

Local Regularity-driven City-scale Facade Detection from Aerial Images

Jingchen Liu¹

Yanxi Liu^{1,2}

¹ Computer Science and Engineering, ² Electrical Engineering
The Pennsylvania State University
University Park, PA 16802, USA
{jingchen, yanxi}@cse.psu.edu

Abstract

We propose a novel regularity-driven framework for facade detection from aerial images of urban scenes. Gini-index is used in our work to form an edge-based regularity metric relating regularity and distribution sparsity. Facade regions are chosen so that these local regularities are maximized. We apply a greedy adaptive region expansion procedure for facade region detection and growing, followed by integer quadratic programming for removing overlapping facades to optimize facade coverage. Our algorithm can handle images that have wide viewing angles and contain more than 200 facades per image. The experimental results on images from three different cities (NYC, Rome, San-Francisco) demonstrate superior performance on facade detection in both accuracy and speed over state of the art methods. We also show an application of our facade detection for effective cross-view facade matching.

1. Introduction

With the increasing popularity of Google maps and Microsoft Bing Maps, high resolution aerial image analysis has become an important yet challenging area of computer vision research. Facades are one of the most essential defining features of buildings. Automatic facade identification from large scale raw imagery becomes increasingly more desirable for city-scale computer modeling and end-user-initiated navigation. Most existing work on semantic facade parsing [7, 3, 22, 21, 13, 16, 17] assumes that the frontal facade views are given, while existing facade detection algorithms ([20, 4, 12, 18, 11, 15]) either require multiple views/depth information or else focus on street view images and segment at most a handful of facades per image. Typical airborne aerial images of urban areas can contain 100+ building facades (Figure 1). None of the published work deals with mass-produced wide-baseline aerial image facade detection at this throughput level, even though it has become an industry standard nowadays.

(a) An aerial view of New York City with detected facades



(b) An aerial view of San Francisco with detected facades

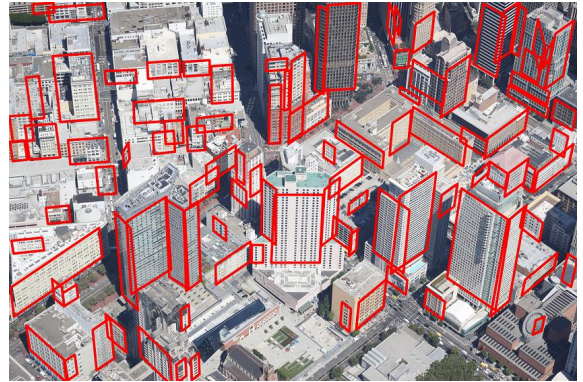


Figure 1. Our sample facade detection results from (a) NYC (top 200 candidates), and (b) San Francisco (top 100 candidates).

Near-regularity is common in many real world data sets (especially in urban scenes) while quantification of such near-regularities is computationally challenging. For facade detection from aerial imagery, we seek a regularity metric that is local, discriminative, efficient to compute and robust to noise. Some of our key observations are:

(1) *Localness*: due to the large scope each aerial image covers, a meaningful regularity measure for facades has to be sufficiently *local* without requiring any type of global homogeneity.

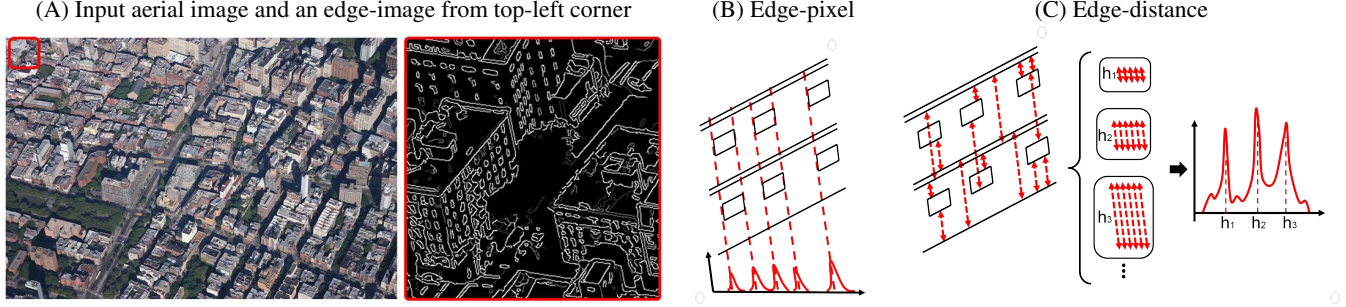


Figure 2. (A): an input aerial image and its Canny-edge extraction from top-left corner. (B) vertical edge-pixel distribution, calculated by projecting edge pixels along the vertical vanishing point direction; (C) vertical edge-distance distribution, calculated by extracting the distance between two non-vertical edge pixels that are along the vertical vanishing point direction.

(2) *Image abstraction and robustness to noise*: the extracted edge image (Fig. 2 (A)) seems to be a proper level of abstraction of the raw data in terms of regularity preservation. Furthermore, statistical distributions of the edge points may provide robustness against lighting variations and geometric deformations.

(3) *Sparsity and regularity*: there is a direct, well-defined relation between regularity captured in an image patch and the sparsity of its corresponding edge feature distributions when projected along specific orientations (Fig. 2 (B) (C)).

(4) *Sparsity and discriminative-ness*: Intuitively, peakier (sparser) edge distributions correspond to higher probability of a facade region, while edges of non-facade regions (e.g. bushes, rivers, roofs) are more randomly distributed. Formally, there exists a set of *distribution sparsity* measures [5], of which *Gini-index* is adapted in this work.

(5) *Computational feasibility*: The (re)projection of massive local edge distributions can be tedious, while in urban scenes there are only two primary directions of interest – vertical and horizontal. The unique vertical vanishing point is determined from calibration. The most plausible horizontal direction can be obtained by finding the direction of maximum sparsity of local edge distributions (Fig. 3).

Therefore, we formulate the facade detection problem into a regularity or sparsity optimization problem, where the *vertical regularity* of local image regions is well-defined and serves as an effective facade indicator. The *horizontal regularity* is at its maximum when its *corresponding sparsity* is optimized. We apply a greedy adaptive region expansion procedure for facade detection and growing, and finally use *integer quadratic programming* (IQP) to remove overlapping facades and achieve an optimal facade coverage.

2. Related Work

Automatic, fast and accurate extraction of facade regions from city-scale wide-baseline aerial images remains a challenging problem [14]. Zhao *et al.*, [20], Delmerico *et al.*, [4], Recky *et al.*, [12], B. Micusik and J. Kosecka [8] extract fa-

acades from ground-level imagery. These works either assume a 3D point cloud is given or depth information at each pixel can be estimated from multiple views. While it is reasonable to assume that street-view image databases are usually captured from a car-mounted camera system, range sensors may be less accurate/available for aerial images where the camera is located far away from the buildings. Furthermore, feature/pixel-based wide baseline stereo is challenged by the repetitive structure of building facades and thus prone to fail in precisely the regions we are most interested in. Bansal *et al.* [1] register satellite images with aerial images to detect roof contours, and then recognize facade regions using the ground plane homography between the satellite image and the aerial image. Wendel *et al.* [18], Park *et al.* [11], Schindler *et al.* [15] locate facade regions and recover their frontal views by detecting repetitive (lattice) patterns. These methods mainly work with large, highly repetitive building facades. The state-of-the-art lattice detection approach in [10] is computationally expensive for dense aerial images. Ceylan *et al.* [2] achieve facade acquisition by exploring the symmetries in a set of line segments. However in aerial views, due to the large distance between camera and facades, image resolution of each facade is limited for accurate line-segment extraction from edge pixels (Fig. 2).

The authors of TILT [19, 9] propose a low-rank seeking method to rectify a near-regular texture to its frontal-view under global affine/perspective transformations. Our work differs from TILT in at least two aspects: (1) **detection vs rectification**: Our algorithm detects facades from aerial images with massive background clutter while TILT requires good initialization of a facade region to recover its frontal view; (2) **distribution sparsity vs matrix sparsity**: we aim at a statistical interpretation of sparsity in relation to regularity where randomness leads to uniform distribution of local features while high regularity leads to peaky/sparse distributions. On the other hand, the maximization of sparsity of a matrix is the computational basis for TILT.

3. Our Approach

Our approach is primarily regularity-driven with the following assumptions: Since the camera is far enough from the scene in aerial urban images, we assume affine deformations on the building facades and approximate each facade as a parallelogram. In general, all facades in the same image share the same vertical direction determined by the vertical vanishing point, while the horizontal orientation of each facade differs and is unknown in advance. Therefore, the basic steps in our computational framework are:

(1) using a sliding window, we define a pair of *vertical* regularity measures along the direction of the vertical vanishing point to compute a likelihood estimation function for potential facade region evaluation; (2) we search for a potential *horizontal* direction of a local facade by maximizing horizontal regularity as evidenced by its edge-distribution sparsity; (3) we detect, grow and group local facade patches with consistent horizontal orientations, and maximize the facade regularity score; (4) we use integer quadratic programming (IQP) to remove overlapping facades, achieving an optimal facade coverage over the entire image.

3.1. Gini-Index for Regularity Evaluation

Gini-index is a measure of statistical dispersion or the inequality within a distribution. We adopt Gini-index in this work since it is invariant to scaling and cloning of the distribution and shown to be the best among six sparsity measures compared in [5].

Given a distribution with histogram bins $\mathbf{c} = (c_1, \dots, c_K)$ in sorted order $c_1 \leq c_2 \leq \dots \leq c_K$, the *Gini-index* is defined as

$$G(\mathbf{c}) = 1 - 2 \sum_{k=1}^K \frac{c_k}{\|\mathbf{c}\|_1} \left(\frac{K - k + \frac{1}{2}}{K} \right), \quad (1)$$

where $\|\mathbf{c}\|_1$ is a normalization (scaling) term, K is the total number of bins. $G(\mathbf{c})$ ranges from 0 (uniform distribution) to 1 (single peak). Since the histogram coefficients c_k directly measure the recurrence, with a higher value indicating stronger ‘regularity’, we define a *regularity likelihood metric* as the un-normalized Gini-index:

$$G_r(\mathbf{c}) = G(\mathbf{c}) \cdot \|\mathbf{c}\|_1 \quad (2)$$

3.2. Vertical Regularity

The vertical vanishing point direction is known at every point in the image from calibration data. Having high regularity along this unique direction is a *necessary condition* for a local region to be part of a facade, since most building facades are upright (Fig. 1). We propose two types of vertical regularity scores: *vertical-edge distribution regularity* and *vertical-distance regularity*, as illustrated in Fig. 2(B)(C). Vertical-edge distribution $\mathbf{c}^{(1)}$ is ob-

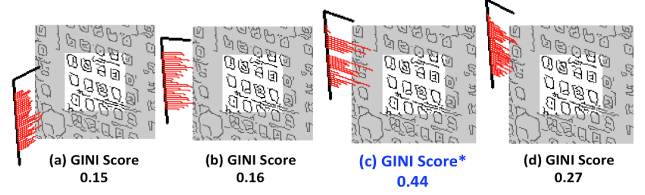


Figure 3. Horizontal regularity (Gini-index) measured along different horizontal directions. The black axis on the left indicates the local vertical direction and the hypothesized horizontal direction of projection. The histogram of projected edges are shown as red bars. The Gini-index score is shown below each subplot. The projection direction in (c) yields the **highest regularity** with the **most sparse distribution**, thus it is closest to the ‘correct’ horizontal direction of the facade.

tained by projecting local edge pixels along the vertical vanishing point direction; vertical distance distribution $\mathbf{c}^{(2)}$ is a histogram of distance between two non-vertical edge pixels that are along the vertical vanishing point direction (the two non-vertical edge pixels and the vertical vanishing point are on the same line) (Fig. 2(B)(C)). Image edge-pixels on vertical edges contribute to $\mathbf{c}^{(1)}$ and edge pixels on parallel horizontal lines contribute to $\mathbf{c}^{(2)}$.

For a local image patch centered at pixel (x, y) , we define its overall *vertical regularity* as the product of edge distribution regularity and distance regularity.

$$s(x, y) = G_r(\mathbf{c}^{(1)}) \cdot G_r(\mathbf{c}^{(2)}). \quad (3)$$

To balance computation and accuracy considerations we evaluate vertical regularity on a spatially sub-sampled grid (on a 16-pixel interval in all our experiments).

3.3. Horizontal Regularity

A key insight that guides our search for plausible horizontal direction of a local facade patch is the direct mapping between distribution sparsity and distribution regularity, as established conceptually in Fig. 2(B,C) and computationally by the Gini-index (Section 3.1). Instead of projecting local edge-pixels to the single vertical direction to calculate vertical regularities (Section 3.2), we project edge pixels to different potential horizontal directions and compare their quantified relative horizontal distribution regularities measured by the Gini-index. The highest Gini-index leads to the maximum regularity and thus the ‘correct’ horizontal direction for the local facade patch. This is found through an efficient coarse-to-fine 1D search within the angular range of $(-\pi/2, \pi/2)$. Fig. 3 illustrates one example where the correct horizontal direction yields the highest Gini-index sparsity value, and thus maximum regularity. Fig. 4 illustrates the top two directions of projection measured by the Gini-index, corresponding to the correct horizontal direction of the facade as desired.

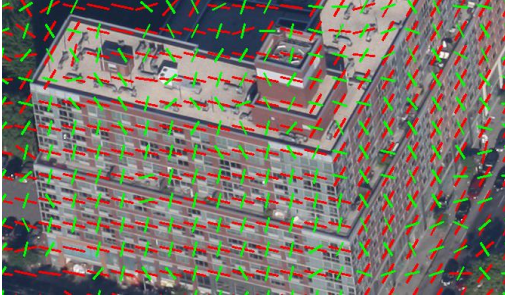


Figure 4. The first (red) and second (green) dominant local horizontal orientations are indicated.

3.4. Facade Detection via Greedy Region Expansion

As a result of the local regularity analysis above, we obtain, at each down-sampled grid point (x, y) , the vertical regularity score $s(x, y)$, the vertical vanishing point direction $\theta_v(x, y)$ and potential horizontal direction $\theta_h(x, y)$. Since each facade region can be represented as a parallelogram bounded by $X_{Min}, X_{Max}, Y_{Min}, Y_{Max}$ (Fig. 5), we can define a *facade* as: $\mathbf{f} \doteq (X_{Min}, X_{Max}, Y_{Min}, Y_{Max}, \dot{\theta}_h, \dot{\theta}_v)$. We can now treat facade region detection (maximization) as a regional regularity/sparsity maximization problem with the following objective function:

$$\mathbf{f}^* = \arg \max_{\mathbf{f}} \sum_{i=1}^I \sum_{j=1}^J s(x_{ij}, y_{ij}) \cdot a_{ij} \quad (4)$$

where $x_{ij} \in [X_{Min}, X_{Max}]$, $y_{ij} \in [Y_{Min}, Y_{Max}]$ satisfying:

$$\sum_j s(x_{ij}, y_{ij}) a_{ij} > \tau_r \cdot \max_k \left\{ \sum_j s(x_{kj}, y_{kj}) a_{kj} \right\} \quad \forall i = 1, \dots, I \quad (5)$$

$$\sum_i s(x_{ij}, y_{ij}) a_{ij} > \tau_r \cdot \max_k \left\{ \sum_i s(x_{ik}, y_{ik}) a_{ik} \right\} \quad \forall j = 1, \dots, J \quad (6)$$

where we introduce a binary indicator $a_{ij} \in \{0, 1\}$ to indicate whether the local dominant horizontal orientation is consistent with the global horizontal orientation within the entire facade, i.e.

$$a_{ij} = H(\tau_a - \angle(\theta_h(x_{ij}, y_{ij}) - \dot{\theta}_h)), \quad (7)$$

where $\angle()$ denotes the absolute angular distance, $H()$ is the Heaviside step function, and the threshold is set as $\tau_a = 10^\circ$. We also adopt a relative threshold τ_r in the constraint of equation (5) and (6) in order to stop the expansion of facade \mathbf{f} when the regularity score accumulated along any row/column drops below $\tau_r = 70\%$ of the peak value. To

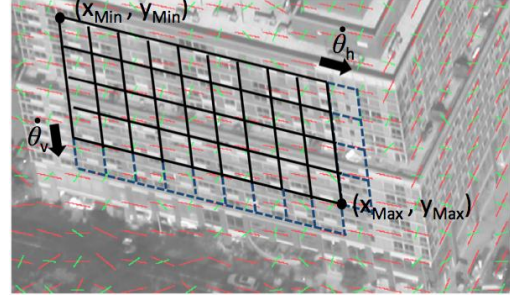


Figure 5. Illustration of facade expansion with moves that add a row or column, where the parallelogram point grid specifies the facade region.

solve this optimization problem computationally, we use a greedy adaptive region expansion strategy. Starting with a randomly chosen facade containing a single point (x, y) , we set $\dot{\theta}_h = \theta_h(x, y)$, $\dot{\theta}_v = \theta_v(x, y)$ and iteratively expand the facade boundary (parallelogram grid) horizontally and vertically as illustrated in Fig. 5.

During each attempt of an expansion (add a row/column to the parallelogram grid), the constraints of equation (5), (6) are examined and one of the following occurs: (1) the regularity score of the new row/column is lower than the adaptive threshold and the expansion is rejected; (2) the regularity score of the new row/column becomes the new maximum which causes existing facade region to dis-satisfy the constraints retrospectively, the facade then re-initializes at the new row/column leading to a facade-shift; (3) the expansion is accepted and the iteration continues until no valid expansion exists.

3.5. Multiple Facade Detection via Integer Quadratic Programming (IQP)

Via greedy adaptive expansion we can obtain a facade region from any initialization within the image. We can thus repeatedly initialize new facades within the regions not yet covered by existing facades and end up with multiple, possibly overlapping, facades $\{\mathbf{f}_i, i = 1, \dots, N\}$. IQP is used to select a subset of facades that have maximum regularity score and facade coverage, with minimum overlap.

Let the subset of facades be denoted as a binary indicator vector x , where $x_i = 1$ indicates the i th facade is selected. We specify a cost matrix $\mathbf{C}_{N \times N}$, where the diagonal entry c_{ii} is the regularity score of the i th facade, and the non-diagonal entry c_{ij} penalizes the overlapping area between facades i and j :

$$c_{ij} = -r \cdot \sum_{(x,y) \in \mathbf{f}_i \cap \mathbf{f}_j} s(x, y), \quad (8)$$

where r adjusts the tolerance of overlapping ($r = \infty$ means strictly no overlapping is allowed). In our experiments we

use $r = 2$, which tolerates small, partial overlapping. This also guarantees $C_{N \times N}$ to be positive-definite as long as there exists no image region being covered by more than two facades simultaneously.

The final solution is given by

$$x^* = \arg \max_x \{x^T C x\}, \quad (9)$$

and is solved by a spectral method [6] that first calculates then binarizes the principal eigenvector of C .

4. Experiments

We evaluate our algorithm on aerial image datasets collected from 3 different cities, New York City (NYC), San Francisco (SF) and Rome, with different camera viewing angles, facade areas and densities. The image resolution is 3744×5616 . Independent human raters manually label facade regions on 5 images of NYC, 2 images of Rome and 10 images of SF, yielding more than 3000 facades. NYC and Rome have relatively small and dense facades (241 facades per image on average), while the SF dataset contains larger facades with an average of 142 facades per image (Fig. 9). We illustrate some sample output of our algorithm below; please refer to more results in our supplemental material¹.

4.1. Regularity-based Per-pixel Facade Likelihood Validation

We first validate the effectiveness of the vertical regularity scores proposed (Section 3.2) as a facade likelihood indicator, *i.e.*, higher regularity (Gini-index) indicates higher likelihood of a facade region. We rank image pixels according to their vertical regularity score together with the per-pixel binary (facade/non-facade) labels, and obtain the precision/recall curves of per-pixel classification performance as shown in Fig. 6. The quantitative results show that the combination of two types of vertical regularities achieves the best performance.

4.2. Facade Detection Evaluation and Comparison

data set	Park <i>et al.</i> overall	ours overall	vertical regularity	horizontal regularity
NYC	114.3	6.0	0.8	4.8
SF	179.0	6.4	0.8	5.3

Table 1. computation time per image (minutes)

A. Comparison to state-of-the-art lattice detection algorithm [11]: Based on region-based (pixel-level) evaluation, we carry out a quantitative comparison with [11] (Fig. 7). The lattice detection approach (in blue) has good precision but a much lower recall rate, especially on the NYC/Rome

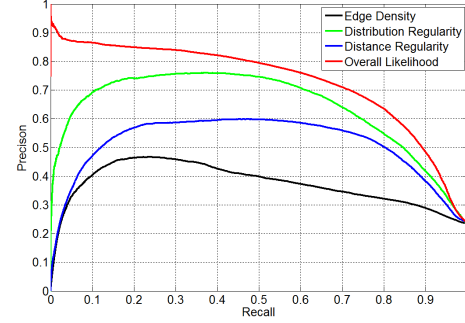
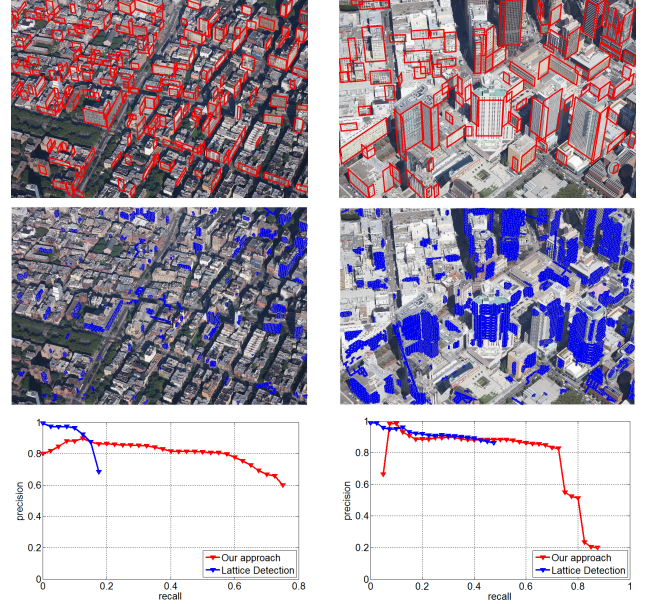


Figure 6. Precision-recall curve comparison of facade region classifications from an SF image. The black curve is a baseline approach that estimates likelihood based on local edge density. The distribution regularity score and distance regularity score performances are plotted in green and blue, respectively. The fused facade likelihood score from equation (3), plotted in red, outperforms the individual regularity score, confirming the synergetic nature of the regularity scores we propose.



(a) New York City

(b) San Francisco

Figure 7. Qualitative and quantitative comparison of our approach with lattice detection ([10]). Top: Our results in red; Middle: Lattice detection results in blue; Bottom: Quantitative comparison of precision-recall curve.

datasets where the buildings are relatively small. Our algorithm performs consistently across datasets with different building densities and facade sizes under a fixed set of parameters. Our approach demonstrates an obvious advantage in speed over [11] (Table 1). Both algorithms are implemented in C++/OpenCV and run on the same machine with a 2.7GHz i7 CPU and a 64-bit OS.

¹<http://vision.cse.psu.edu/research/facade/index.shtml>

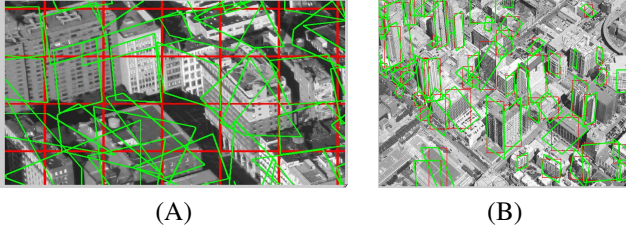


Figure 8. Sample results (green) using TILT [9], initialized from (A) a predefined grid in red; (B) the detection output from our algorithm in red.

B. Comparison against TILT [9]

The TILT algorithm [9] has been shown to effectively rectify a near-regular texture into its frontal-view when a rectangular bounding box in the image is given for initialization. We simulate such input by (1) segmenting each aerial image into equal grid-cells (Fig. 8(A) red) and (2) outlining facade locations discovered by our algorithm (Fig. 8(B) red), then evaluate the output of TILT respectively (Fig. 8 green). Such input contains background clutter (*e.g.*, roofs, streets) and geometric deformation which seem to have considerably hindered the performance of TILT.

C. Qualitative and quantitative results of our algorithm

Some sample qualitative results of our facade detection are given in Fig. 9. The precision/recall plots can be viewed in Fig. 7 and the computation time in Table 1. Our algorithm detects facades in city-scale aerial test images with above 80% accuracy and close to 80% recall rates in minutes on average.

4.3. Application: Cross-view Matching

Given the facades detected from each single view and the calibration information, we are able to determine the 3D orientation of the facade (with a depth ambiguity) and the frontal view of the facade in each aerial image. This information simplifies facade matching across views: (1) via epipolar constraints and facade orientation consistency, we are able to remove a large proportion of false-positive matches; (2) when evaluating the matching score of two facades from different views, we can resolve the depth ambiguity. Thus the recovered frontal view facades from different images have a consistent orientation and scaling. Therefore, the appearance matching can be achieved via straightforward template matching or normalized cross-correlation (NCC). Fig. 10 shows the top-10 facade matching results from two different cities and the automatically rectified frontal view facades using our algorithm.

5. Conclusion

Urban scene data sets used in this research have provided a testbed for us to explicitly explore and validate the rela-

tionship between local regularity in an image and sparsity as statistical distributions of low level features (edges, edge lengths). This offers an alternative perspective of computational regularity in general and the effectiveness of local regularity in urban scene applications in particular. Our work on automated extraction of facade regions from high-resolution aerial views is a timely effort for urban scene analysis and a concrete step forward. We have presented a novel, local-regularity-driven method that differs from and complements existing work on facade rectification, analysis and synthesis. Our approach leverages properties of regularity of building facades, namely vertical and horizontal alignment of edge features. Vertical regularities are used as a likelihood indicator of the presence of a facade region. Search for an appropriate local horizontal direction is also guided by exploring the horizontal regularities. The optimization of facade regions is performed using a greedy adaptive expansion that maximizes regularity within parallelogram-shaped image regions. A final set of facade regions is found by suppressing excessively-overlapping facade hypotheses using binary quadratic programming. We have compared experimentally and quantitatively with state of the art algorithms for facade extraction and show superior performance in both speed and accuracy.

We hypothesize that frontal view building facade patches will prove to be an effective intermediate-level representation of urban scenes, suitable for many tasks such as fly-through visualization, geolocation, and wide baseline stereo reconstruction. As a preliminary result we have demonstrated an example of cross-view matching of extracted facades from different images (Fig. 10).

Acknowledgment This work is partially funded under NSF grants IIS-1248076 and IIS-1144938. We thank Google for the image data and A. Pope, M.H. Lin for their insightful comments.

References

- [1] M. Bansal, H. S. Sawhney, H. Cheng, and K. Daniilidis. Geo-localization of street views with aerial image databases. In *Proc. MM*, pages 1125–1128. ACM, 2011. 2
- [2] D. Ceylan, N. J. Mitra, H. Li, T. Weise, and M. Pauly. Factored facade acquisition using symmetric line arrangements. In *Proc. EUROGRAPHICS*, volume 31, pages 671–680, 2012. 2
- [3] D. Dai, M. Prasad, G. Schmitt, and L. V. Gool. Learning domain knowledge for facade labelling. In *Proc. ECCV*, pages 1–14, 2012. 1
- [4] J. A. Delmerico, P. David, and J. J. Corso. Building facade detection, segmentation, and parameter estimation for mobile robot localization and guidance. In *Proc. Intelligent Robots and Systems (IROS)*, pages 1–8, 2011. 1, 2
- [5] N. Hurley and S. Rickard. Comparing measures of sparsity. In *Proc. MLSP*, pages 1–16, 2008. 2, 3



Figure 9. Top row: human labeled ground-truth facades (blue). Middle and bottom rows: facades detected (red) by our local-regularity driven algorithm.

- [6] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. In *Proc. CVPR*, pages 1–8, 2005. [5](#)
- [7] A. Martinovic, M. Mathias, J. Weissenberg, and L. Gool. A three-layered approach to facade parsing. In *Proc. ECCV*, pages 1–14, 2012. [1](#)
- [8] B. Micusik and J. Kosecka. Multi-view superpixel stereo in man-made environments. *IJCV*, 89(1):106–119, 2010. [2](#)
- [9] H. Mobahi, Z. Zhou, A. Yang, and Y. Ma. Holistic 3d reconstruction of urban structures from low-rank textures. In *Proc. ICCV workshop*, pages 1–8, 2011. [2](#), [6](#)
- [10] M. Park, K. Brocklehurst, R. Collins, and Y. Liu. Deformed lattice detection in real-world images using mean-shift belief propagation. *T-PAMI*, 31(10):1804–1816, 2009. [2](#), [5](#)
- [11] M. Park, K. Brocklehurst, R. Collins, and Y. Liu. Translation-symmetry-based perceptual grouping with applications to urban scenes. In *Proc. ACCV*, pages 1–14, 2010. [1](#), [2](#), [5](#)
- [12] M. Recky, A. Wendel, and F. Leberl. Facade segmentation in a multi-view scenario. In *Proc. 3DIMPVT*, pages 358–365, 2011. [1](#), [2](#)
- [13] H. Riemenschneider, U. Krispel, W. Thaller, M. Donoser, S. Havemann, D. Fellner, and H. Bischof. Irregular lattices for complex shape grammar facade parsing. In *Proc. CVPR*, pages 1–8, 2012. [1](#)
- [14] R. Roberts, S. N. Sinha, R. Szeliski, and D. Steedly. Structure from motion for scenes with large duplicate structures. In *Proc. CVPR*, pages 1–8, 2011. [2](#)
- [15] G. Schindler, P. Krishnamurthy, R. Lublinerman, Y. Liu, and F. Dellaert. Detecting and matching repeated patterns for au-

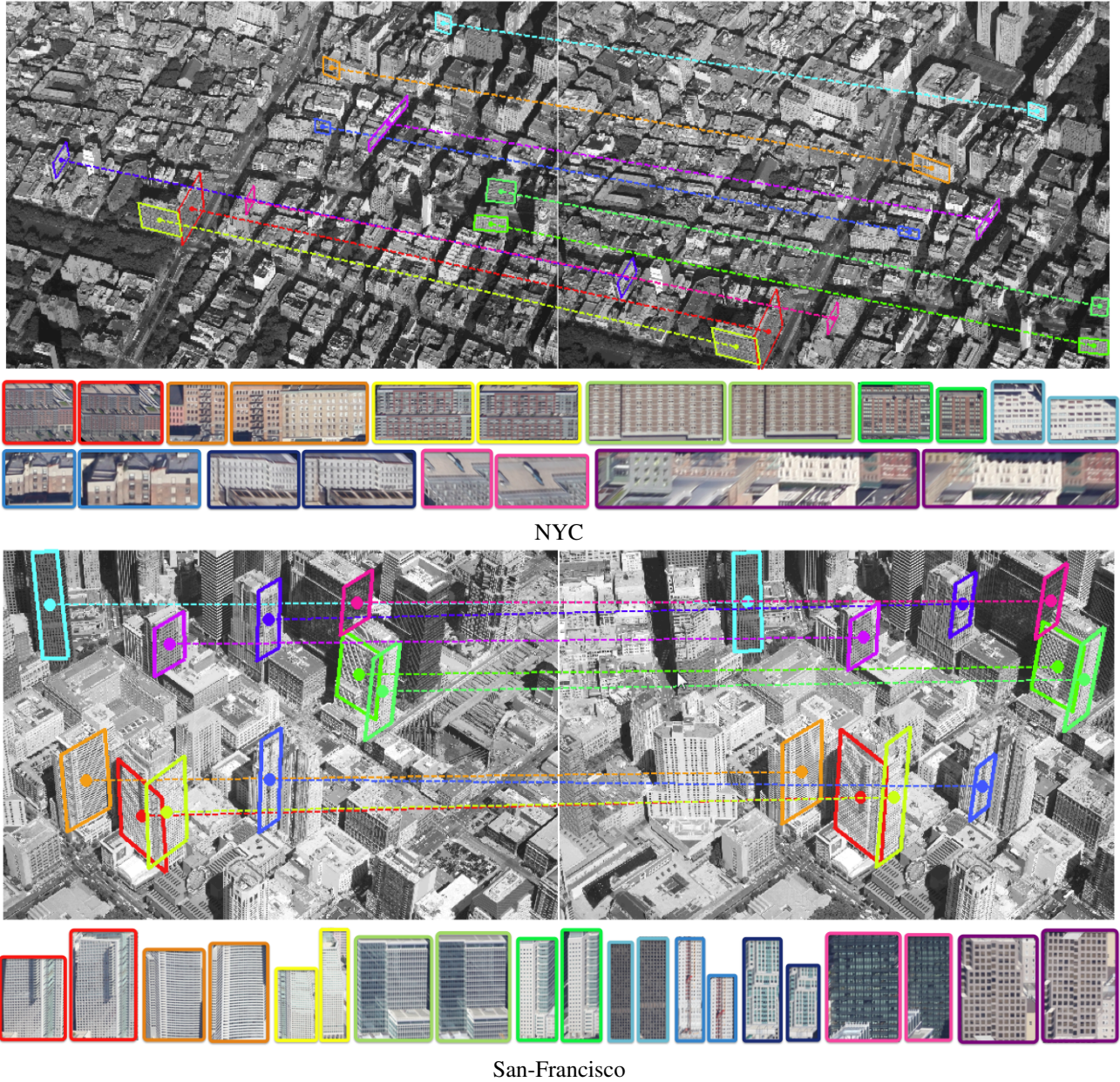


Figure 10. Examples of cross-view facade matching from different aerial images. The facades are detected and rectified automatically by our local regularity-based algorithm.

- automatic geo-tagging in urban environments. In *Proc. CVPR*, pages 1–8, 2008. 1, 2
- [16] O. Teboul, I. Kokkinos, L. Simon, P. Koutsourakis, and N. Paragios. Shape grammar parsing via reinforcement learning. In *Proc. CVPR*, pages 1–8, 2011. 1
- [17] O. Teboul, L. Simon, P. Koutsourakis, and N. Paragios. Segmentation of building facades using procedural shape priors. In *Proc. CVPR*, pages 1–8, 2010. 1
- [18] A. Wendel, M. Donoser, and H. Bischof. Unsupervised facade segmentation using repetitive patterns. In *Proc. DAGM*, volume 6378, pages 51–60, 2010. 1, 2
- [19] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma. Tilt: Transform invariant low-rank textures. *International Journal of Computer Vision (IJCV)*, 99(1):1–24, 2012. 2
- [20] P. Zhao, T. Fang, J. Xiao, H. Zhang, Q. Zhao, and L. Quan. Rectilinear parsing of architecture in urban environment. In *Proc. CVPR*, pages 1–8, 2010. 1, 2
- [21] P. Zhao and L. Quan. Translation symmetry detection in a fronto-parallel view. In *Proc. CVPR*, pages 1–8, 2011. 1
- [22] P. Zhao, L. Yang, H. Zhang, and L. Quan. Per-pixel translational symmetry detection, optimization, and segmentation. In *Proc. CVPR*, pages 1–8, 2012. 1