Translation Symmetry Detection: A Repetitive Pattern Analysis Approach

Yunliang Cai and George Baciu
GAMA Lab, Department of Computing
The Hong Kong Polytechnic University, Hong Kong, China
{csyunlcai,csgeorge}@comp.polyu.edu.hk

Abstract

Translation symmetry is one of the most important pattern characteristics in natural and man-made environments. Detecting translation symmetry is a grand challenge in computer vision. This has a large spectrum of real-world applications from industrial settings to design, arts, entertainment and eduction. This paper describes the algorithm we have submitted for the Symmetry Detection Competition 2013. We introduce two new concepts in our symmetric repetitive pattern detection algorithm. The first concept is the bottom-up detection-inference approach. This extends the versatility of current detection methods to a higher level segmentation. The second concept is the framework of a new theoretical analysis of invariant repetitive patterns. This is crucial in symmetry/non-symmetry structure extraction but has less coverage in the previous literature on pattern detection and classification.

1. Introduction

This paper discusses the algorithm we submitted to the translation symmetry detection contest in Symmetry Detection Competition. Translation symmetry detection is widely used in the analysis of higher-level visual structures, such as buildings, cloth and fabric patterns, and crystal-structure materials as well as in the analysis of bio-medical imaging. The detection technique is applied to applications such as image retrieval, shape reconstruction, and texture rendering.

Traditional translation symmetry detection is often modeled as a top-down matching process. A global deformable template is defined, and is then continuously tuned until the shape of template can align the features from the target image. The top-down approach is fast and efficient, but have limitations in versatility and robustness.

From top-down to bottom-up. Unlike the traditional methods, we try to propose a bottom-up detection approach. The bottom-up approach starts without a prior template. It collects a subset of repetitive patterns from a given image, and assigns a meaningful structure to describe the spatial organization of the repetitive patches in the image. The inference is based on the relative locations of the patterns.

The bottom-up approach requires additional time in structure inference, but extend the detection of symmetry types. The structure inference is possible to analyze more than one symmetry type in the current image. For example, the inference is allowed to extract both translation symmetry and rotation symmetry simultaneously. In this paper, for the purpose of algorithm evaluation, the inference is simplified and limited to translation symmetry structure only.

Invariant repetitive pattern. Another main ingredient of our detection algorithm is the analysis of invariant repetitive pattern. A set of local image patches are considered as repetitive patterns if they share the same image content.

In our algorithm, we use deformable quadrilaterals to represent repetitive patterns. The shapes of the quadrilaterals are determined by the joint registration of all repetitive patches. The invariant component extracted from the aligned patches can serve as image templates to detect more repetitive patterns in the same image or others.

Interactive pattern detection. We introduce user-interactions in our algorithm. The interactive idea is inspired by interactive image segmentation, i.e. the graph-cut method. Our algorithm allows the user to draw a small set (at least one) of local image patches on the input image as initial repetitive patterns. The algorithm then aligns the patches and obtains an invariant patch template for detection. The template can be incrementally updated when new repetitive patterns are detected and aligned.

2. Related work

Baseline algorithm. The main baseline algorithm used for evaluation is Park's deformed lattice detection method [3]. Park's method used deformed lattice as a global template and KLT features for grouping and alignment. The result of the detection method is a lattice structure, which encodes the location of the extracted features and specifies dimensions of the lattice grid. The Park's method was later improved as a interactive detection method using SIFT/SURF features [4]. In contrast, our algorithm does

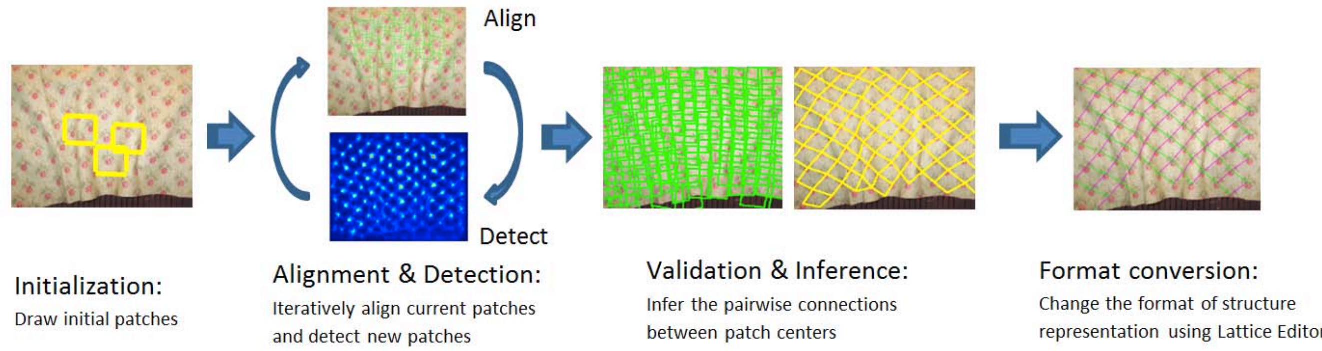

Figure 1. The overview of the proposed repetitive pattern detection and bottom-up structure inference algorithm in this paper.

not provide a valid lattice structure. The results obtained from our algorithm is a set of quadrilateral patches and their pairwise connections (i.e. adjacent matrix). We manually convert our results to the lattice structure formats using the Lattice Editor [1]. The other algorithm that can be used for evaluation is Wu's frieze-like pattern detection [5]. It is specified for detecting repeated patterns in building scenes. Particularly, in the symmetry detection for the front-view of a building, Zhao et al [6] has proposed a solution using an image segmentation technique which is similar to our algorithm output. But these methods are difficult to extend to highly deformed patterns.

Our algorithm is based on our previous work published in [2] [1] with slight modifications. In this paper, we introduce the HOG descriptor in our registration model and simplify the structure inference to lattice structure only.

3. Algorithm overview

The overall algorithm is an iterative propagation and update process, it continuously detects the perspective new repeated patterns and aligns the involved patches to maintain the invariant representation. The complete algorithm is shown in Figure 1 and is described in the following steps:

Initialization: The user draws a small number of quadrilateral patches (one or more than one patches) as initial repetitive patches. The algorithm exploits the joint alignment model in Section 4 in order to align the patches and extracts the initial pattern template. If only one patch is provided, this patch will be directly assigned as the template.

Detection: For each repetitive patch detected, we deform the invariant template to align with the detected patch. We then perform a normalized cross-correlation (NCC) around the local area of the particular detected patch, detecting new patches which share the content and shape of the deformed invariant template.

Alignment: For all the detected patches, we apply a joint registration (congealing) as described in Section 4. By tuning the shapes of the quadrilateral patches, all patches are deformed and aligned. After all the patches are matched, we can update the invariant template.

Validation and inference: If there are new patches detected during the detection stage, we go back to the detection phase again in order to perform the invariant template alignment. Otherwise, we go to the structure inference phase. The inference estimates the local connection orientations using the filtering of beamlet functions. The orientations of the beamlets are determined by the NCC map obtained in the detection stage. Only two major orientations are considered for translation symmetry detection.

4. Detection and joint alignment method

For an image $I : \Omega \to \mathbb{R}$, $\mathbf{I}_p : \Omega \to \mathbb{R}^{d(p)}$ is a vector-valued function which represents the local patch centered at $p \in \Omega$. $d(p)$ is the number of pixels in $\mathbf{I}_p$ which solely depends on the shape of $\mathbf{I}_p$. $\mathbf{I}_p$ is free-form quadrilateral.

Let G_p be the affine transform that warps an $m \times m$ regular square patch to quadrilateral $\mathbf{I}_p$, such that $\mathbf{I}_p \circ G_p^{-1}$ is an $m \times m$ square. Our first assumption is the generative assumption. There exists an $m \times m$ invariant template $\mathbf{T}$ that satisfies

$$||\mathbf{T} - \mathbf{I}_p \circ G_p^{-1}||^2 < \epsilon \tag{1}$$

for p in the image area with ϵ sufficiently small. G_p can be written as a 3×3 matrix obtained by exponential map:

$$G_p = \sum_{k=1}^{6} \mathrm{Exp}(a_p^k E_k) \tag{2}$$

as E_k for $k = 1, \ldots, 6$ are the matrix basis for the tangent space [2]. Another assumption is repetition assumption. If $\mathbf{I}_p$ and $\mathbf{I}_q$ are repetitive patterns where $p \neq q$, we will have

$$||\mathbf{I}_p \circ G_p^{-1} - \mathbf{I}_q \circ G_q^{-1}||^2 < \epsilon \tag{3}$$

The joint alignment model. If both the generative and repetition assumptions hold, given a set of patches $\{\mathbf{I}_p\}_{p \in \Lambda}$, we can deform $\mathbf{I}_p$ by tuning parameters $\{a_k\}_{1,\ldots,6}$ in (2). For each $p \in \Lambda$, we can formulate the joint registration as the minimization of the following functional:

[1] http://vision.cse.psu.edu/research/latticeEditor/latticeEditor.shtml

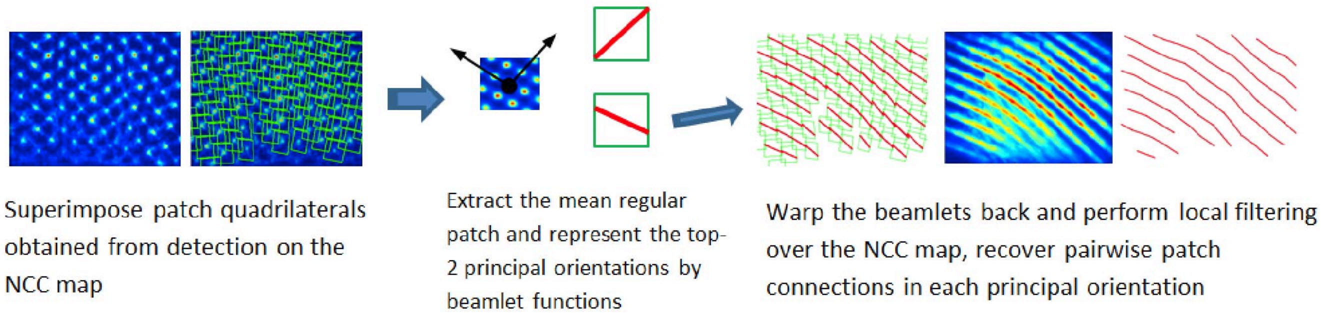

Figure 2. The structure inference algorithm.

$$F_p(\mathbf{a}_p) = \sum_{q \in N(p)} ||\phi(\mathbf{I}_p) - \phi(\widetilde{\mathbf{I}}_q(\mathbf{a}_p))||^2 + \sum_{q \in N(p)} ||\phi(\mathbf{I}_p) - \phi(\widetilde{\mathbf{T}}(\mathbf{a}_p))||^2 + \sum_{q \in N(p)} ||\mathcal{L}A_p||_F^2 \quad (4)$$

where ϕ is a dense image feature descriptor, i.e., HOG descriptor. $\mathbf{a}_p = (a_p^1, a_p^2, \ldots, a_p^6)$ is used as parametrization for representing $\widetilde{\mathbf{T}}(\mathbf{a}_p) = \mathbf{T} \circ G_p(\mathbf{a}_p)$ and $\widetilde{\mathbf{I}}_q(\mathbf{a}_p) = \mathbf{I}_q \circ G_q^{-1} \circ G_p(\mathbf{a}_p)$. $\mathcal{L}A_p$ is the Laplacian of $A_p = \mathrm{Log}(G_p)$ over the matrix space spanned by $E_1, \ldots, E_6$.

The detection model. For the current repetitive patch set $\{\mathbf{I}_p\}_{p\in\Lambda}$, we can obtain the corresponding deformation set $\{G_p\}_{p\in\Lambda}$ using joint alignment (4). By the generative assumption, we can generate deformed a template set $\{\mathbf{T} \circ G_p\}_{p\in\Lambda}$. Using these deformed templates, we can calculate the NCC correlation in local image area near positions in Λ. The new repetitive patterns can thus be detected and their patch centers are added to Λ. In other word, the repetitive set $\{\mathbf{I}_p\}_{p\in\Lambda}$ propagates as new positions are involved after each detection stage.

5. Structure inference

The structure inference is based on the NCC map obtained during the detection phase. The goal of structure inference is to recover the pairwise connections between the detected patches. Shape information of the set $\{\mathbf{I}_p\}_{p\in\Lambda}$ is also used for analyzing the oriented connections. The main steps in the structure inference are illustrated in Figure 2.

Principal orientation extraction. We first extract the dominant local orientations. Denote the NCC map as R. For each $\mathbf{I}_p$, the corresponding quadrilateral is superimposed on p in R, producing local patch $\mathbf{R}_p$. As all patches are aligned already, the peak positions enclosed in each quadrilateral are aligned accordingly. This indicates that the relative positions of the local neighbors also satisfy our generative assumption. We then rectify all superimposed quadrilaterals over R back to $m \times m$ regular patches, and obtain the set $\{\mathbf{R}_p \circ G_p^{-1}\}_{p\in\Lambda}$. We can have

$$\overline{\mathbf{R}} = \frac{1}{|\Lambda|} \sum_{p\in\Lambda} \mathbf{R}_p \circ G_p^{-1}. \quad (5)$$

By using the Radon transform, one can easily extract the top two orientations, θ_1 and θ_2 in patch $\overline{\mathbf{R}}$. This leads to two beamlet functions b_1 and b_2. Each beamlet function is a 2D function defined on $[-m/2, m/2]^2$ where there is a line segment crossing the origin with orientation θ_1 and θ_2 respectively. The beamlet is required to have non-zero values on the line segment and zeros elsewhere. The beamlet is normalized so that $\int b_1(x)dx = \int b_2(x)dx = 1$.

Beamlet filtering. For position set Λ over R, we can have $\{b_1 \circ G_p\}_{p\in\Lambda}$ and $\{b_2 \circ G_p\}_{p\in\Lambda}$. In order to extract the first orientation of a translation symmetry structure, one can take $\{b_1 \circ G_p\}_{p\in\Lambda}$ as local deformed templates to conduct local filtering over R. For each p, we let $b_1 \circ G_p$ as filter to convolute the image area around $\mathbf{I}_p$. By simply adding all the local filter response, one can have the first orientation of the desired lattice structure. Similarly, by substituting $b_1 \circ G_p$ with $b_2 \circ G_p$, the second orientation can be obtained. Both orientations form a global lattice-like structure, which represents the desired translation symmetry structure.

6. Algorithm evaluation

The proposed algorithm is tested using the dataset provided by the Symmetry Detection Competition 2013. Our results are manually converted to the lattice format using the Lattice Editor, with the points in a lattice are validated by our pairwise patch connections. The runtime of our algorithm depends on the number of repeated patches detected. The average runtime on a 2.7GHz dual-core parallel computing environment is 5 min. Some images in the urban scenes category take longer than 30 min due to the large number (≥ 800) of repetitive patterns detected.

6.1. Comparison with baseline algorithm

We compare our method with the suggested baseline algorithm [3]. As the baseline algorithm is an automatic method, we also point out some extreme cases for the completeness of our discussion.

Frieze patterns. The frieze pattern images contain a large number of repetitive patterns in most affine translation symmetry structure. For images in this category, one input

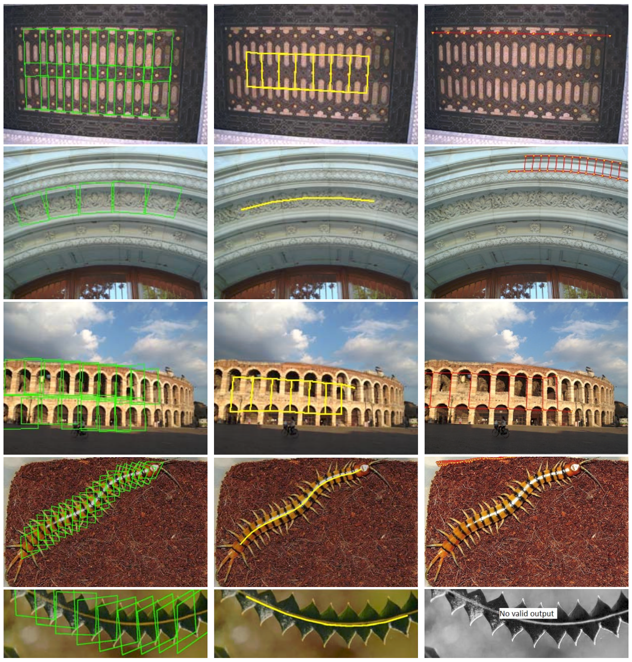

Figure 3. The selected results of frieze pattern test images. Left to right: the detected repetitive patterns of our algorithm, the inferred symmetry structures, and the results of the baseline algorithm.

patch, i.e. an initial regular patch drawn by user, is often enough for detection. A number of selected results are shown in Figure 3. Because a large number of images in this category do not fit the deformed lattice structure, the baseline algorithm could be affected by the lack of lattice features. Our patch-based algorithm can pass the non-lattice cases.

General wallpaper test images. The wallpaper pattern images include patterns that can form smooth surface geometries. For most cases, our algorithm can successfully reconstruct the spatial organizations as the baseline algorithm. However, because the baseline algorithm is fully automatic, it might be trapped in the detection of less significant structures in an image and would fail consequently. Selected results are presented in Figure 4.

Fence-like wallpaper test images. The fence-like pattern images contain look-through lattice structures. Our algorithm requires additional input patches (3 to 5 patches) in order to extract a stable template. Our algorithm can still infer the correct lattice structures in most cases. However, for this image category, the structure inference is more difficult than other categories. The repetitive patches in fence-like wallpaper images have very high variety in their contents, it is then more difficult for the alignment model to group the quadrilaterals. The selected results can be seen in Figure 5.

Urban scene test images. The images in urban scene category contain often contain large number of repetitive patterns as shown in Figure 6. This costs our algorithm longer time in computing the patch alignment, but this

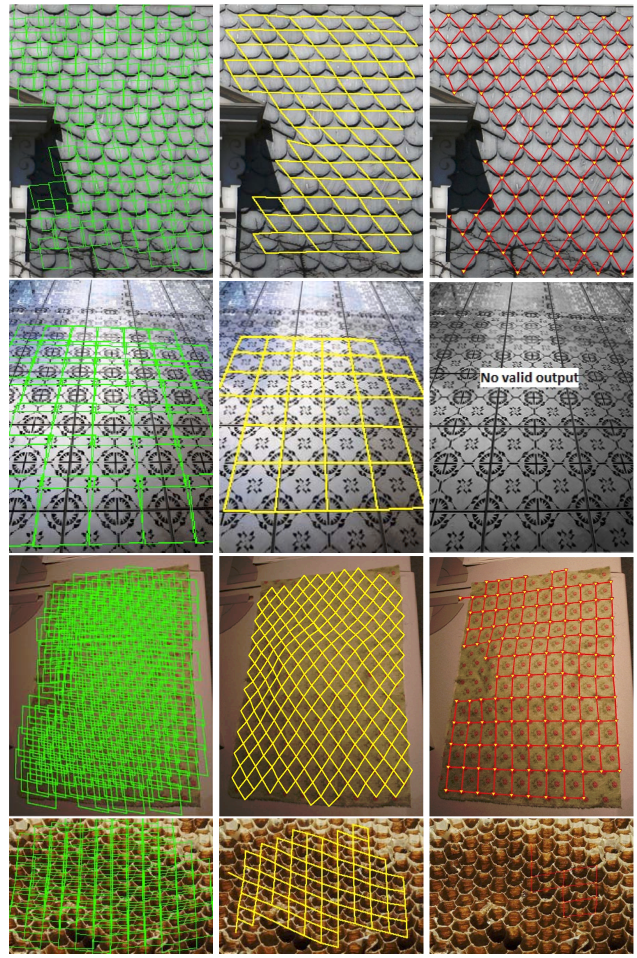

Figure 4. Selected results of general wallpaper pattern images. Our detection and inference method (1st and 2nd column) can detect highly distorted geometric scenes and have better performance than the baseline algorithm (3rd column).

alignment allows us to analyze structures that suffer from large distortions. In addition, our interactive initialization allows user to select the desired structures he/she may be interested in as urban scene images often contain multiple structures. By combining separated user-initialized inference results, it is easy to generalize the single structure inference to multiple structure inference.

6.2. Test on multi-structure detection

For images contain multiple translation symmetry structures, a multi-structure grouping is needed. We compare our method with the perceptual grouping method [4] in multi-structure detection.

The extension of our method from single structure inference to multiple structure inference is straightforward. We draw different input patches for individual structure in an image, and let the detection and inference be conduct separately. The resulting multiple structures are obtained by

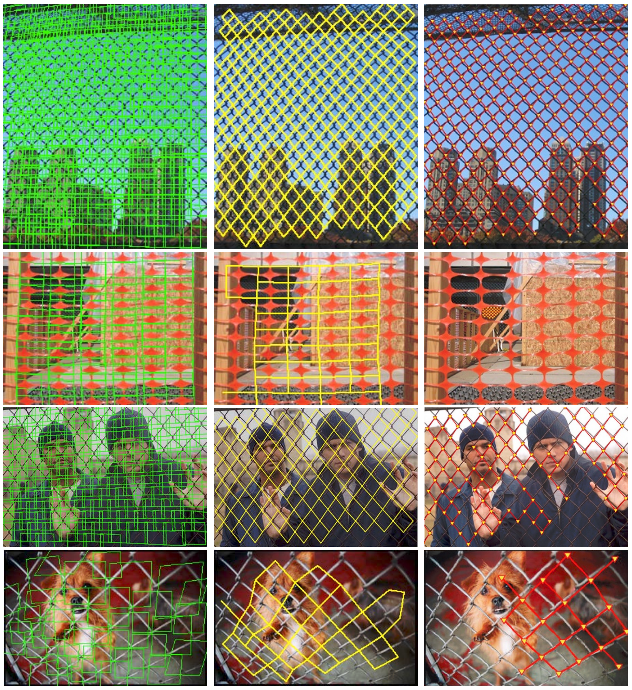

Figure 5. Selected results of fence-like pattern test images. Our detection and inference method (1st and 2nd column) suffer from the high variations of the patch contents, but can still infer the correct pairwise connections in most cases. The baseline algorithm (3rd column) is affected by the same problem.

simply combing all separately inferred structures. A test result is shown in Figure 7. The further tests with comparison to Park's perceptual grouping method can be found in Figure 8.

7. Conclusions

We proposed a translation symmetry detection method based on the analysis of invariant repetitive patterns. Unlike the popular top-down matching method, our method provides a bottom-up inference approach for estimating the translation symmetry structure. The algorithm has better detection performance in highly curved or distorted symmetry structures than the traditional ones.

8. Acknowledgement

The authors would like to acknowledge the partial support of the Hong Kong Innovation Technology Council HKRITA Grant ITP/002/10TP, the Hong Kong Research Grants Council Grant GRF PolyU 5101/11E, and PolyU 5100/12E.

References

[1] Y. Cai and G. Baciu. Higher level segmentation: Detecting and grouping of invariant repetitive patterns. In *CVPR 2012*. 2

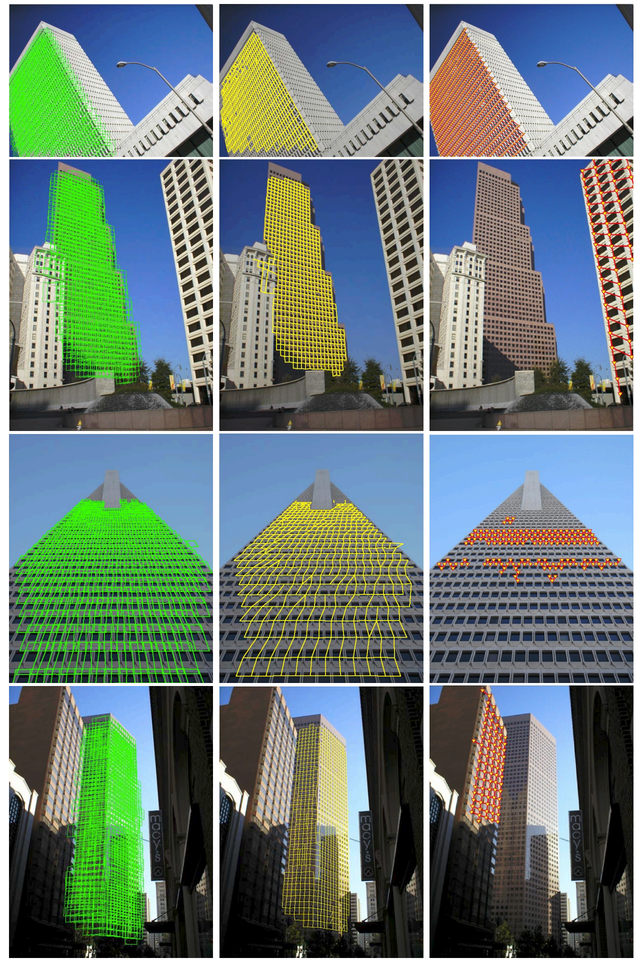

Figure 6. Selected results of urban scene test images. Our algorithm (first two columns) has better performance than baseline (3rd column) in largely distorted scenes.

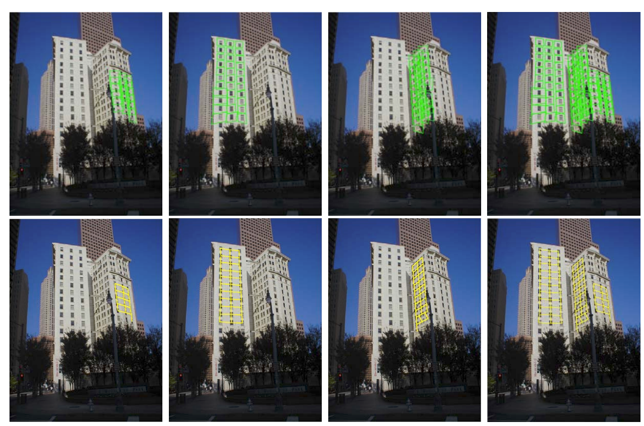

Figure 7. The combination of separately initialized results. Top: the individual detected patches and the combined result at last. Bottom: the separately inferred structures and the combined structure at last.

[2] Y. Cai and G. Baciu. Detecting, grouping, and structure inference for invariant repetitive patterns in images. *IEEE Trans on*

Figure 8. Selected results of interactive grouping for multiple structures. Our method (1st and 2nd column) can perform the perceptual grouping for different structures like the method in [4] (3rd column).

Image Processing, (doi:10.1109/TIP.2013.2251649), 2013. 2

[3] M. Park, K. Brocklehurst, R. T. Collins, and Y. Liu. Deformed lattice detection in real-world images using mean-shift belief propagation. *Trans on PAMI*, 31(10):1804–1816, 2009. 1, 3

[4] M. Park, K. Brocklehurst, R. T. Collins, and Y. Liu. Translation-symmetry-based perceptual grouping with applications to urban scenes. In *ACCV*, pages 329–342. Springer, 2010. 1, 4, 6

[5] C. Wu, J. Frahm, and M. Pollefeys. Repetition-based dense single-view reconstruction. In *CVPR 2011*. 2

[6] P. Zhao and L. Quan. Translation symmetry detection in a fronto-parallel view. In *CVPR 2011*. 2