}_6$

$(1,1) \to \mathbf{W}_{3,4}$

1	2	3	4	5	6
0.6	0.8	4.1	6.1	8.1	9.1
2.5	2.7	6	8	10	11
3.5	3.7	7	9	11	12
6.5	6.7	10	12	14	15
7.5	7.7	11	13	15	16
11.5	11.7	15	17	19	20
$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_2$	$\mathbf{u}_1$	$\mathbf{u}_6$

$(2,1) \to \mathbf{W}_{4,4}$

1	2	3	4	5	6
0.6	0.8	4.1	6.1	8.1	9.1
2.5	2.7	6	8	10	11
3.5	3.7	7	9	11	12
6.5	6.7	10	12	14	15
7.5	7.7	11	13	15	16
11.5	11.7	15	17	19	20
$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_2$	$\mathbf{u}_1$	$\mathbf{u}_6$

$(1,2) \to \mathbf{W}_{3,3}$

1	2	3	4	5	6
0.6	0.8	4.1	6.1	8.1	9.1
2.5	2.7	6	8	10	11
3.5	3.7	7	9	11	12
6.5	6.7	10	12	14	15
7.5	7.7	11	13	15	16
11.5	11.7	15	17	19	20
$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_2$	$\mathbf{u}_1$	$\mathbf{u}_6$

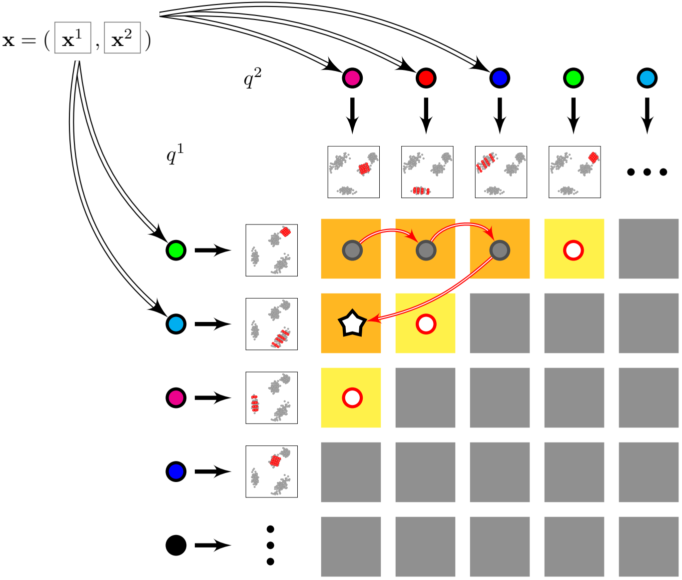
$(2,2) \to \mathbf{W}_{4,3}$

1	2	3	4	5	6	
0.6	0.8	4.1	6.1	8.1	9.1	$\mathbf{v}_4$
2.5	2.7	6	8	10	11	$\mathbf{v}_3$
3.5	3.7	7	9	11	12	$\mathbf{v}_5$
6.5	6.7	10	12	14	15	$\mathbf{v}_2$
7.5	7.7	11	13	15	16	$\mathbf{v}_6$
11.5	11.7	15	17	19	20	$\mathbf{v}_1$
$\mathbf{u}_3$	$\mathbf{u}_4$	$\mathbf{u}_5$	$\mathbf{u}_2$	$\mathbf{u}_1$	$\mathbf{u}_6$	

$(1,3) \to \mathbf{W}_{3,5}$

Multi-LOPQ

[Kalantidis & Avrithis, CVPR 2014]



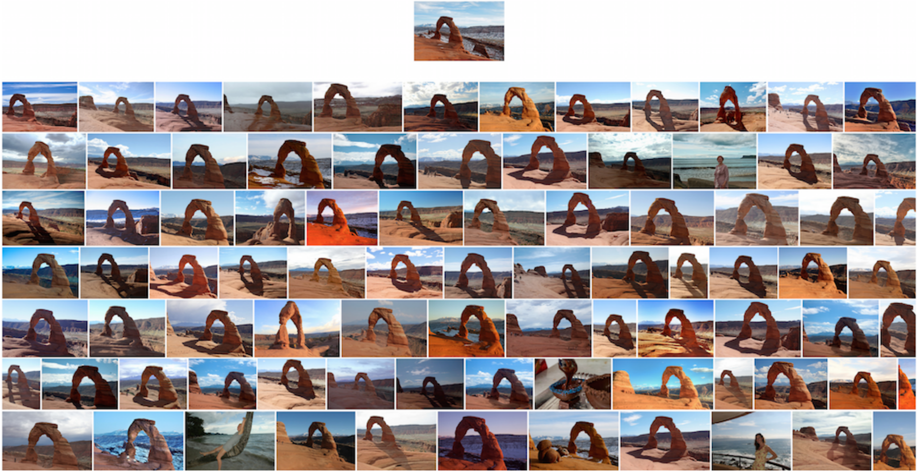
comparison to state of the art

on SIFT1B, 128-bit codes

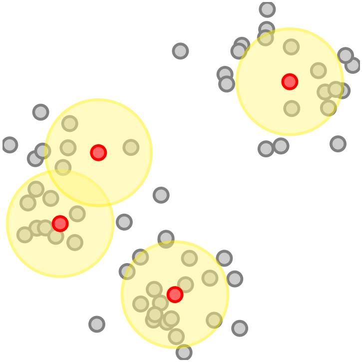
T	Method	$R = 1$	10	100
20K	IVFADC+R [Jégou *et al.* '11]	0.262	0.701	0.962
	LOPQ+R [Kalantidis & Avrithis '14]	0.350	0.820	0.978
10K	Multi-D-ADC [Babenko & Lempitsky '12]	0.304	0.665	0.740
	OMulti-D-OADC [Ge *et al.* '13]	0.345	0.725	0.794
	Multi-LOPQ [Kalantidis & Avrithis '14]	0.430	0.761	0.782
30K	Multi-D-ADC [Babenko & Lempitsky '12]	0.328	0.757	0.885
	OMulti-D-OADC [Ge *et al.* '13]	0.366	0.807	0.913
	Multi-LOPQ [Kalantidis & Avrithis '14]	0.463	0.865	0.905
100K	Multi-D-ADC [Babenko & Lempitsky '12]	0.334	0.793	0.959
	OMulti-D-OADC [Ge *et al.* '13]	0.373	0.841	0.973
	Multi-LOPQ [Kalantidis & Avrithis '14]	0.476	0.919	0.973

image query on Flickr 100M

deep learned features, 4k → 128 dimensions

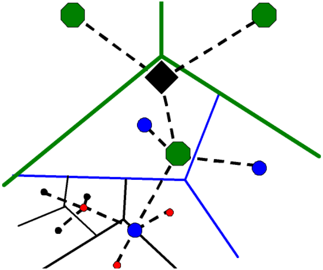


credit: Y. Kalantidis



clustering

ANN search - clustering connection

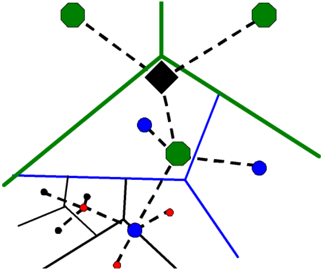


hierarchical k-means

[Nister & Stewenius 2006]

use k-means tree for ANN search

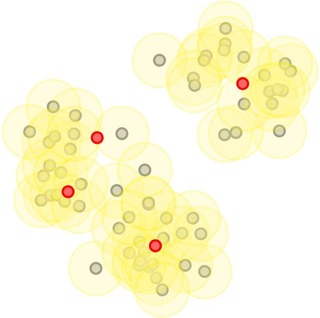
ANN search - clustering connection



hierarchical k-means

[Nister & Stewenius 2006]

use k-means tree for ANN search

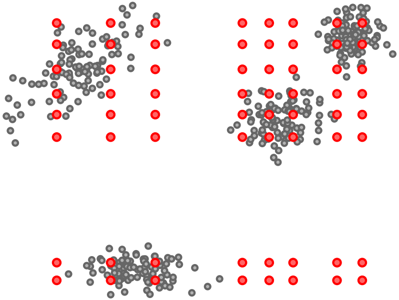


approximate k-means

[Philbin *et al.* 2007]

use ANN search to accelerate assignment step

ANN search - clustering connection

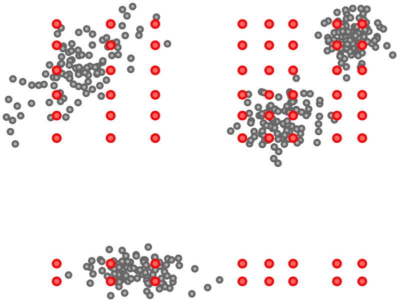


product quantization

[Jégou *et al.* 2010]

use k-means on subspaces to accelerate ANN search

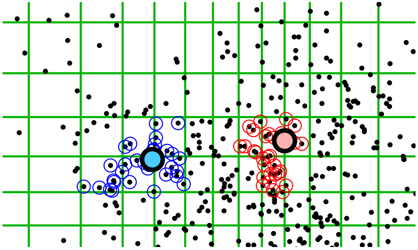
ANN search - clustering connection



product quantization

[Jégou *et al.* 2010]

use k-means on subspaces to accelerate ANN search



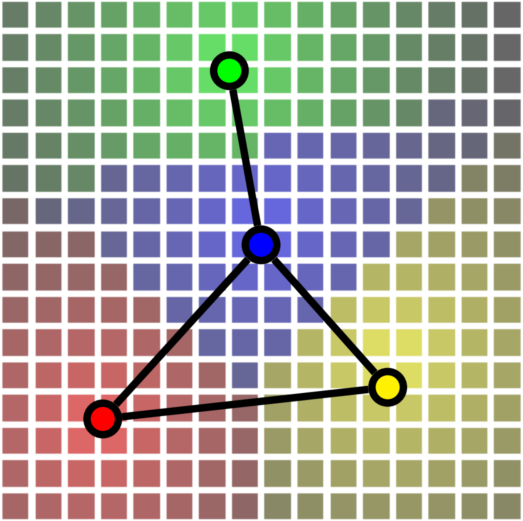
inverted multi-index

[Babenko & Lempitsky 2012]

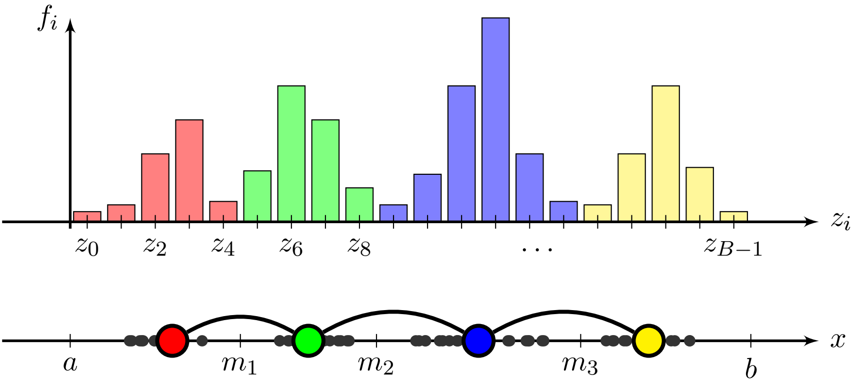
exhaustively search on subspaces before searching on entire space

dimensionality-recursive vector quantization

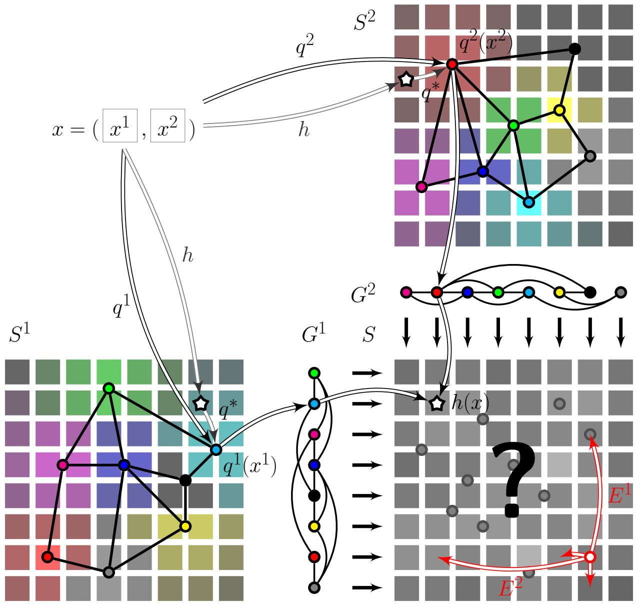
[Avrithis, ICCV 2013]



DRVQ base case: $d = 1$



DRVQ recursion: $d \to 2d$



DRVQ: vector quantization

k	16k	8k	4k	2k	1k	512
approximate (μs)	0.95	0.83	0.80	0.73	0.80	0.90
exact (ms)	1.19	0.79	0.51	0.26	0.21	0.11

averaged over the $n = 75$k SIFT descriptors of the 55 cropped query images of *Oxford 5k*

DRVQ: clustering

k	$\log k_p$ $(d = 2^p)$						time (m)
	1	2	4	8	16	32	
16k	6	7	8	9	11	14	129.96
8k	6	7	8	9	11	13	119.43
4k	6	7	8	9	10	12	20.07
2k	5	6	7	8	9	11	2.792
1k	5	6	7	8	9	10	2.608
512	4	5	6	7	8	9	0.866
4k	approximate k-means						504.2

4 codebooks at $d = 32$ dimensions each on $n = 12.5$M 128-dimensional SIFT descriptors of *Oxford 5k*

DRVQ: clustering

k	$\log k_p$ $(d = 2^p)$						time (m)
	1	2	4	8	16	32	
16k	6	7	8	9	11	14	129.96
8k	6	7	8	9	11	13	119.43
4k	6	7	8	9	10	12	20.07
2k	5	6	7	8	9	11	2.792
1k	5	6	7	8	9	10	2.608
512	4	5	6	7	8	9	0.866
4k	approximate k-means						504.2

4 codebooks at $d = 32$ dimensions each on $n = 12.5$M 128-dimensional SIFT descriptors of *Oxford 5k*

inverted-quantized k-means

[Avrithis et al. ongoing]

clustering of 100M images in less than one hour on a single core

mining, location & instance recognition

http://viral.image.ntua.gr

query

result

Estimated Location Similar Image, Incorrectly geo-tagged Unavailable

Suggested tags: Buxton Memorial Fountain, Victoria Tower Gardens, London
Frequent user tags: Victoria Tower Gardens, Buxton Memorial Fountain, Winchester Palace, Architecture, Victorian gothic

Similar Images

Similarity: 0.619
Details Original

Similarity: 0.491
Details Original

Similarity: 0.397
Details Original

Similarity: 0.385
Details Original

suggested tags

Suggested tags: Buxton Memorial Fountain, Victoria Tower Gardens, London
Frequent user tags: Victoria Tower Gardens, Buxton Memorial Fountain, Winchester Palace, Architecture, Victorian gothic

related wikipedia articles

New features Log in / create account

Article Discussion Read Edit View history Search

Main page
Contents
Featured content
Current events
Random article

Interaction
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Donate to Wikipedia
Help

Toolbox
What links here
Related changes
Upload file
Special pages
Permanent link
Cite this page

Print/export

Buxton Memorial Fountain

From Wikipedia, the free encyclopedia

The **Buxton Memorial Fountain** is a memorial and drinking fountain in London, the United Kingdom, that commemorates the emancipation of slaves in the British Empire in 1834.

It was commissioned by Charles Buxton MP, and was dedicated to his father Thomas Fowell Buxton along with William Wilberforce, Thomas Clarkson, Thomas Babington Macaulay, Henry Brougham and Stephen Lushington, all of whom were involved in the abolition. It was designed by Gothic architect Samuel Sanders Teulon (1812–1873) in 1865 coincidently with the passing of the Thirteenth Amendment to the United States Constitution, which effectively ended the western slave-trade.[1]

It was originally constructed in Parliament Square, erected at a cost of £1,200. As part of the postwar redesign of the square it was removed in 1949 and not reinstated in its present position in Victoria Tower Gardens until 1957.[2] There were eight decorative figures of British rulers on it, but four were stolen in 1960 and four in 1971. They were replaced by fibreglass figures in 1980. By 2005 these were missing, and the fountain was no longer working. Between autumn 2006 and February 2007 restoration works were carried out. The restored fountain was unveiled on 27 March 2007 as part of the commemoration of the 200th anniversary of the act to abolish the slave trade [3].

A memorial plaque commemorating the 150th anniversary of the Anti-Slavery Society was added in 1989.

The Buxton Memorial Fountain, designed by Samuel Sanders Teulon, celebrating the emancipation of slaves in the British Empire in 1834, in Victoria Tower Gardens, Millbank,

Description [edit]

The base is octagonal, about twelve feet in diameter, having open arches on the eight sides, supported on clustered shafts of polished Devonshire marble around a large central shaft, with four massive granite basins. Surmounting the pinnacles at the angles of the octagon are eight figures of bronze, representing the different rulers of England; the Britons represented by Caractacus, the Romans by Constantine, the Danes by Canute, the Saxons by Alfred, the Normans by William the Conqueror, and so on, ending with Queen Victoria. The fountain bears an inscription to the effect that it is "intended as a memorial of those members of Parliament who, with Mr. Wilberforce, advocated the abolition of the British slave-trade, achieved in 1807; and of those members of Parliament who, with Sir T.

related wikipedia articles

New features Log in / create account

Article | Discussion | Read | Edit | View history | Search

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate

Interaction
About Wikipedia
Community portal
Recent changes
Contact Wikipedia
Help

Toolbox
What links here
Related changes
Upload file
Special pages
Permanent link
Cite this page

Print/export

Victoria Tower Gardens

From Wikipedia, the free encyclopedia

Coordinates: 51°29′49.0″N 0°7′30.0″W

Victoria Tower Gardens is a public park along the north bank of the River Thames in London. As its name suggests, it is adjacent to the Victoria Tower, the south-western corner of the Palace of Westminster. The park, which extends southwards from the Palace to Lambeth Bridge, sandwiched between Millbank and the river, also forms part of the Thames Embankment.

Victoria Tower Gardens, 2005, with the Buxton Memorial Fountain at the front and the Palace of Westminster in the background

Contents [hide]

Features [edit]

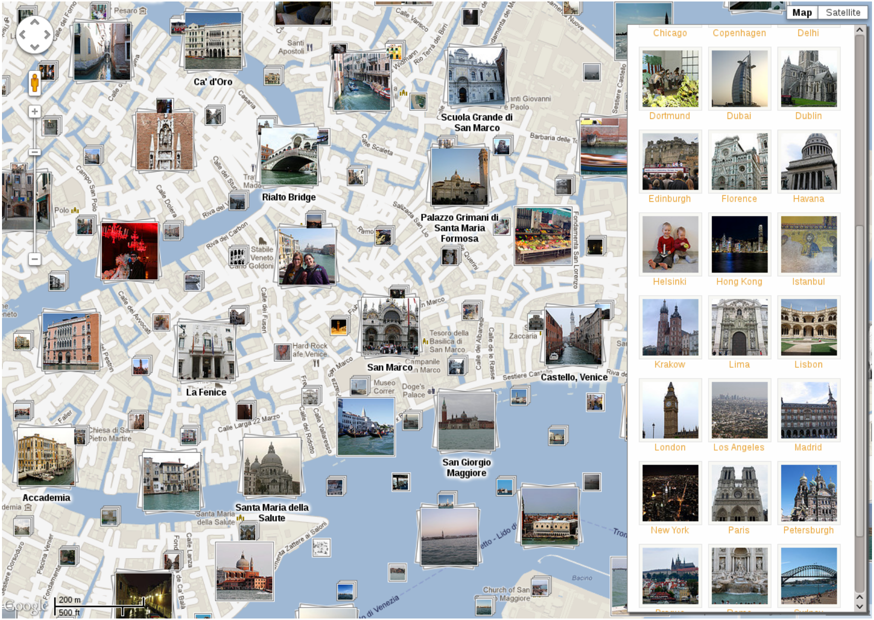
The park features:

- A reproduction of the sculpture *The Burghers of Calais* by Auguste Rodin, purchased by the British Government in 1911 and positioned in the Gardens in 1915.
- A 1930 statue of the suffragette Emmeline Pankhurst, by A.G. Walker.
- The Buxton Memorial Fountain – originally constructed in Parliament Square, this was removed in 1940 and placed in its present position in 1957. It was commissioned by Charles Buxton MP to commemorate the emancipation of slaves in 1834, dedicated to his father Thomas Fowell Buxton, and designed by Gothic architect Samuel Sanders Teulon (1812–1873) in 1865.
- A stone wall with two modern-style goats with kids – situated at the southern end of the Gardens.

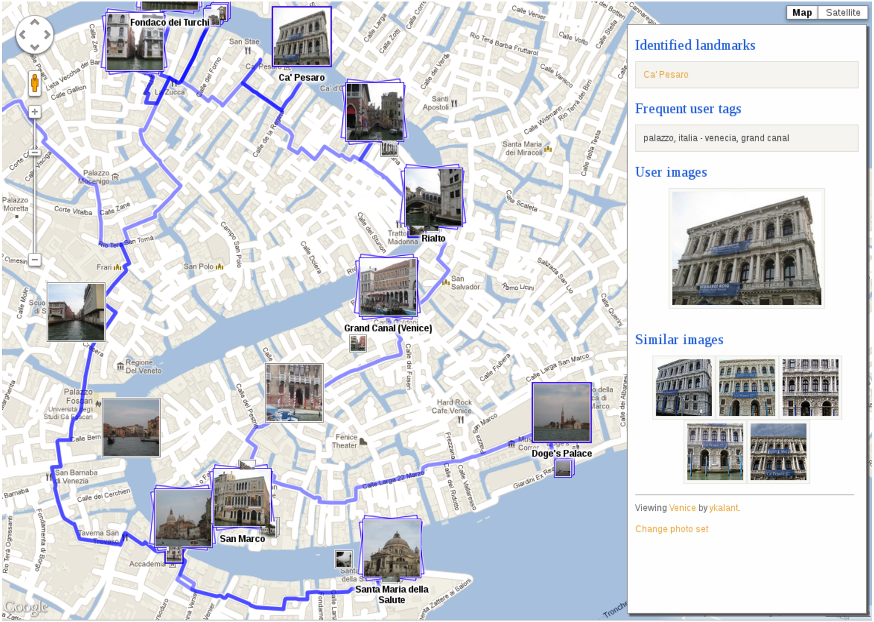
Transport [edit]

VIRaL explore

VIRaL explore



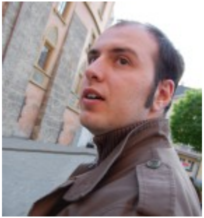
VIRaL routes



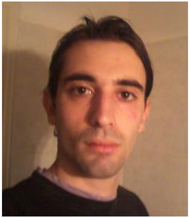
credits

Yannis Kalantidis

Giorgos Tolias



Christos Varitimidis



Kimon Kontosis

Marios Phinikettos

Kostas Rapantzikos

http://image.ntua.gr/iva/research/

thank you!