Translation-Symmetry-based Perceptual Grouping with Applications to Urban Scenes

Minwoo Park[1], Kyle Brocklehurst[1], Robert Collins[1], Yanxi Liu[1,2]

[1]Dept. of Comp. Sci. & Eng. [2]Dept. of Elec. Eng.

The Pennsylvania State University, USA

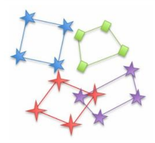
ABSTRACT

An important finding in our understanding of the human vision system is **perceptual grouping**, the mechanism by which visual elements are organized into coherent groups. Though grouping is generally acknowledged to be a crucial component of the mid-level visual system, its building blocks are largely ignored in computer vision, primarily due to the scarcity of mid-level cues and computational difficulties in constructing feature detectors for such cues. **We propose a novel mid-level visual feature detector** where the visual elements are based on the 2D translation subgroup of a wallpaper pattern. Different from previous state-of-the-art lattice detection algorithms for near-regular wallpaper patterns, our proposed method **can detect multiple, semantically relevant 2D lattices in a scene simultaneously**, achieving an effective translation-symmetry-based segmentation. Our experimental results on urban scenes demonstrate the use of translation-symmetry-based perceptual grouping for super-resolution and orientation estimation of building facades from a single view.

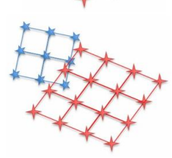
APPROACH

1) Low-level features such as KLT, MSER, and SURF points provide distinctive image locations and meaningful appearance descriptors.

2) Appearance descriptors allow us to separate the points into groups having visual similarity.

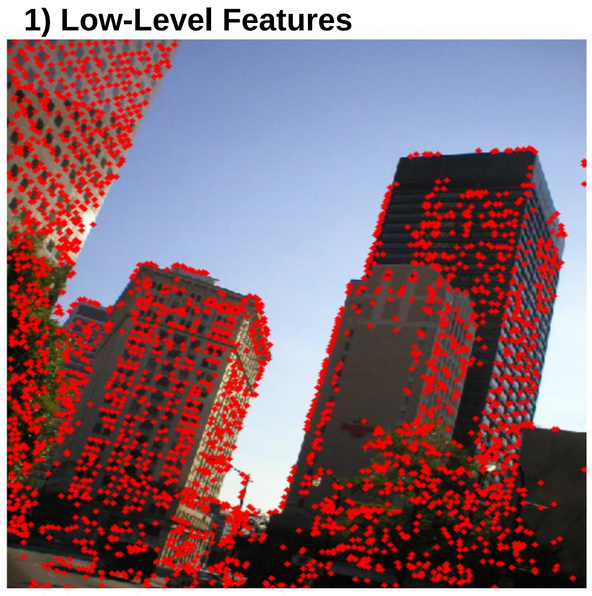


3) Bases of (t1,t2) vector pairs are identified by grouping evenly spaced points in a RANSAC procedure.

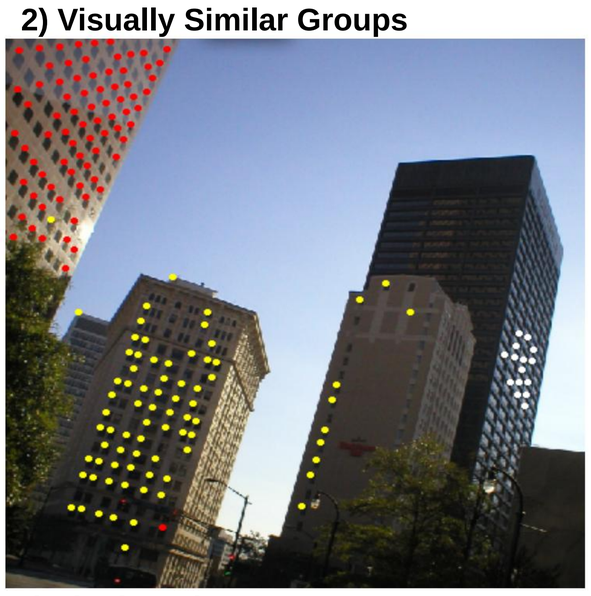


4) Lattices are grown iteratively from the vector pairs, grouping the points into one or more lattices.

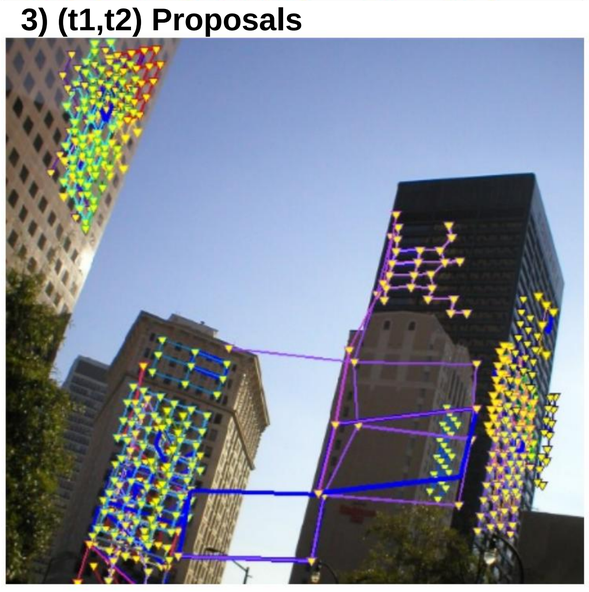
1) Low-Level Features



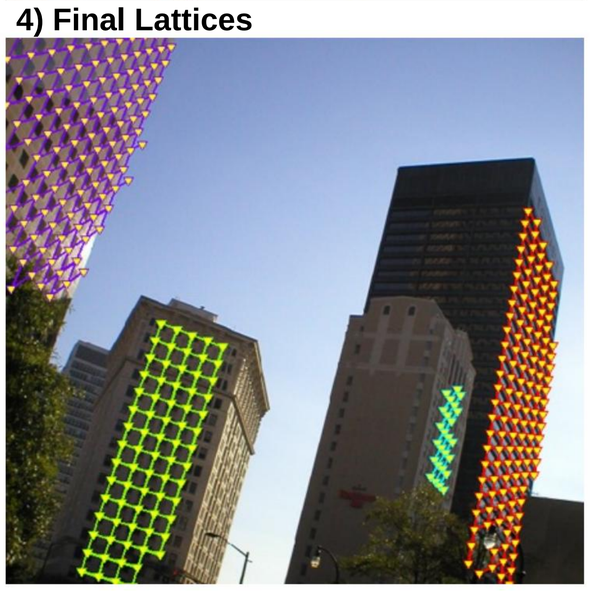
2) Visually Similar Groups



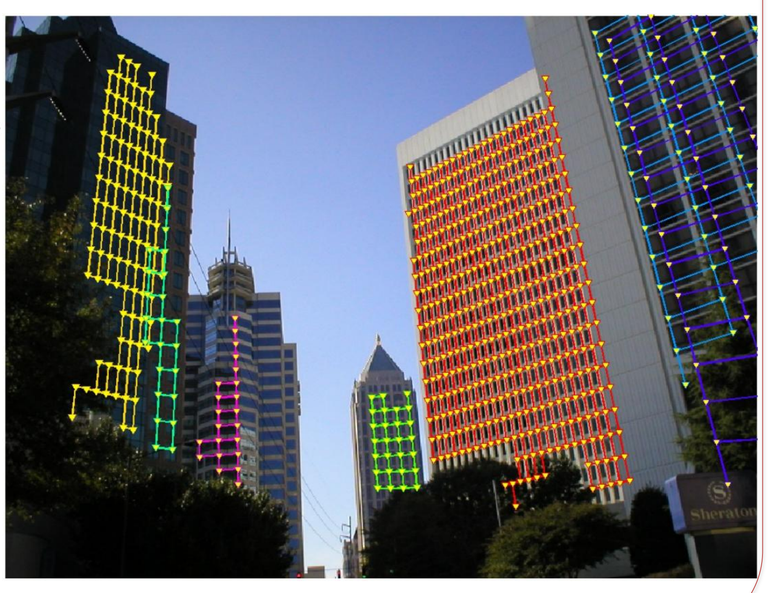
3) (t1,t2) Proposals



4) Final Lattices

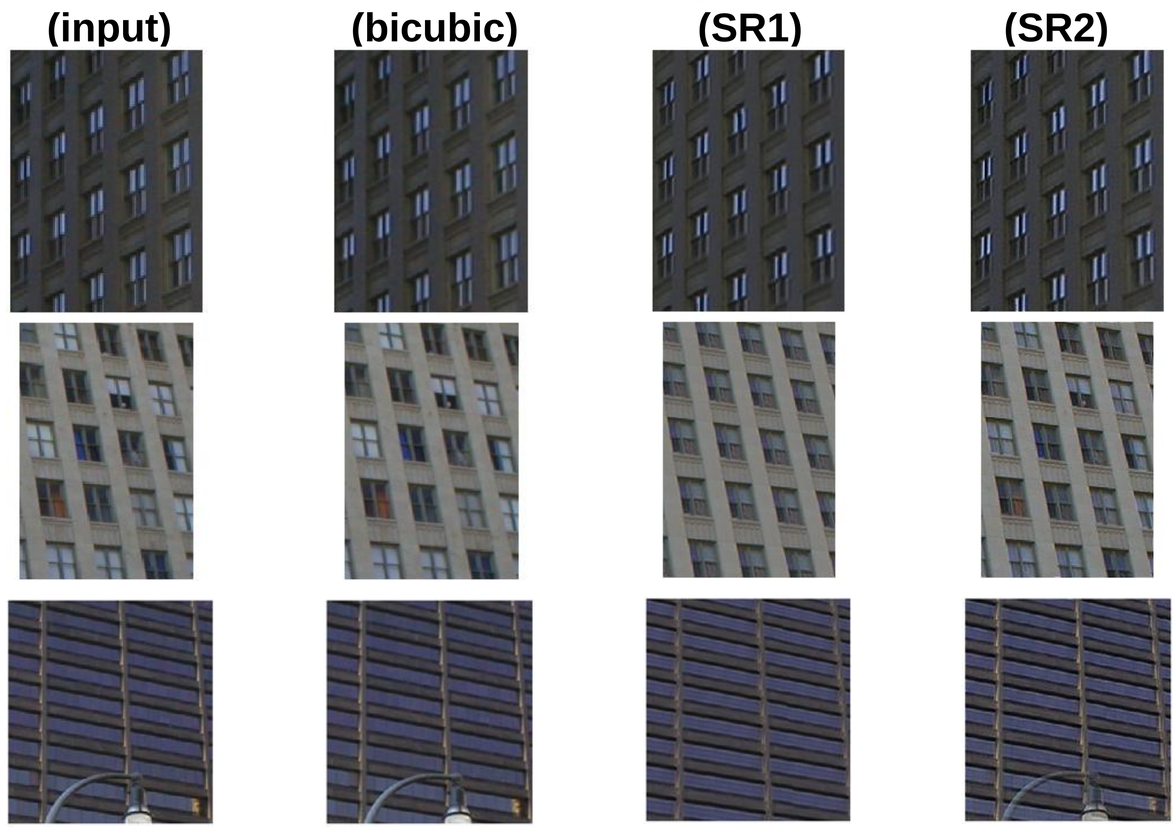


For each (t1,t2) proposal, a lattice is grown iteratively. In each iteration, a planar projective rectification is performed to bring all lattice points to a regular spacing. Then, points that have similar appearance and consistent spacing are included into the lattice. These help to refine the projective rectification and bring more points into alignment. This achieves a detailed and complete identification of multiple lattices, as seen on the right.



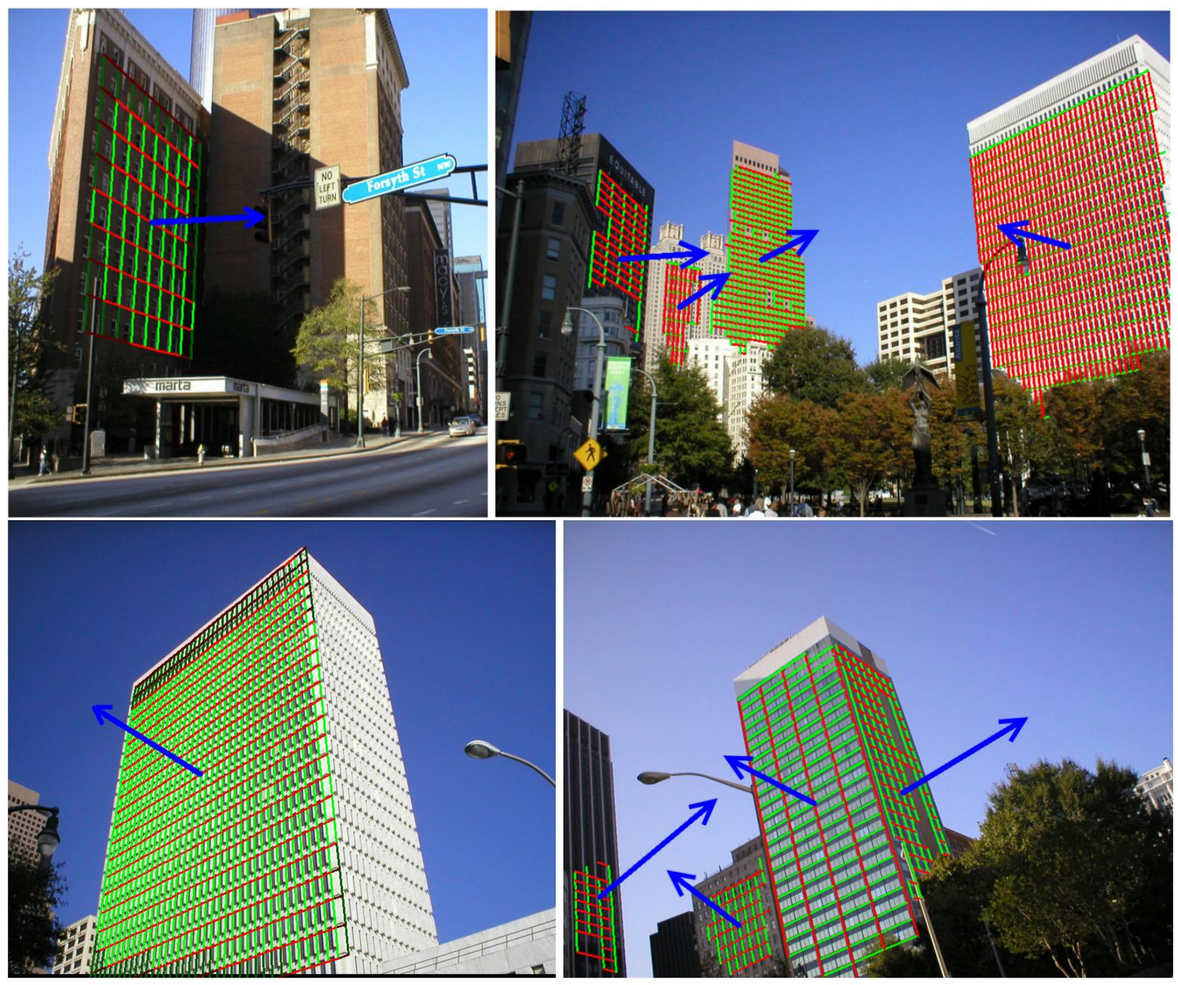
APPLICATIONS

1) SUPER-RESOLUTION - Recurrence of similar patches makes super-resolution from a single image possible. **Finding the correspondence between patches necessary to align them is the principle difficulty in this process, but is accomplished by identifying the lattice structure.** We rectify each quadrilateral in the 2D lattice to the same shape, compute a median texel, and perform deconvolution to get a high resolution (HR) image. In our basic approach (SR1), we map the HR image back on top of each observed texel. Our more advanced approach (SR2) uses a discrete cosine transform to map only the high-frequency components of the HR image back to the observed image, which preserves the lighting, shadows, and occlusions.



2) FACADE NORMAL VECTOR ESTIMATION - A perspectively distorted lattice that has been identified by our method will converge to two vanishing points, one in each direction of 2D repetition. Estimation of the orientation of a plane is straightforward if two vanishing points with a known angle between them are observed. By aligning our (t1,t2) vectors with the reflection axes of the building facade, we ensure that for all typical architecture our vanishing points are orthogonal.

We can accurately recover the normal of each facade relative to the camera. This also allows 3D reconstruction up to a scale factor to be accomplished for the scene. Facades that share a common edge can be scaled together. Relative scale between other facades must be performed by hand.



Detected Lattices: (red & green) **Recovered Normal:** Vectors (blue)

CONCLUSION

We present a novel algorithm that performs 2D translation-symmetry-based perceptual grouping and mid-level feature extraction. We show superior performance in detecting single and multiple lattices in an image over the state-of-the-art algorithm. We have demonstrated that the detected lattice structures are useful for estimating surface normal vectors and performing single view super-resolution. We plan to extend this work to single view 3D urban scene reconstruction and to use mid-level visual features for object categorization.

Acknowledgment: This work was supported in part by an NSF grant IIS-0729363 and a Google Research Award to Dr. Liu.

Project Page: http://vision.cse.psu.edu/research/perceptualGroupingUrbanAnalysis/index.shtml