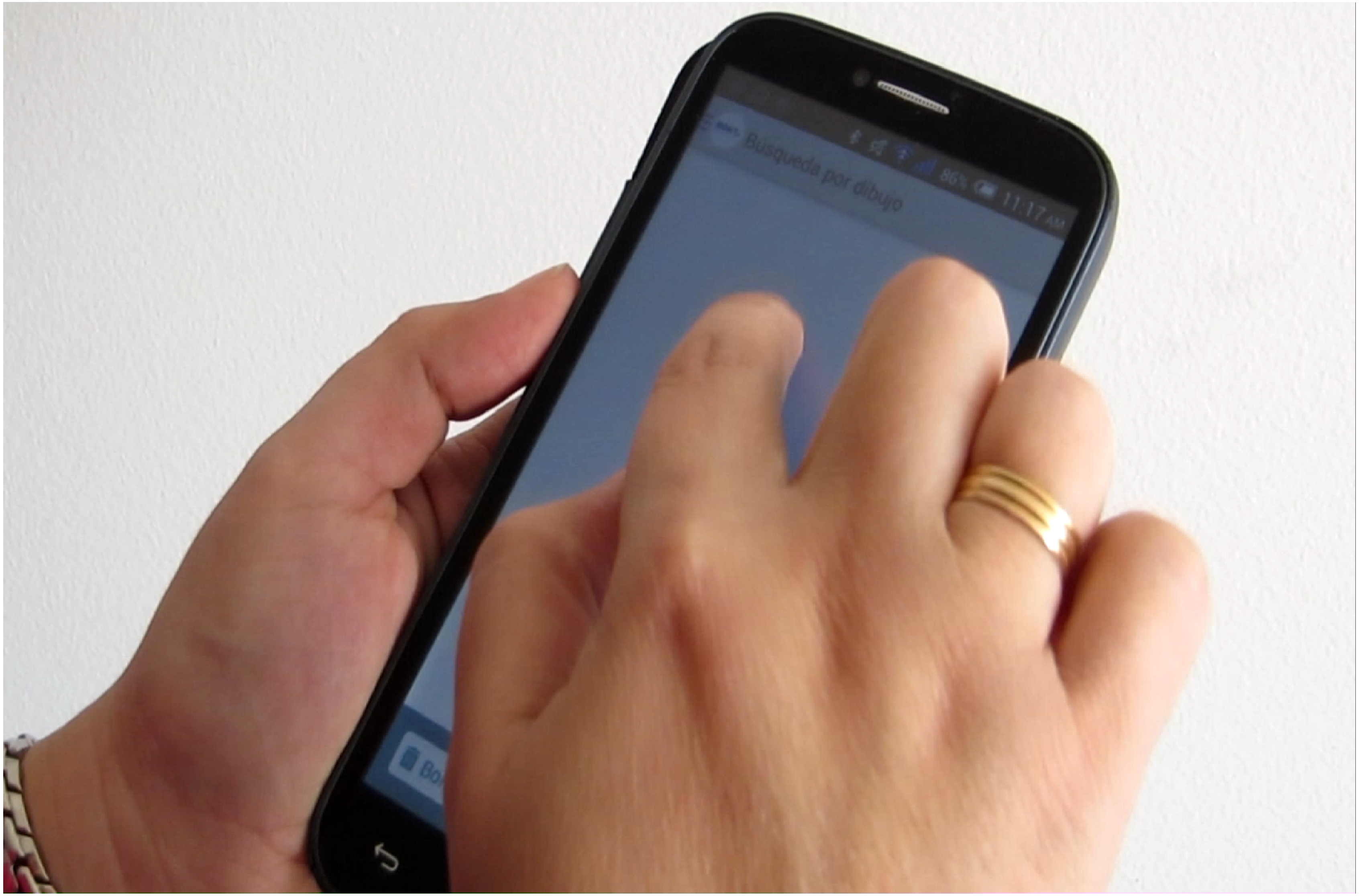
Sketch based Image Retrieval using Learned KeyShapes (LKS)

José M. Saavedra (PhD.)

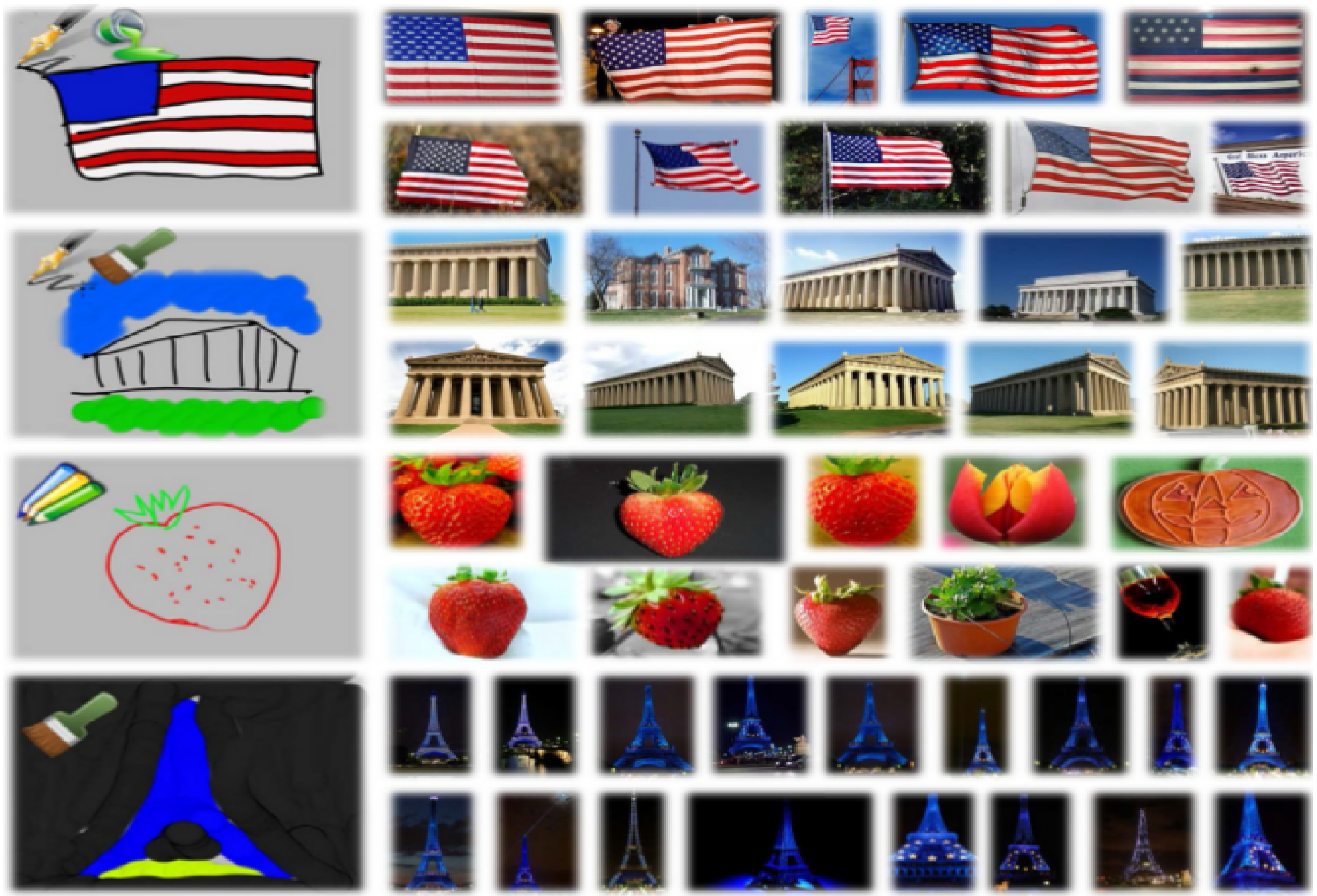
orand
Ingeniería – Innovación - Software

Jose M. Saavedra, Juan Manuel Barrios. Sketch based Image Retrieval using Learned KeyShapes (LKS) . 26th British Machine Vision Conference (BMVC), Swansea, UK, 2015

Sketch Based Image Retrieval



Sketch Based Image Retrieval



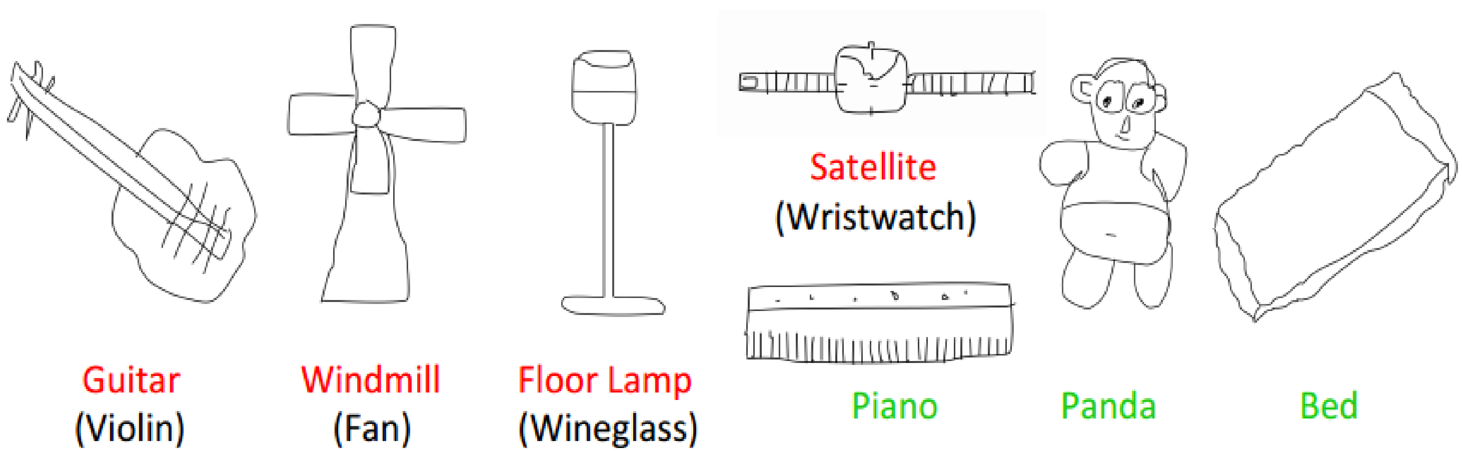
Xinghai Sun, Changhu Wang, Avneesh Sud, Chao Xu, and Lei Zhang. 2013. MagicBrush: image search by color sketch. In Proceedings of the 21st ACM international conference on Multimedia (MM '13). ACM, New York, NY, USA

Sketch Based Image Retrieval

A *sketch* is a simple hand-made drawing composed of just a sequence of strokes. A *sketch* may lack of color and texture.

Sketch Based Image Retrieval

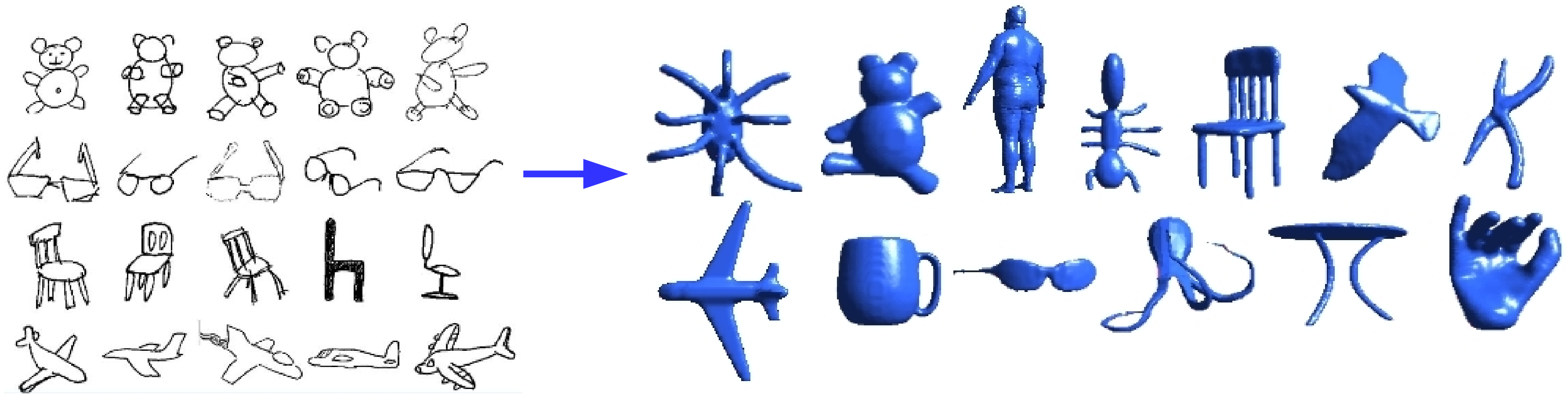
Related applications:



Sketch Classification

Sketch Based Image Retrieval

Related applications:

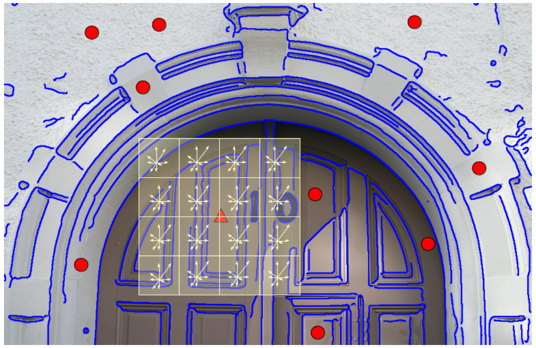


Sketch Classification

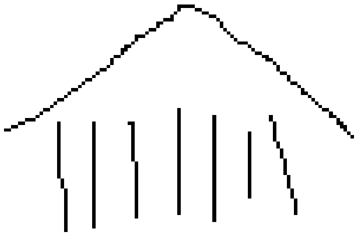
Sketch Based Image Retrieval

Related Work: Commonly using low-level features.

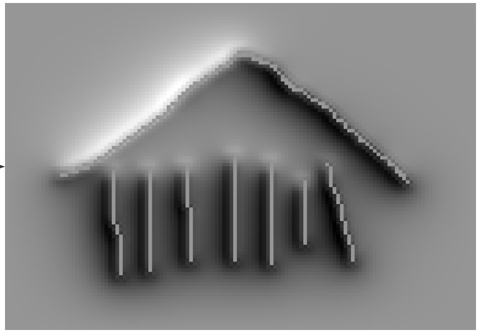
- Gradient Field [Hu et al. 2011, 2013]
- Spark Descriptor [Eitz et al. 2011]
- DoIGOH Descriptor [Eitz et al. 2011]



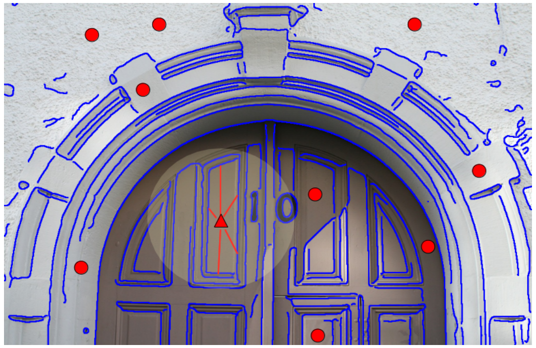
DoIGOH



Sketch



Gradient field

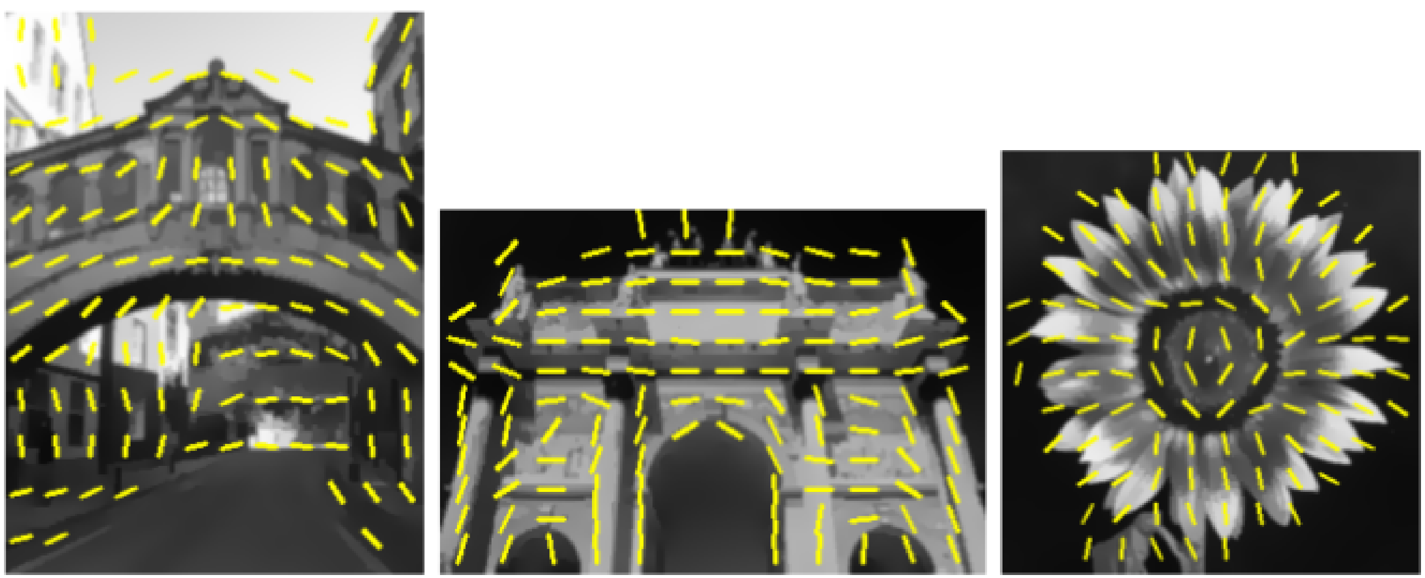


Spark

Sketch Based Image Retrieval

Related Work: Commonly using low-level features.

- HELO (Histogram of Edge Local Orientations) [Saavedra, 2010]
- SHELO (Soft HELO) [Saavedra, 2014]



Sketch Based Image Retrieval

Huebel and Wiesel cat experiment on Visual Perception

Sketch Based Image Retrieval

How can we apply MID-LEVEL features in the SBIR problem?

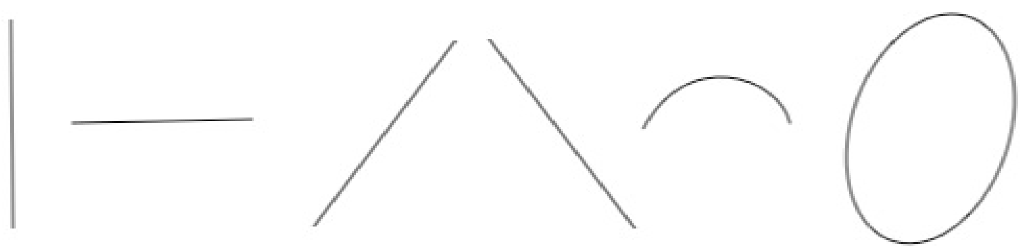
↓

KEYSHAPES

Sketch Based Image Retrieval

Related Work: Using keyshapes

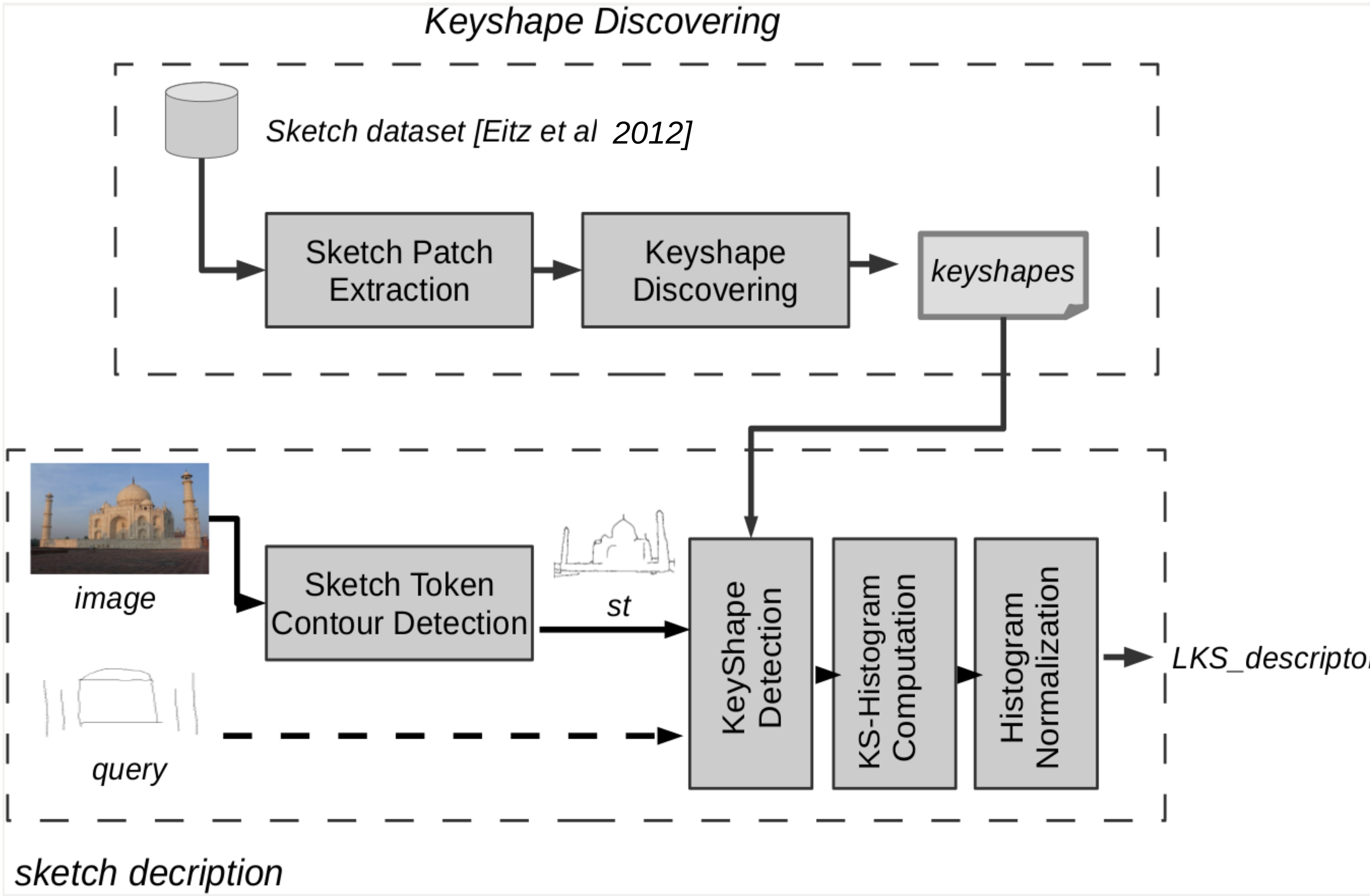
- Sketch-based image retrieval using keyshapes [Saavedra, 2011(14)]



A set of shapes manually defined

Sketch based Image Retrieval using Learned KeyShapes (LKS)

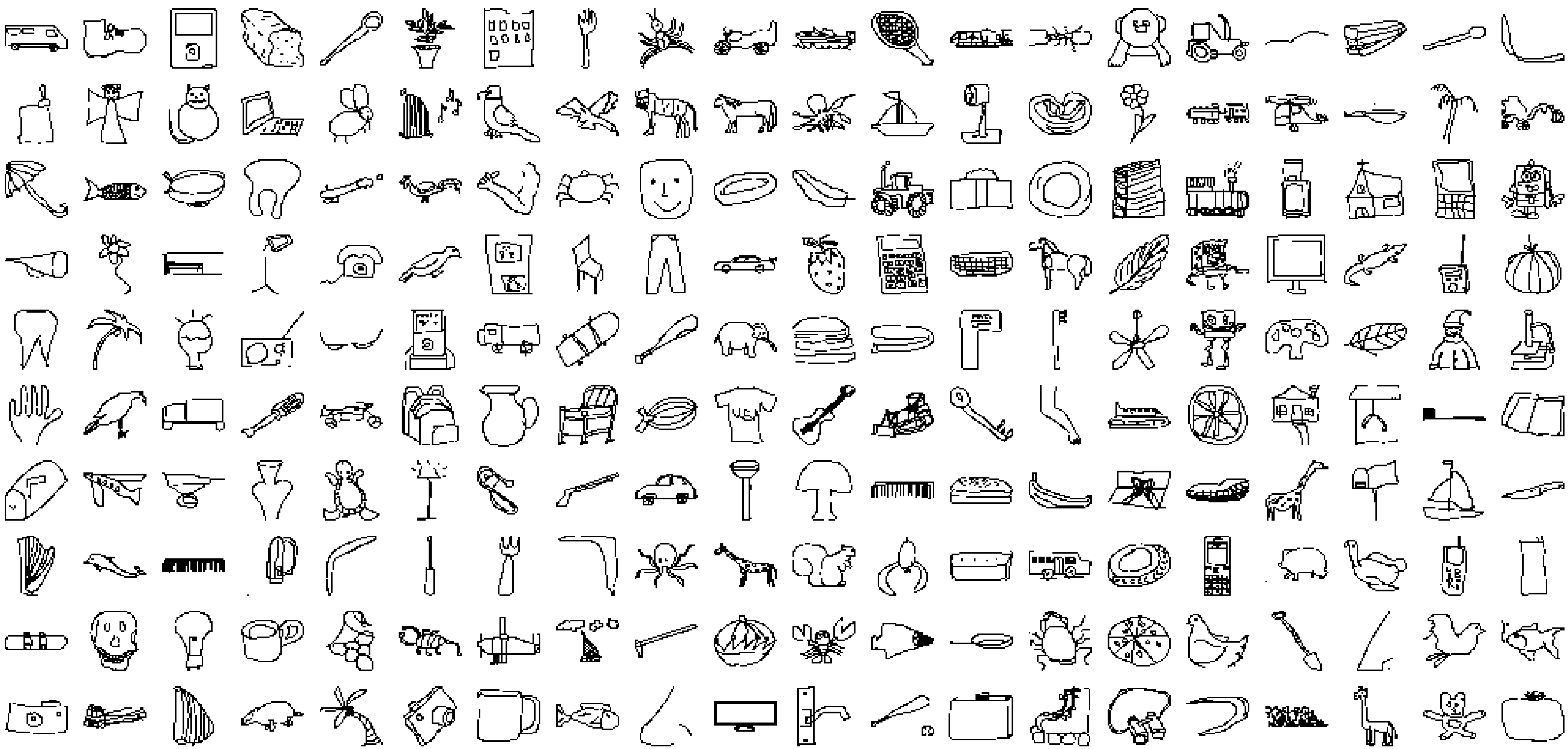
Scheme



Jose M. Saavedra, Juan Manuel Barrios. Sketch based Image Retrieval using Learned KeyShapes (LKS) . 26th British Machine Vision Conference (BMVC), Swansea, UK, 2015

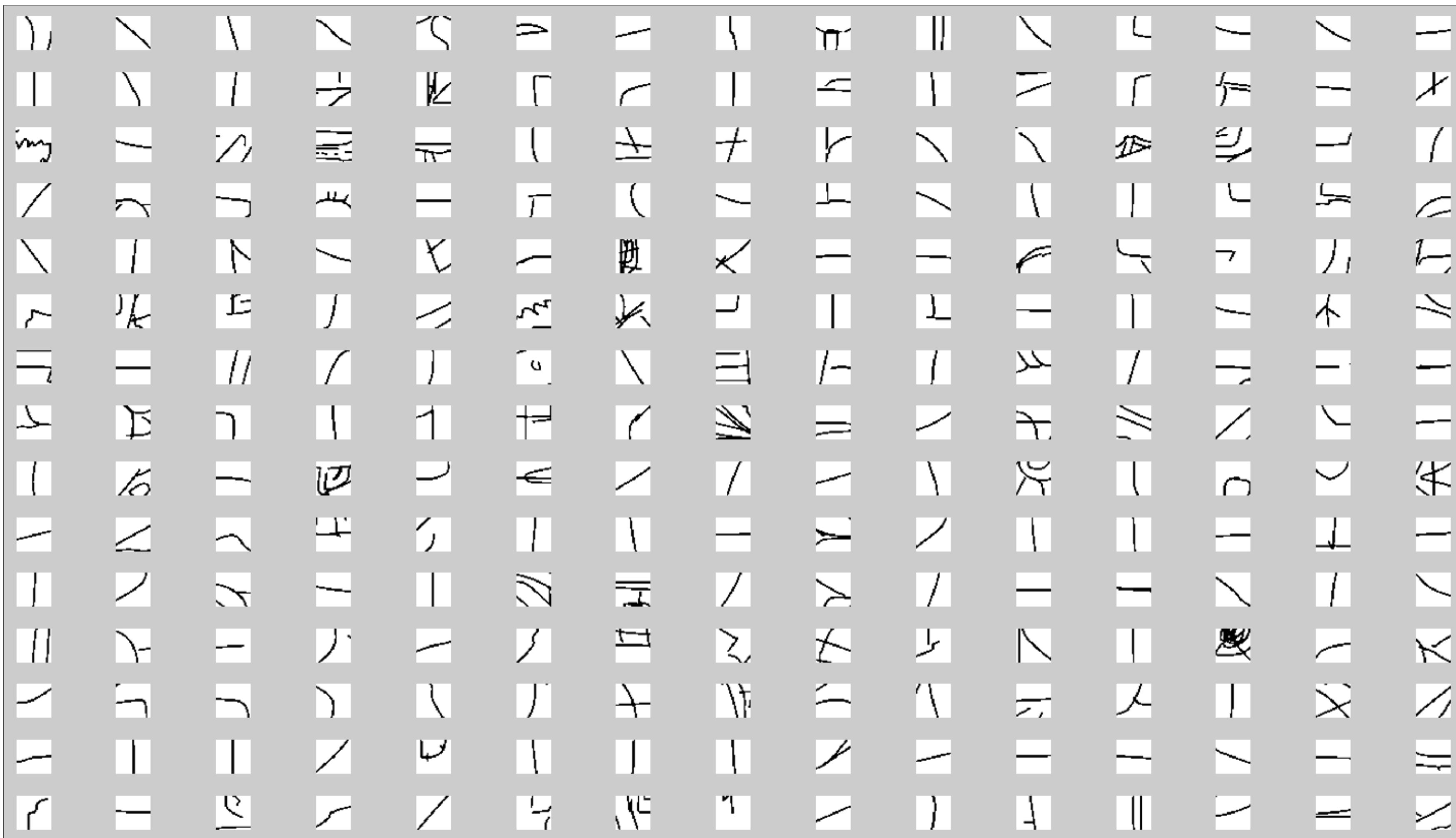
LKS-Keyshape Discovering

Training Dataset: 250 classes, 80 sketches/class, total = 20K sketches.



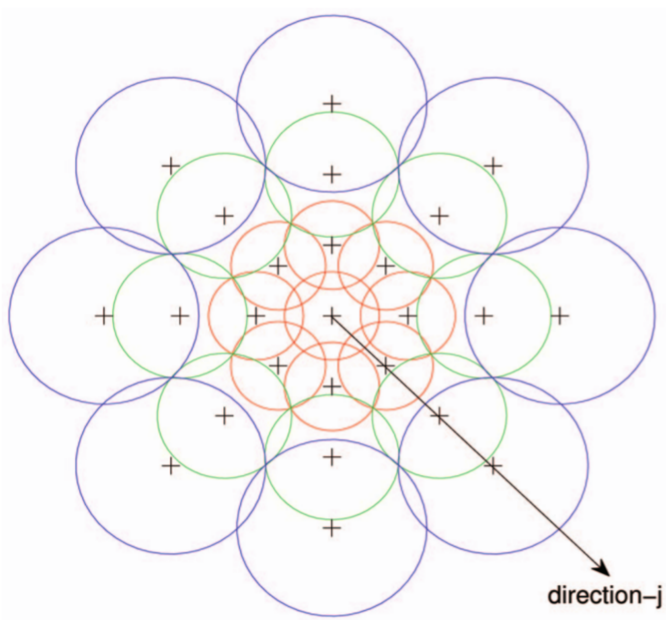
LKS-Keyshape Discovering

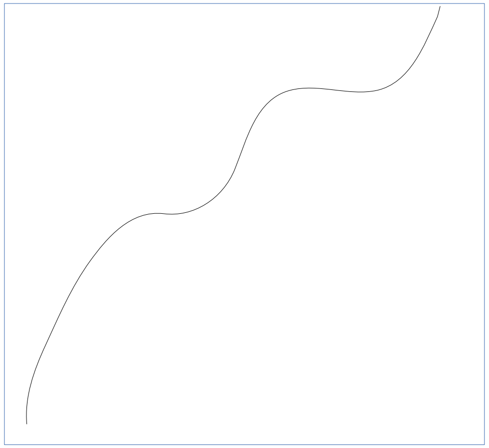
Patch Extraction



LKS-Keyshape Discovering
Patch Description

- ✔ Sketch patches are represented by DAISY descriptors



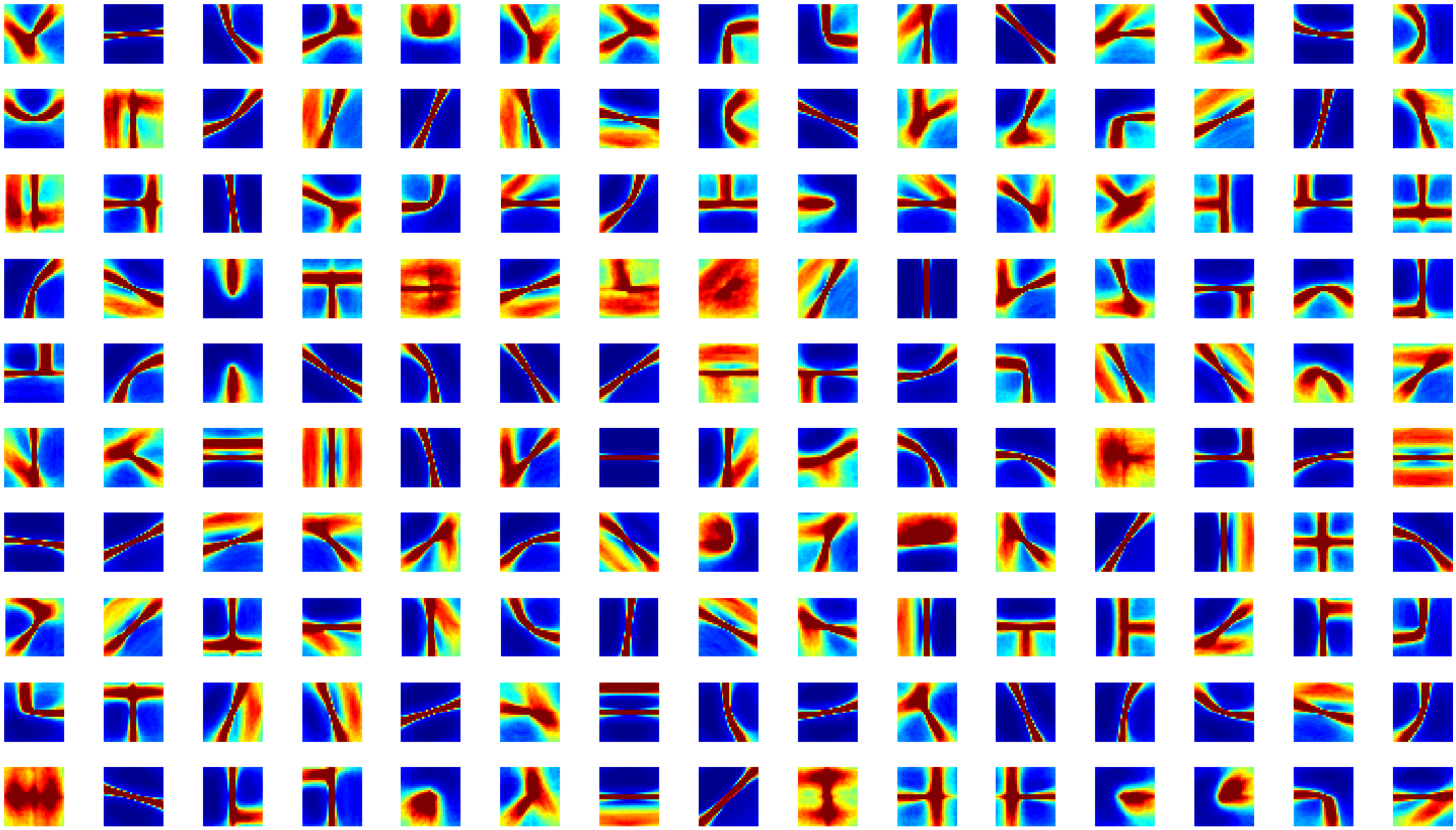


$$\begin{aligned}
&\mathcal{D}(u_0, v_0) = \\
&\quad [\widetilde{\mathbf{h}}^\top_{\Sigma_1}(u_0, v_0), \\
&\quad \widetilde{\mathbf{h}}^\top_{\Sigma_1}(\mathbf{l}_1(u_0, v_0, R_1)), \ldots, \widetilde{\mathbf{h}}^\top_{\Sigma_1}(\mathbf{l}_T(u_0, v_0, R_1)), \\
&\quad \widetilde{\mathbf{h}}^\top_{\Sigma_2}(\mathbf{l}_1(u_0, v_0, R_2)), \ldots, \widetilde{\mathbf{h}}^\top_{\Sigma_2}(\mathbf{l}_T(u_0, v_0, R_2)), \\
&\quad \ldots \\
&\quad \widetilde{\mathbf{h}}^\top_{\Sigma_Q}(\mathbf{l}_1(u_0, v_0, R_Q)), \ldots, \widetilde{\mathbf{h}}^\top_{\Sigma_Q}(\mathbf{l}_T(u_0, v_0, R_Q))]^\top
\end{aligned}$$

Tola, Engin; Lepetit, V.; Fua, P., "DAISY: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.32, no.5, pp.815,830, May 2010

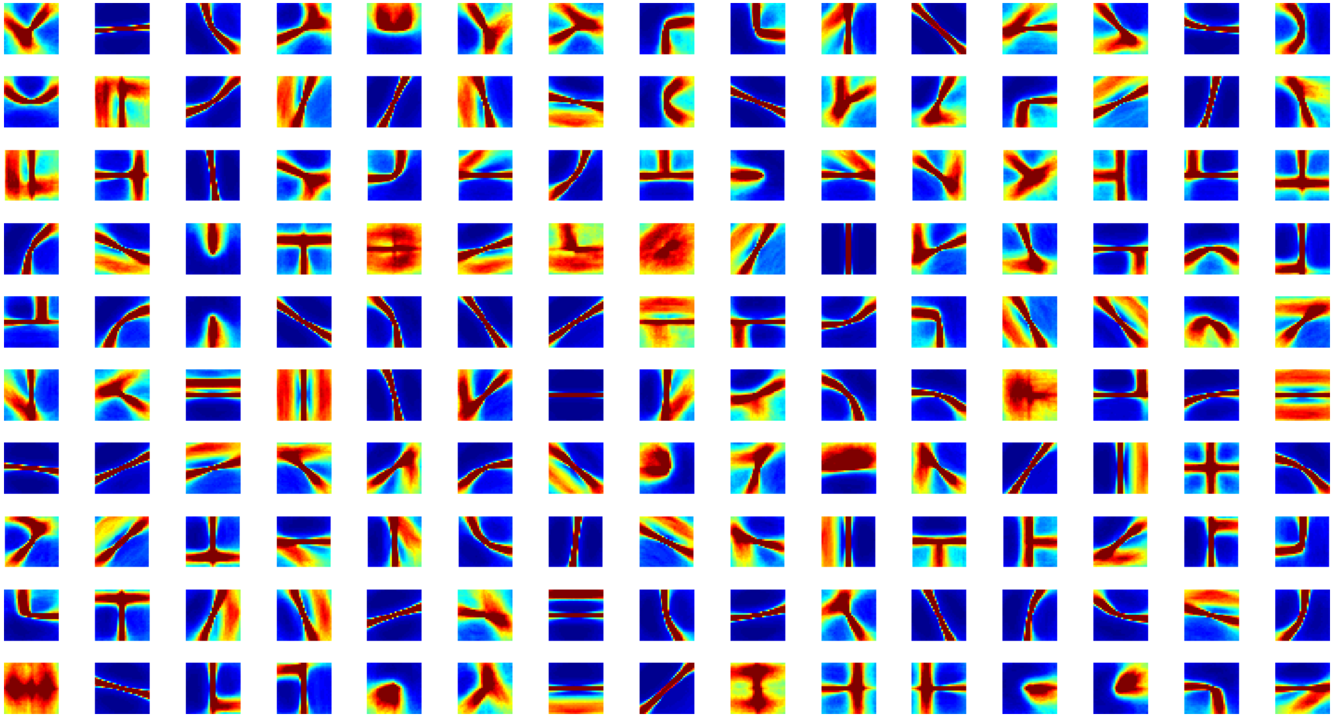
LKS-Keyshape Discovering

Clustering (on 1M patches)



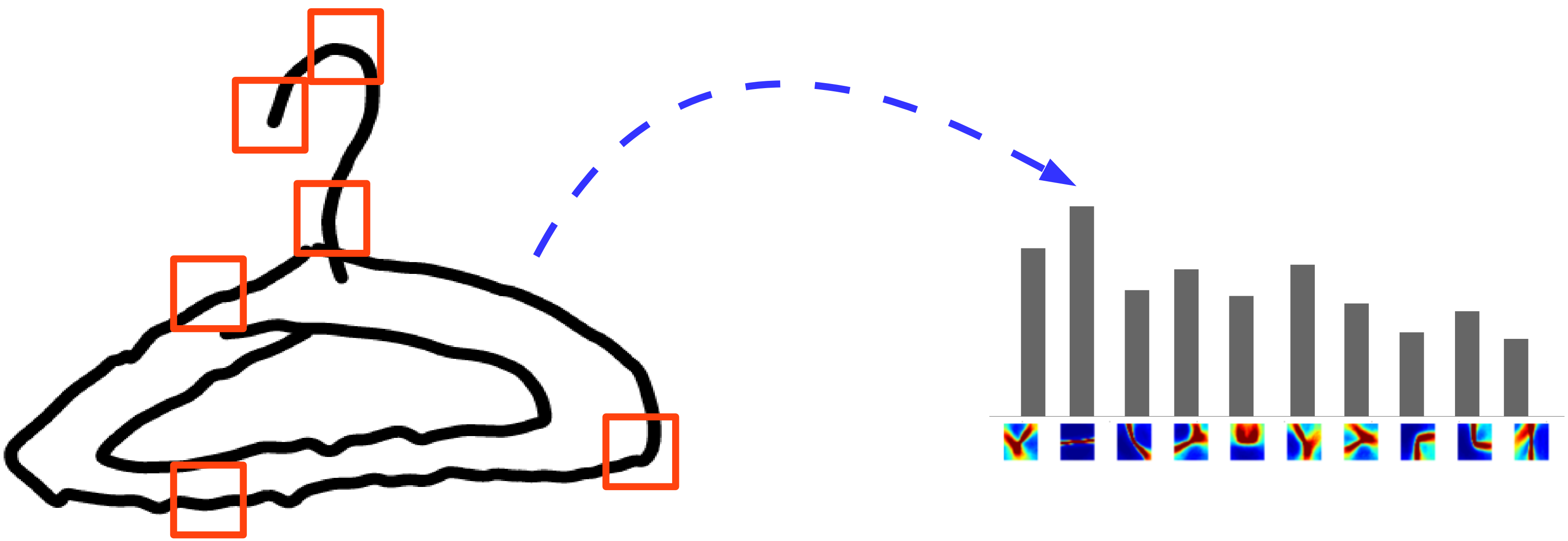
LKS-Keyshape Discovering

Clustering



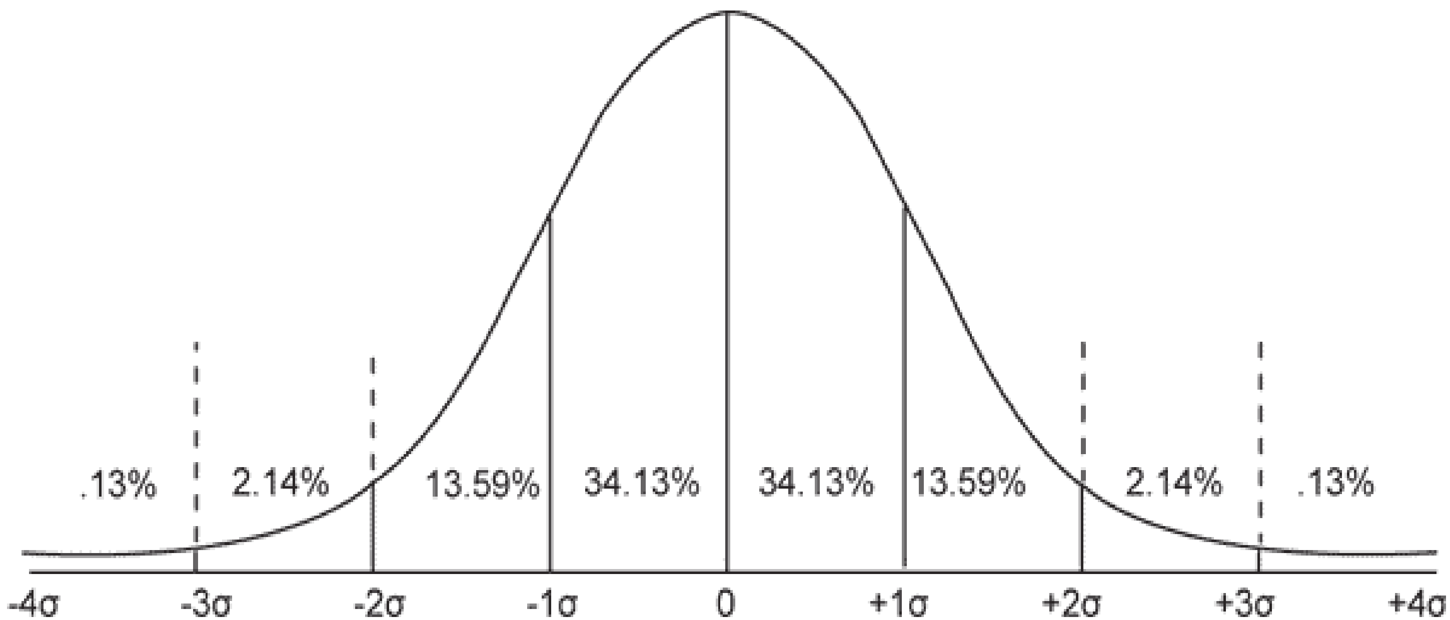
$$Clusters_K = \{(\mu_1, r_1), \cdots, (\mu_K, r_K)\}$$

Sketch Representation



$$KS_P^i = \{(p_1^i, \delta_1^i), \cdots (p_P^i, \delta_P^i)\}$$

Voting Process

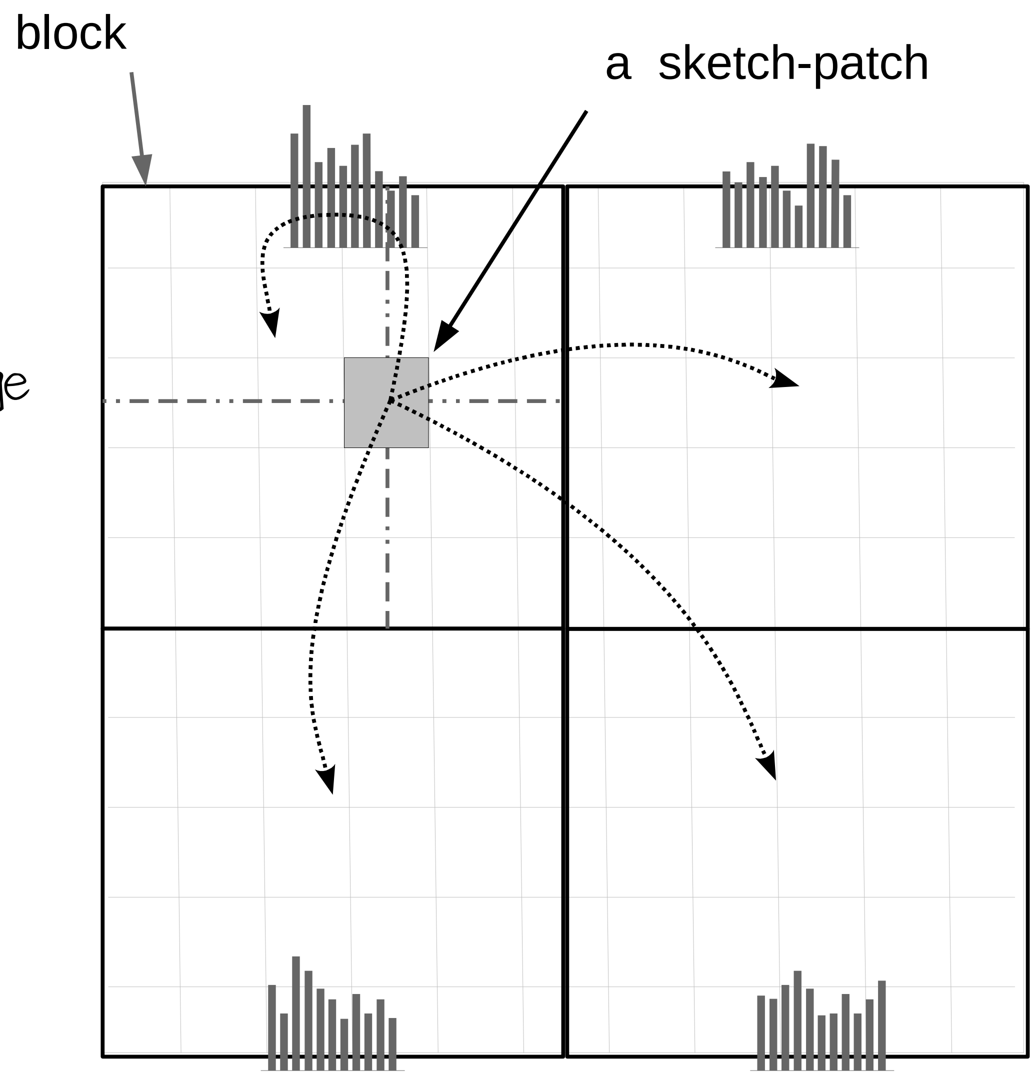


$$G(S_i, p_j^i) = \frac{1}{\sigma_{p_j^i}\sqrt{2\pi}} \exp\left(-0.5\frac{{\delta_j^i}^2}{{\sigma_{p_j^i}}^2}\right)$$

$$vote_j^i = \frac{G(S_i, p_j^i)}{\sum_{k=1}^{P} G(S_i, p_k^i)}$$

$$h_{LKS}(p_j^i) += vote_j^i$$

Spatial Division



Divide the image into BxB blocks and compute a ks-histogram for each block

bi-interpolation

Experimental Evaluation

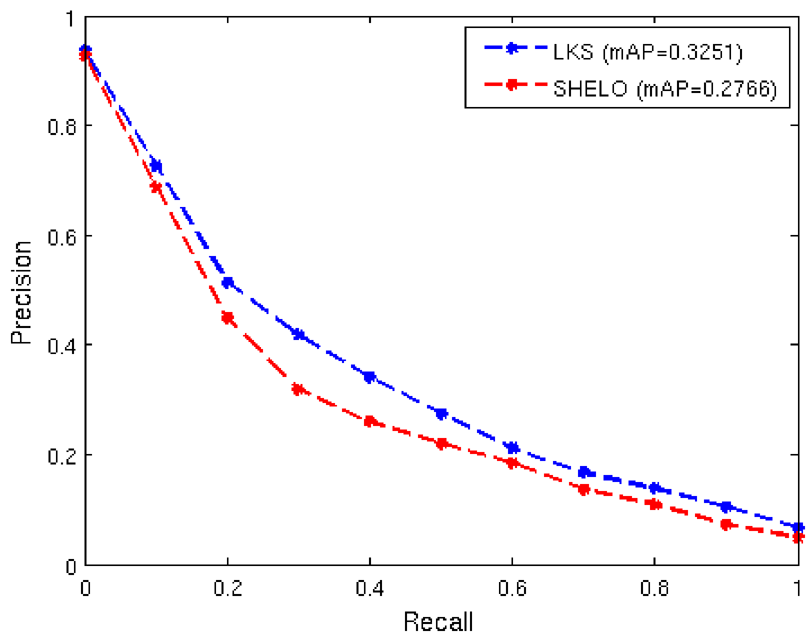
- **Using two different datasets**
  Saavedra (1326 images, 53 queries)
  Flickr 15K (14660 images, 330 queries)
- **Evaluation Metrics**
  MAP: Mean Average Precision
  Recall-Precision Graphic

Experimental Evaluation

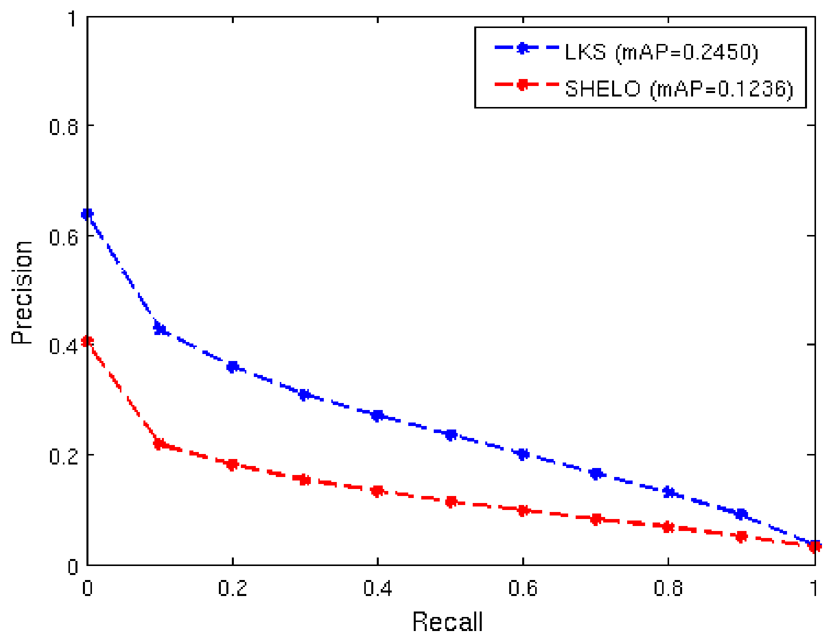
	HOG	GF-HOG[15]	SHELO[25]	**LKS**	gain
Saavedra's	0.2355	*unreported*	0.2766	**0.3251**	17.5%
Flickr15K	0.0771	0.1222	0.1236	**0.2450**	98.2%

mAP : Mean Average Precision

Experimental Evaluation



On Saavedra's dataset



On Flickr dataset

Experimental Evaluation

	K=50	K=100	K=150	K=200	K=300
B=3, P=10	0.2802	0.2901	0.2911	0.2908	0.2949
B=4, P=10	0.2925	0.3186	**0.3251**	0.3248	0.3180
B=3, P=15	0.2773	0.2888	0.2900	0.2911	0.2916
B=4, P=15	0.2828	0.3076	0.3210	0.3241	0.3250
B=3, P=20	0.2783	0.2863	0.2886	0.2878	0.2893
B=4, P=20	0.2789	0.3019	0.3144	0.3191	0.3220

mAP for a sample of different parameters in the Saavedra's dataset.

Experimental Evaluation

	K=50	K=100	K=150	K=200	K=300
B=3, P=10	0.2246	0.2407	**0.2450**	0.2431	0.2381
B=4, P=10	0.2208	0.2370	0.2351	0.2272	0.2104
B=3, P=15	0.2124	0.2348	0.2437	0.2452	0.2472
B=4, P=15	0.2056	0.2330	0.2390	0.2375	0.2300
B=3, P=15	0.2058	0.2276	0.2387	0.2446	0.2497
B=4, P=15	0.1964	0.2255	0.2361	0.2402	0.2387

mAP for a sample of different parameters in the Flick15k dataset.

Experimental Evaluation

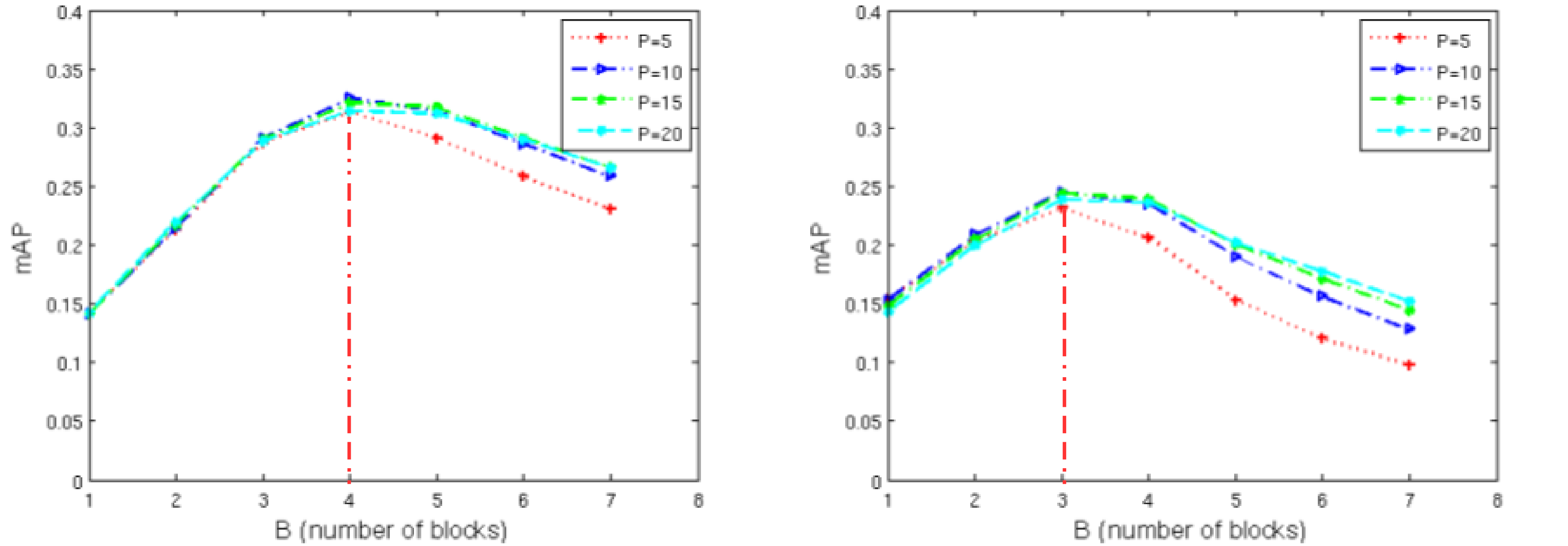


Figure 6: Performance of LKS varying the number of blocks for spatial division and the number of P for voting when $K = 150$. On the left, we see the performance on the Saavedra's dataset. On the right, we show the performance on the Flick15k dataset

Experimental Evaluation

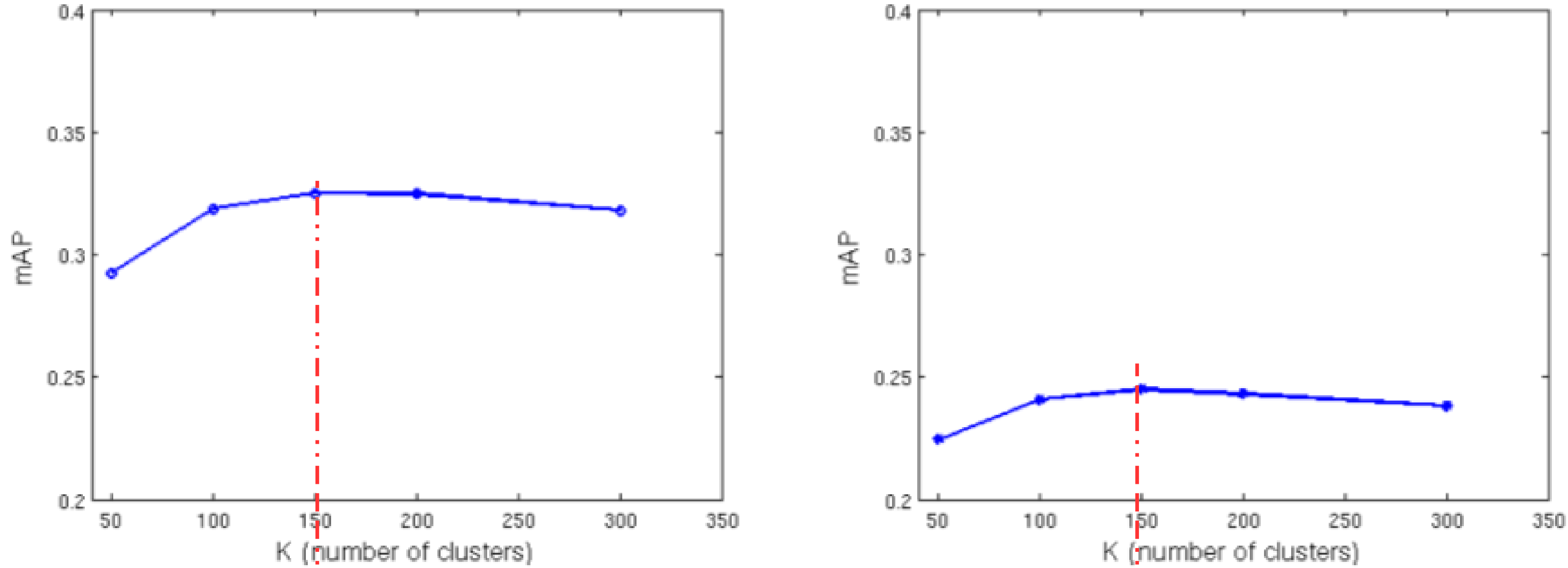


Figure 7: Performance of LKS varying the number of clusters on the Saavedra's dataset (on the left), and Flickr15k (on the right).

Experimental Evaluation

More efficiency?

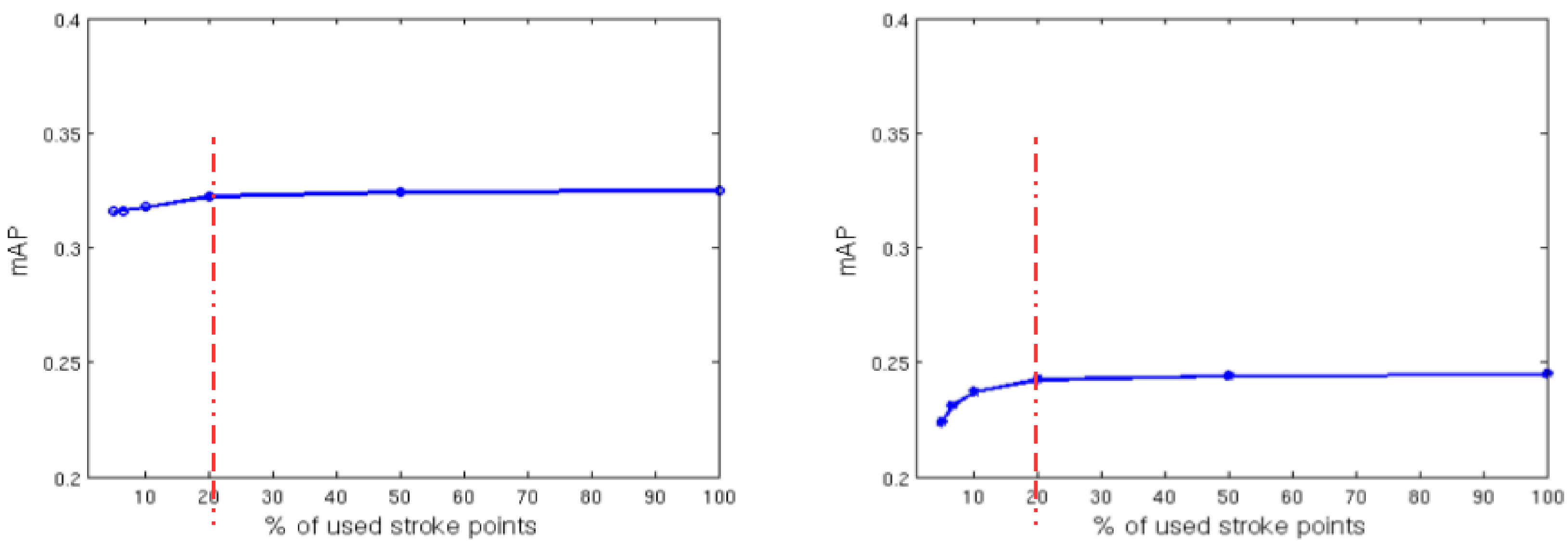
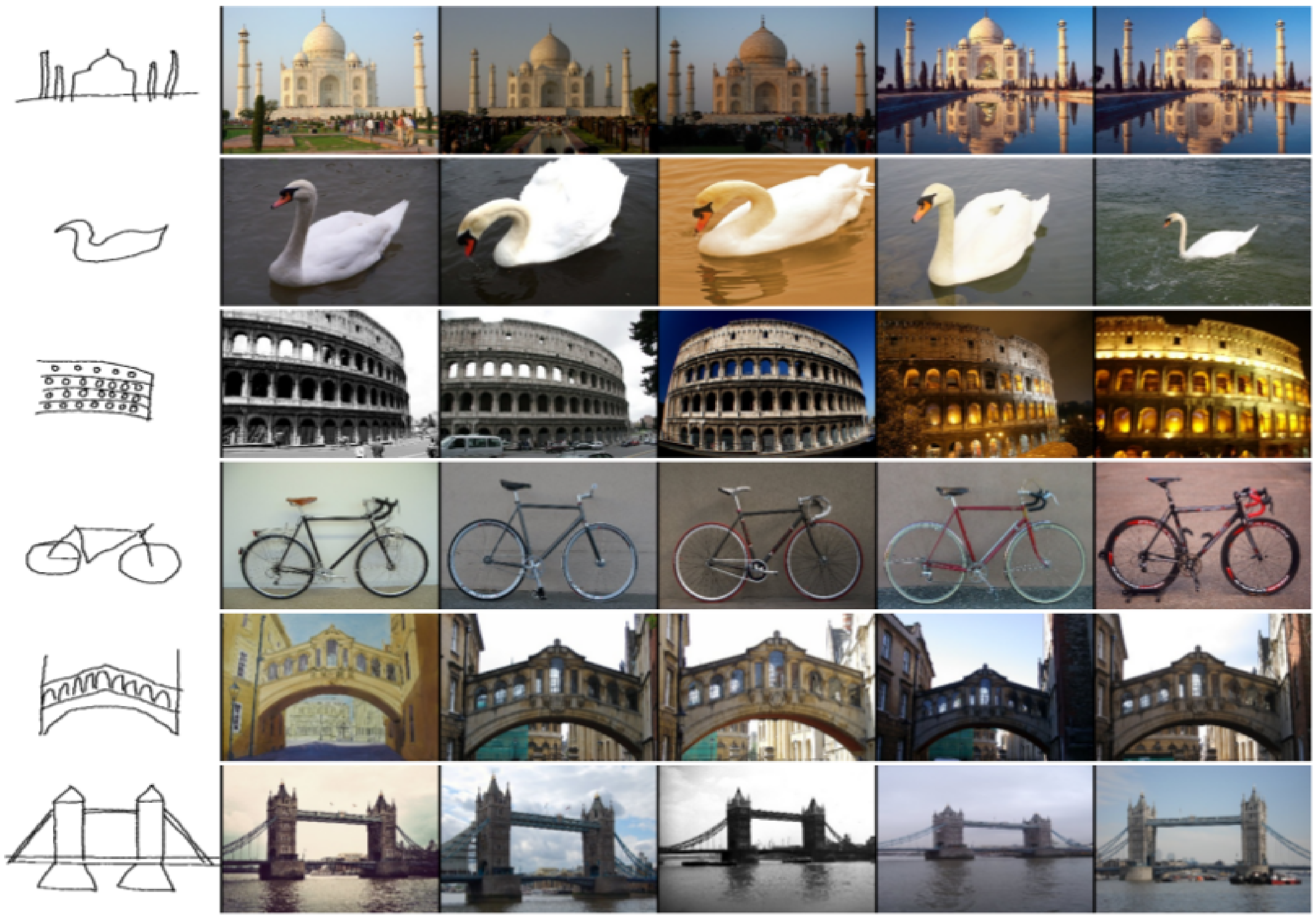


Figure 8: Impact of reducing the number of patches for computing the LKS histogram. On the left, the performance on the Saavedra's dataset. On the right, the performance on Flickr15k.

Results (own dataset)

Results (Flickr15K)



Thanks!!!