3D Visual Phrases for Landmark Recognition

Qiang Hao, Rui Cai, Zhiwei Li, Lei Zhang, Yanwei Pang, Feng Wu

天津大学 TIANJIN UNIVERSITY

Microsoft Research

Problem

- ❑ Utilize reconstructed **3D models** for landmark recognition
- ❑ Design effective **representation of 3D points** in 3D models
- ❑ Identify occurrences of 3D points from images

Limitations of existing solutions

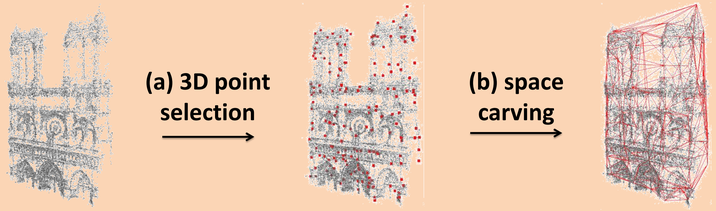
❑ ***Image-to-Image Matching***
- Sensitive to noises in DB images
- Lacking robustness to full projective transformations
- Relying on geometric verification

❑ ***3D-to-2D / 2D-to-3D Matching***
- Utilizing appearance of individual 3D points but **ignoring geometric structure** among them
- Relying on geometric verification

Our solution

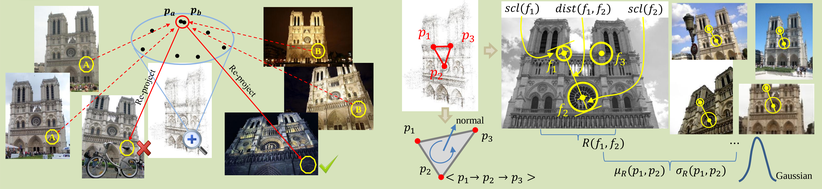
- ❑ ***Discover*** **3D Visual Phrases (3DVP)** from a 3D point cloud
- ❑ ***Describe*** appearance and geometric structure of 3DVPs
- ❑ ***Detect*** 3DVPs to identify landmarks in 2D images

1 Discovering 3DVPs from a point cloud



a) Select **popular** 3D points to **spatially cover** the model
b) Construct 3D surface → **triangular facets as 3DVPs**

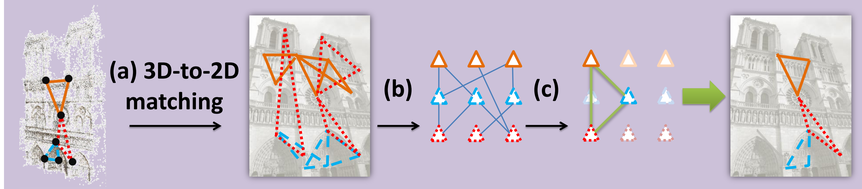
2 Describing 3DVPs (i.e., 3D point triplets)



❖ **Appearance of 3D points**
- ✓3D-to-2D re-projection
- ✓clusters of SIFT descriptors

❖ **Geometry between 3D points**
- ✓cyclic order of vertices
- ✓scale-distance correlation

3 Detecting 3DVPs from a 2D image



a) Find **putative 3DVPs** by 3D-to-2D matching (**appearance & geometry**)
b) Construct a **graph** over putative 3DVPs (by **linking non-conflicting pairs**)
c) Identify 3DVPs by finding maximal cliques from the graph

Evaluation

❑ **Dataset**
- **10 landmarks**, each with ~**1.5K DB images + 200 test images**
- Negative test images from **Oxford5K** dataset

❑ **Overall performance** (averaged over 10 landmarks)

3DVP boosts Precision & F-score, even WITHOUT geometric verification

	Matching (2D-to-2D / 3D-to-2D)				*Matching + Geometric Verification*			
	Bow	*2DVP*	*P2F*	*3DVP*	*Bow*	*2DVP*	*P2F*	*3DVP*
P	0.11	0.39	0.59	**0.94**	0.22	0.85	**1.00**	**1.00**
R	0.89	0.89	**0.90**	0.88	**0.80**	0.74	0.63	0.70
F	0.19	0.55	0.71	**0.91**	0.35	0.79	0.78	**0.83**

Dataset available at **http://landmark3d.codeplex.com**

Arc de Triomphe | *Big Ben* | *Church of Tyn* | *Dresden Frauenkirche* | *Independence Hall* | *Leaning Tower of Pisa* | *Notre Dame de Paris* | *St. Basil's Cathedral* | *St. Louis Cathedral* | *Trevi Fountain*