Exploiting Repetitive Object Patterns for Model Compression and Completion

L.Spinello[1,2], R.Triebel[2], D.Vasquez[3], K.O.Arras[1], R.Siegwart[2]

[1]: Social Robotics Lab, University of Freiburg, Germany [2]: Autonomous Systems Lab, ETH Zurich, Switzerland [3]: ITESM Campus Cuernavaca, Mexico

Abstract

Many man-made and natural structures consist of similar elements **arranged in regular patterns**. **In this paper we present an unsupervised approach for discovering and reasoning on repetitive patterns of objects in a single image.** We propose an unsupervised detection technique based on a voting scheme of image descriptors. We then introduce the concept of ***latticelets***: minimal sets of arcs that generalize the connectivity of repetitive patterns. **Latticelets** are used for building polygonal cycles where the smallest cycles define the sought groups of repetitive elements. The proposed method can be used for **pattern prediction and completion** and high-level image compression. **Conditional Random Fields** are used as a formalism to predict the location of elements at places where they are partially occluded or detected with very low confidence. Model compression is achieved by extracting and efficiently representing the repetitive structures in the image. Our method has been **tested on simulated and real data** and the quantitative and qualitative result show the effectiveness of the approach.

The Paper in a Nutshell

- **Unsupervised** approach on **single images** for **discovering** repetitive patterns of objects
- **pattern prediction** and **completion**

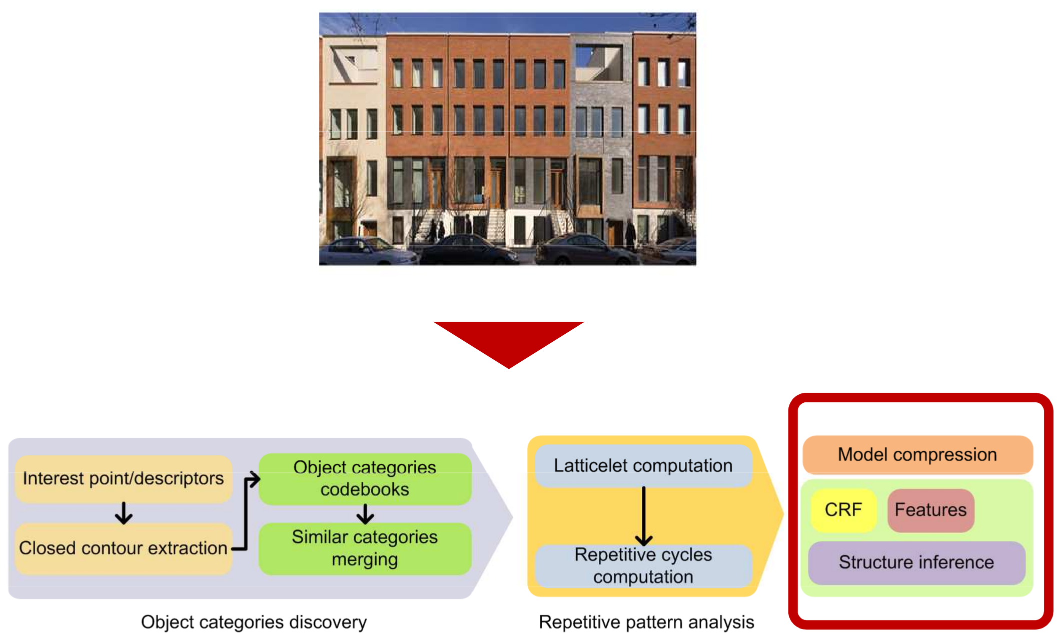
- Model completion
  - **Latticelets**: generalize pattern connectivity through a graph connectivity analysis
  - **Conditional Random Field** built on repetitive cycles for predicting missing elements

- Model compression
  - **Efficient representation** of repetitive structures in the image

- Tested on simulated and real data

Flow Diagram

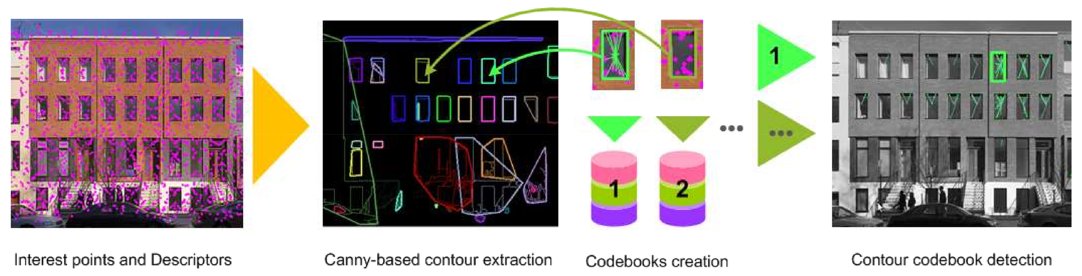
Input (single image)



Processing Pipeline

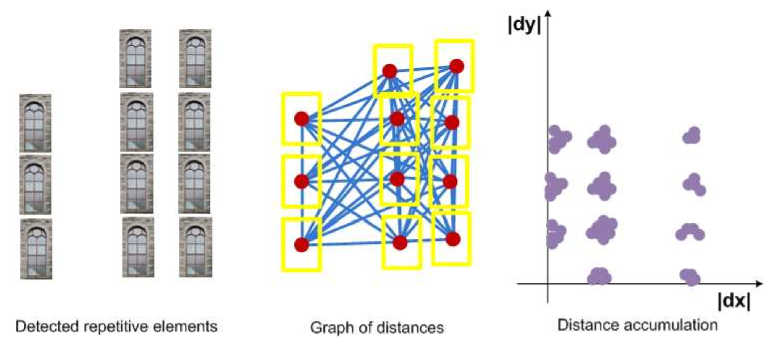
Detection and Analysis of Repetitive Object Patterns

Unsupervised Detection of Repetitive Objects



Extraction of mutually similar objects. For each closed contour, a **codebook of descriptors** is created that contains relative displacements to the object centers (votes). Then, the descriptors of each object are matched against the descriptors in the image.

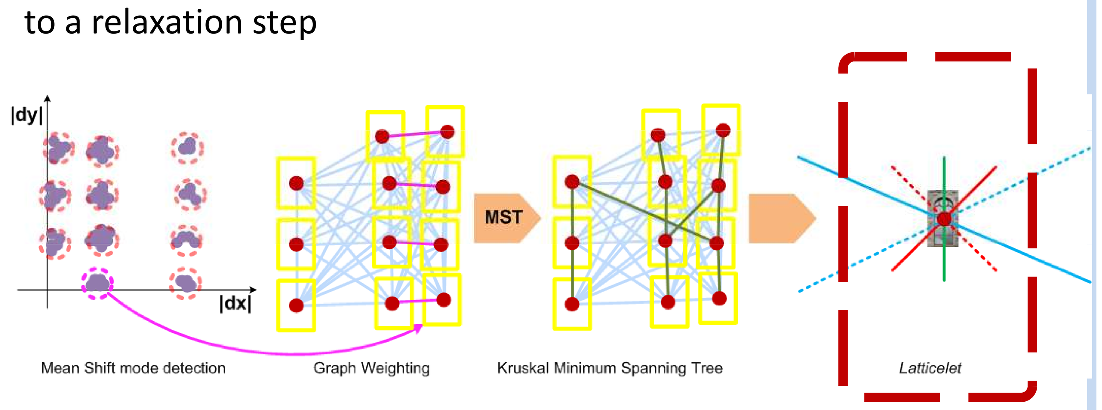
Analysis of Repetitive Objects



Objects of the same kind are detected. A **complete graph is built** and the **relative distances are accumulated** in Cartesian plane.

1. Latticelets

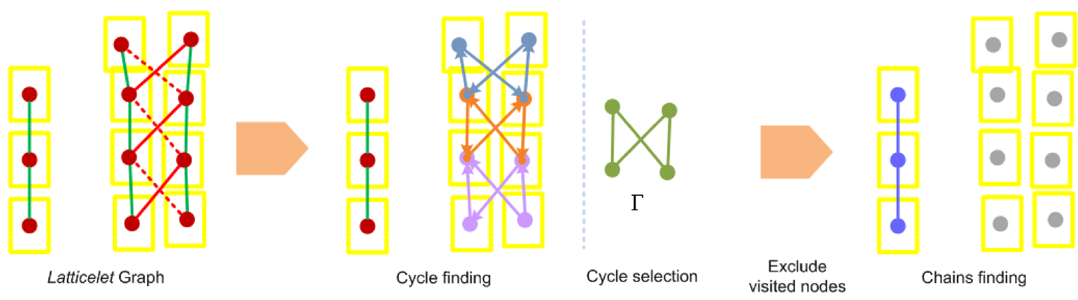
We call **latticelet**, the minimal set of repetitive arcs that are needed to represent the original lattice. Each object kind is associated to a latticelet that generalize its repetition pattern. Our method is able to cope with small perspective distortions thanks to a relaxation step



Repetitive distances in x and y are **clustered via mean-shift**, the arcs are **reweighed** by their mode convergency score. The solid and dotted lines in the **latticelet figure represent the possible directions** espressed by the selected |dx| and |dy|.

2. Cycles and Chains

In order to incorporate **higher level knowledge of the repetitive pattern** of the neighborhood, we use **cycles** composed of **latticelets arcs**. Aim: find minimal size repetitive polygons.



From the **graph created by an incremental set of latticelet's arcs**, small repetitive cycles are selected by using a **Breadth-first Search** algorithm. Chains are created on the remaining nodes that have not been satisfied by any polygonal cycles G .

Structure Inference using Conditional Random Fields (CRF)

In many cases, objects can not be detected due to occlusions or low contrast in the image. In general, the problem of these false negative detections **can not be solved**, as there is **not enough evidence of the occurrence of an object**. In our case, we can use the **additional knowledge** that **all objects of the same kind are grouped according to a repetitive pattern**. Using this information, we can infer the existence of an object, even if its detection quality is very low.

Conditional Random Fields formalism

CRF are applied on a lattice built with polyg. cycles found in the previous step. Aim: **infer presence of an object in each node.**

Object type: τ

Presence of object in location x (bin.var): $l_\tau(\mathbf{x})$

$$p(\mathbf{l}_\tau \mid \mathbf{z}) = \frac{1}{Z(\mathbf{z})} \prod_{i=1}^{N} \varphi(\mathbf{z}_i, l_{\tau i}) \prod_{(i,j)\in\mathscr{E}} \psi(\mathbf{z}_i, \mathbf{z}_j, l_{\tau i}, l_{\tau j}) \qquad \textbf{(A)}$$

Node potential: $\varphi(\mathbf{z}_i, l_{\tau i}) = e^{\mathbf{w}_n \cdot \mathbf{f}_n(\mathbf{z}_i, l_{\tau i})}$

Edge potential: $\psi(\mathbf{z}_i, \mathbf{z}_j, y_i, y_j) = e^{\mathbf{w}_e \cdot \mathbf{f}_e(\mathbf{z}_i, \mathbf{z}_j, l_{\tau i}, l_{\tau j})}$

Features are related to detection quality and to detection consistency in the polygonal neighborhood

$$\mathbf{f}_e(q_i, q_j, l_{\tau i}, l_{\tau j}) = \begin{cases} \frac{1}{\gamma}(f_{e1} \quad f_{e2}) & \text{if } l_{\tau i} = l_{\tau j} \\ (0 \quad 0) & \text{else} \end{cases} \quad \text{with} \quad \begin{array}{l} f_{e1} = \max(\mathbf{f}_n(q_i, l_{\tau i}), \mathbf{f}_n(q_j, l_{\tau j})) \\ f_{e2} = \max_{G \in \mathscr{G}_{ij}}(\mathbf{f}_n(\eta(G), l_{\tau i})), \end{array}$$

$$\mathbf{f}_n(q_i, l_{\tau,i}) = 1 - l_{\tau,i} + (2 l_{\tau,i} - 1) q_i$$

Learn a Generalized Network Structure

The standard way to apply CRFs to our **problem would consist in collecting a large training data set where all objects are labeled by hand** and for each object type τ a pair of node and edge features is learned so that (A) is maximized. Issues:

- For a given τ, there are different kinds of lattice structures in which the objects may appear in the training data.
- Only objects of categories that are present in the training data can be detected.

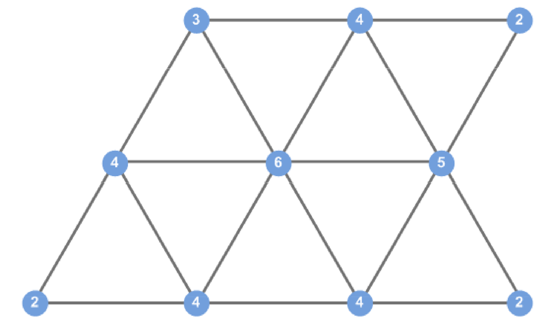
In training phase we only obtain a set of node and edge weights, which do not depend on the network geometry but only on **its topology**, we can generate training instances by setting up networks with a given topology and assigning combinations of low and high detection qualities to the nodes. The advantage of this is that **we can create a high variability of possible situations** than seen in real data and thus obtain a **high generalization of the algorithm**.

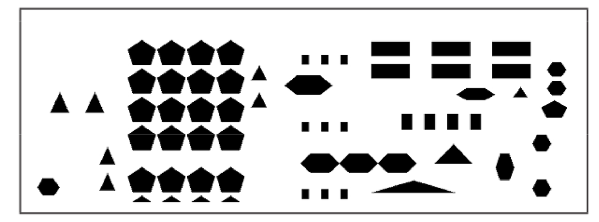
Model Compression

We reduce the image to a set of detected object types, their repetition scheme, and the background. More in detail: each object category is stored **as a set of codebook descriptors** and **vote vectors**, a **rectangular colorscale bitmap** resulting from averaging the image areas inside the detected elements bounding boxes, and a **simplified median color background**.

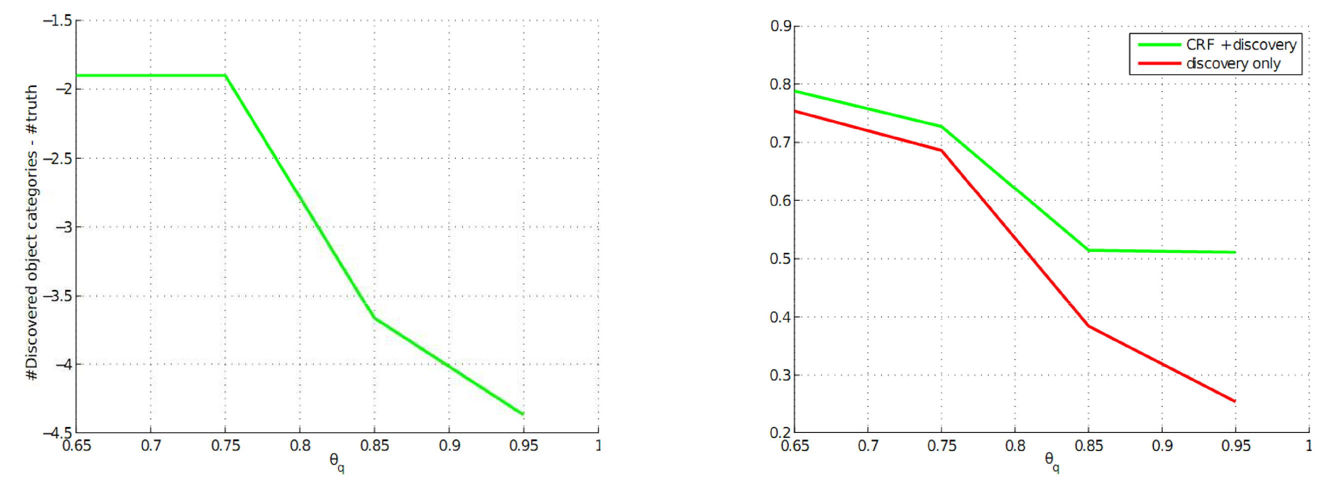
We consider that our approach **useful** in all those cases where the main goal is to **identify images** where repetitive patterns are present, although it is **not as well suited** to provide **detailed reconstructions** of the represented objects

Experiments



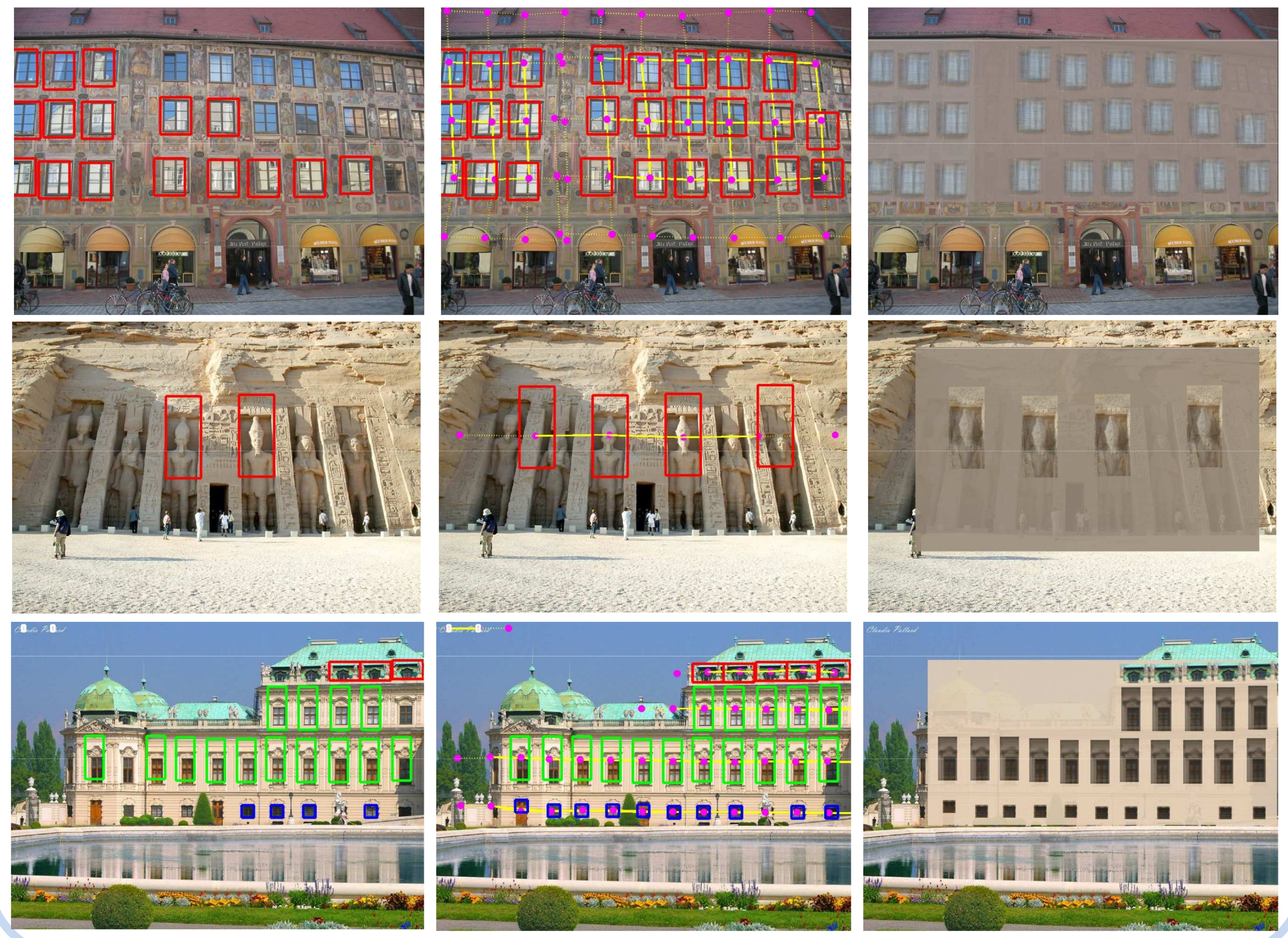


Left: Best lattice topology (triangular) used for training the CRF. **Right:** Sample from the evaluation data set used to assess the quality of object type distinction and grouping.



Left: Average difference between the **number of detected categories and annotated categories**. The algorithm tends to under-explain the data trying to not overfit single detections. **Right: Discovery only detection and discovery + CRF detection.** The contribution of CRF for detecting missing elements is particularly evident when a low detection rate is obtained. Graphs are plotted with respect to the minimum detection quality θb needed for each node.

Detected objects | CRF completed pattern | Model compression



Conclusions

We presented a **probabilistic technique** to discover and reason about **repetitive patterns of objects** in a **single image**.We introduced the concepts of **latticelets**, generalized building blocks of repetitive patterns. For high-level inference on the patterns, **CRFs** are used to soundly couple low-level detections with **high-level model information**. For the task of object detection by model prediction and completion, the experiments showed that the method is able to **significantly improve detection rate** by reinforcing weak detection hypotheses. For the task of model compression, a very high **compression ratio of up to 98%** with respect to the raw image has been achieved. We see applications of this method in **image inpainting, environment modeling** of urban scenes and robot navigation in **man-made buildings**.

ACKNOWLEDGEMENTS: This work has been party supported by German Research Foundation (**DFG**) under contract number ***SFB/TR-8*** and **EU Project *EUROPA-FP7-231888***.