

Visual Localization Across Seasons using Image Sequences

Cyrill Stachniss

Joint work with Tayyab Naseer,
Luciano Spinello, Wolfram Burgard

Place Recognition Across Seasons



Is this the same place?

Place Recognition Across Seasons



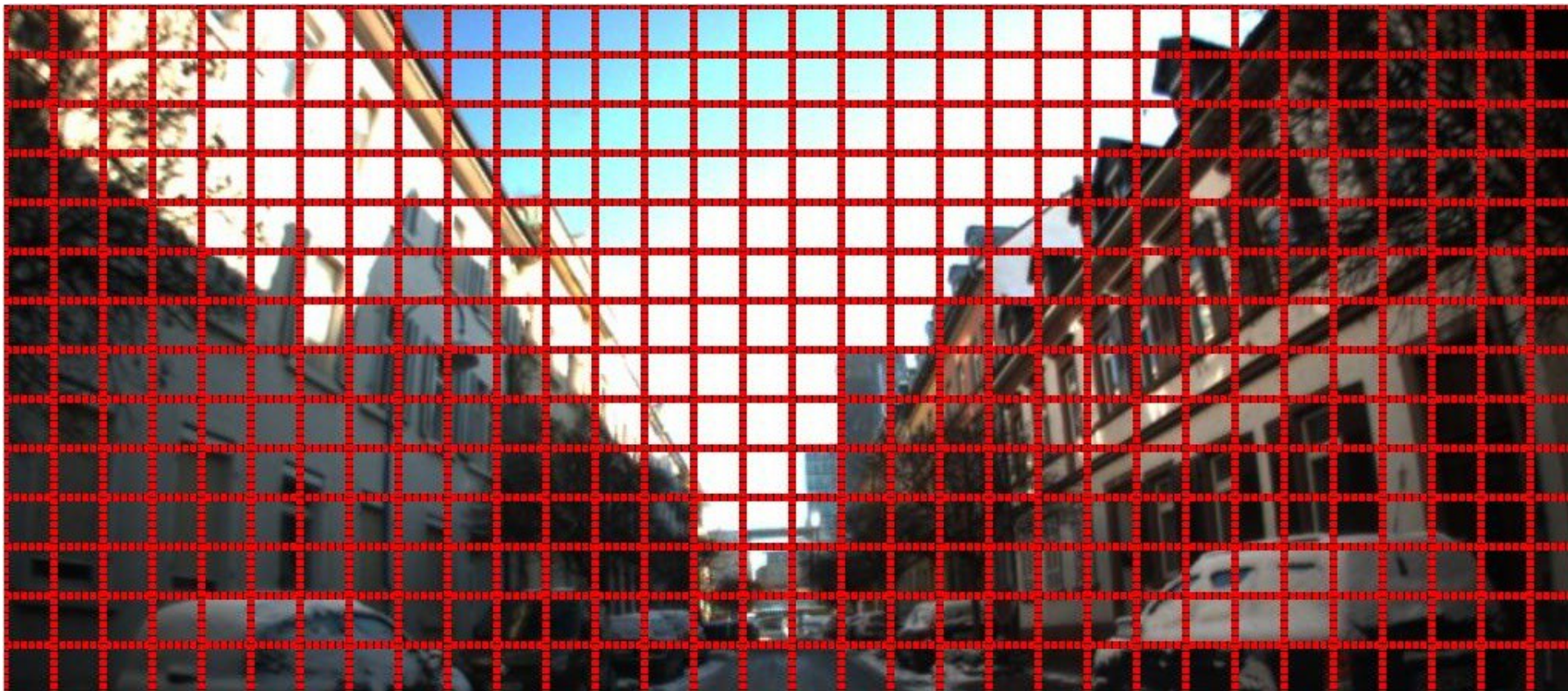
Is this the same place?

Matching SIFT Features Fails

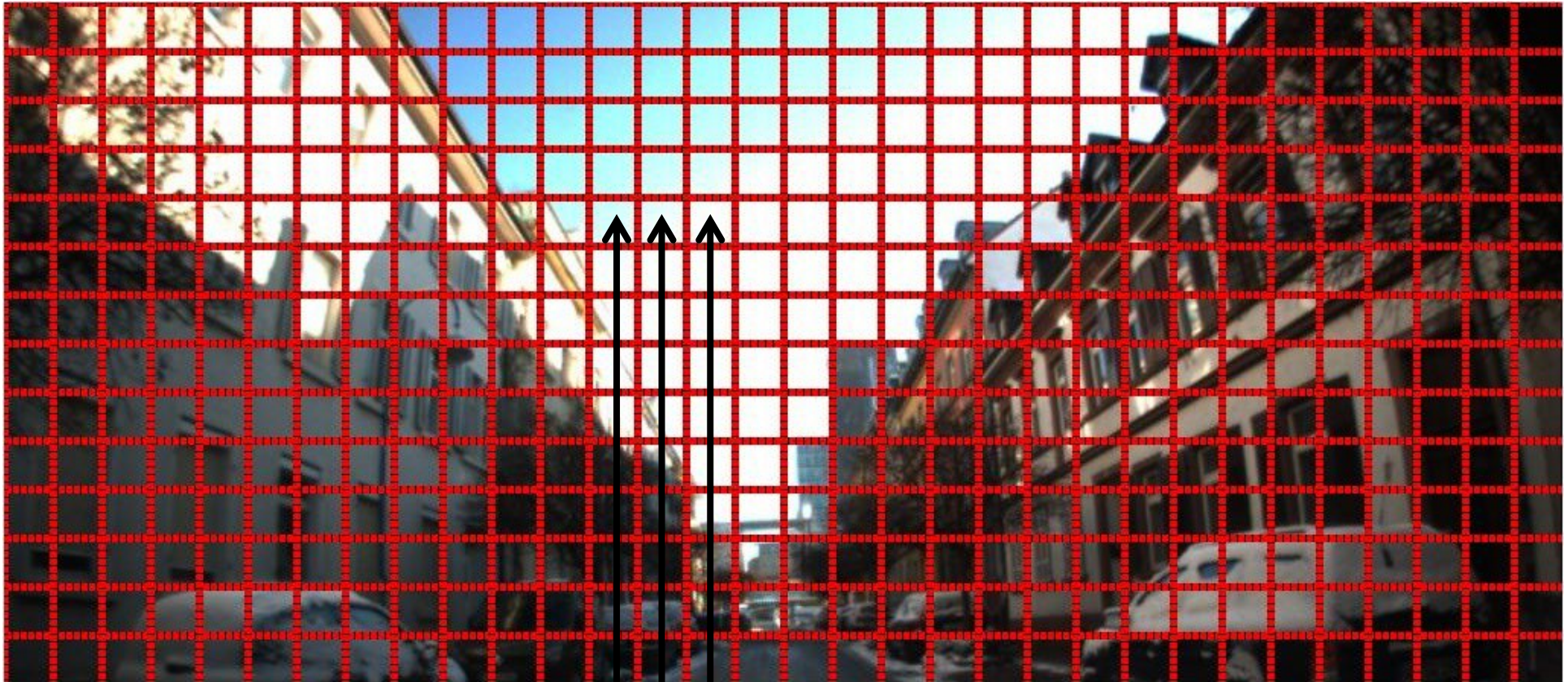


not well suited for across season matching

Global Image Descriptor

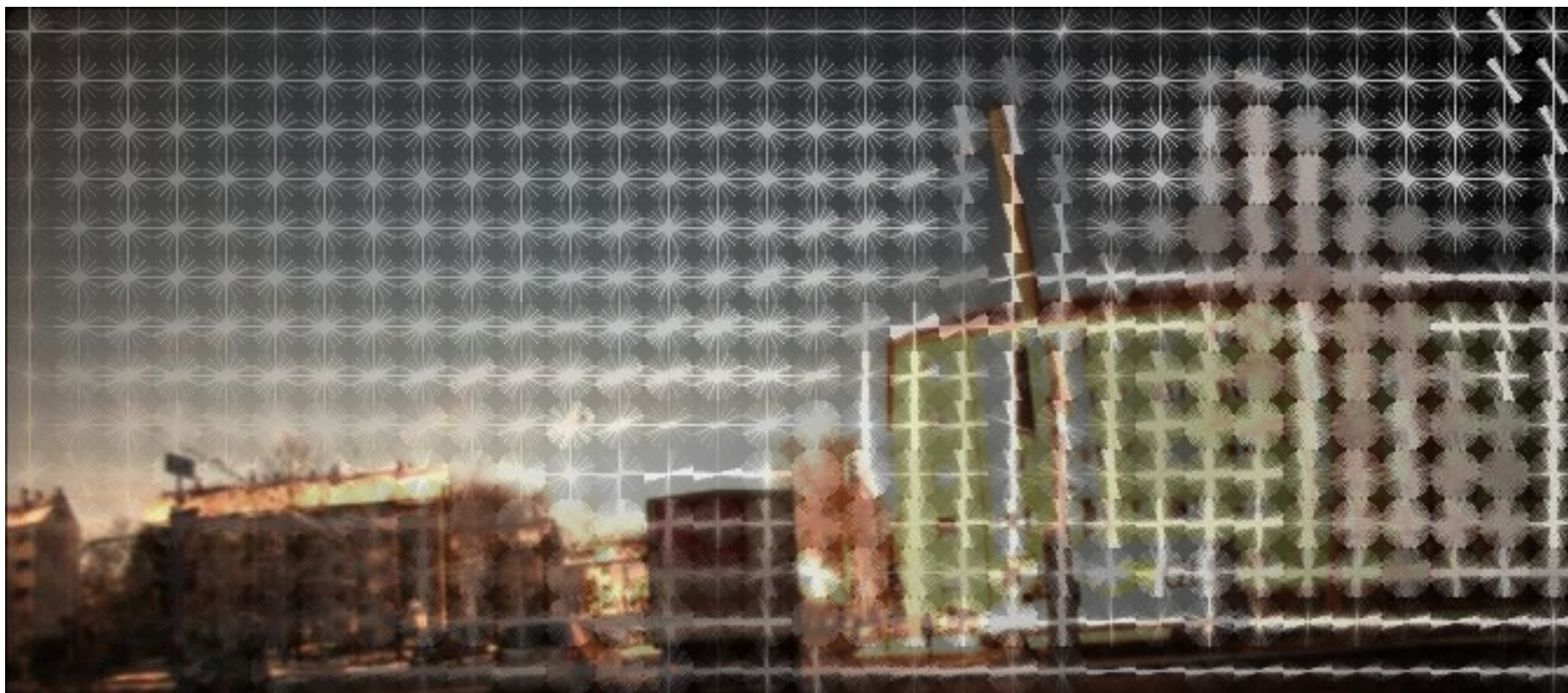


Global Image Descriptor



global descriptor for each cell

Global Image Descriptor



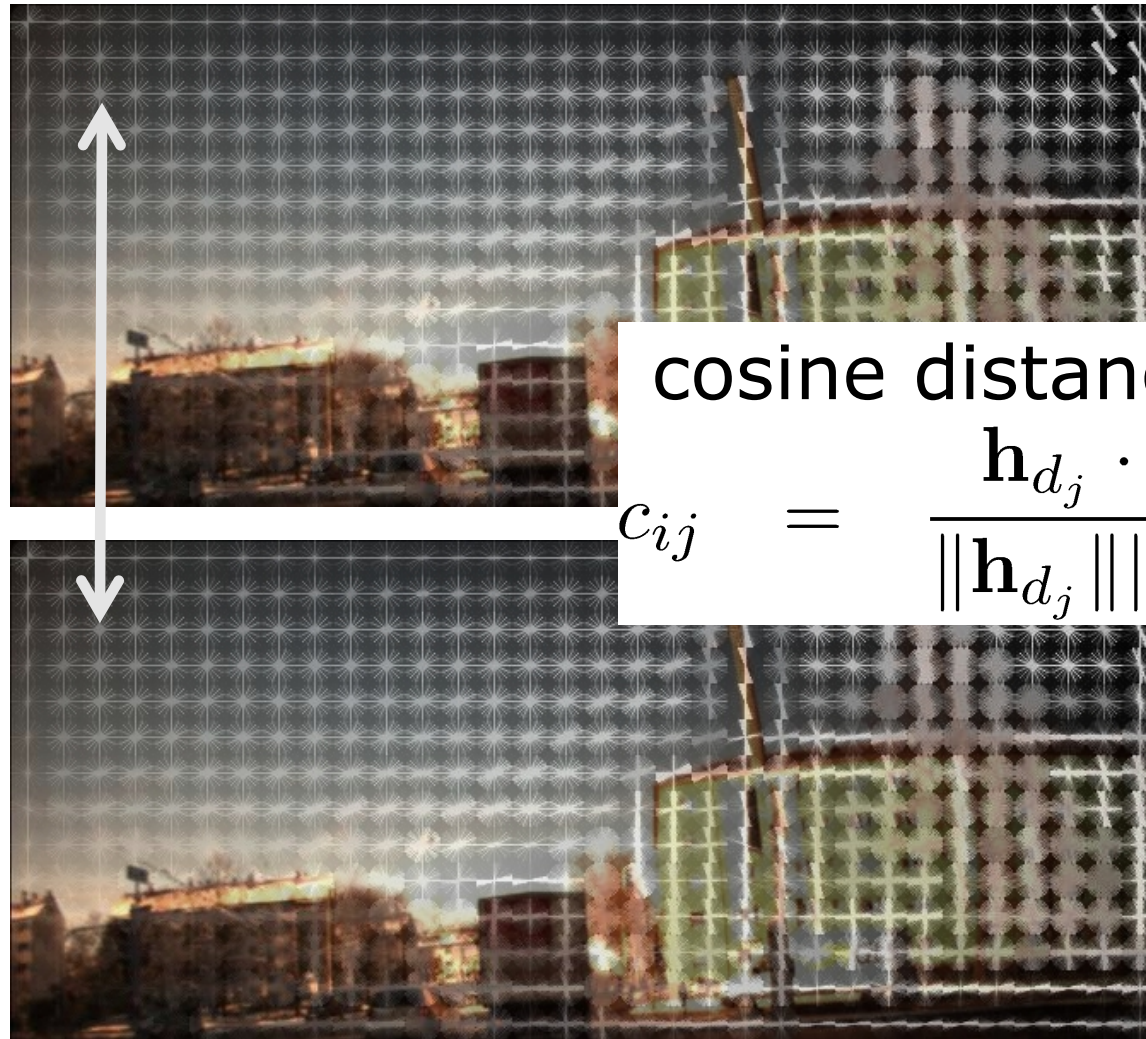
 h_i

Why HOG?

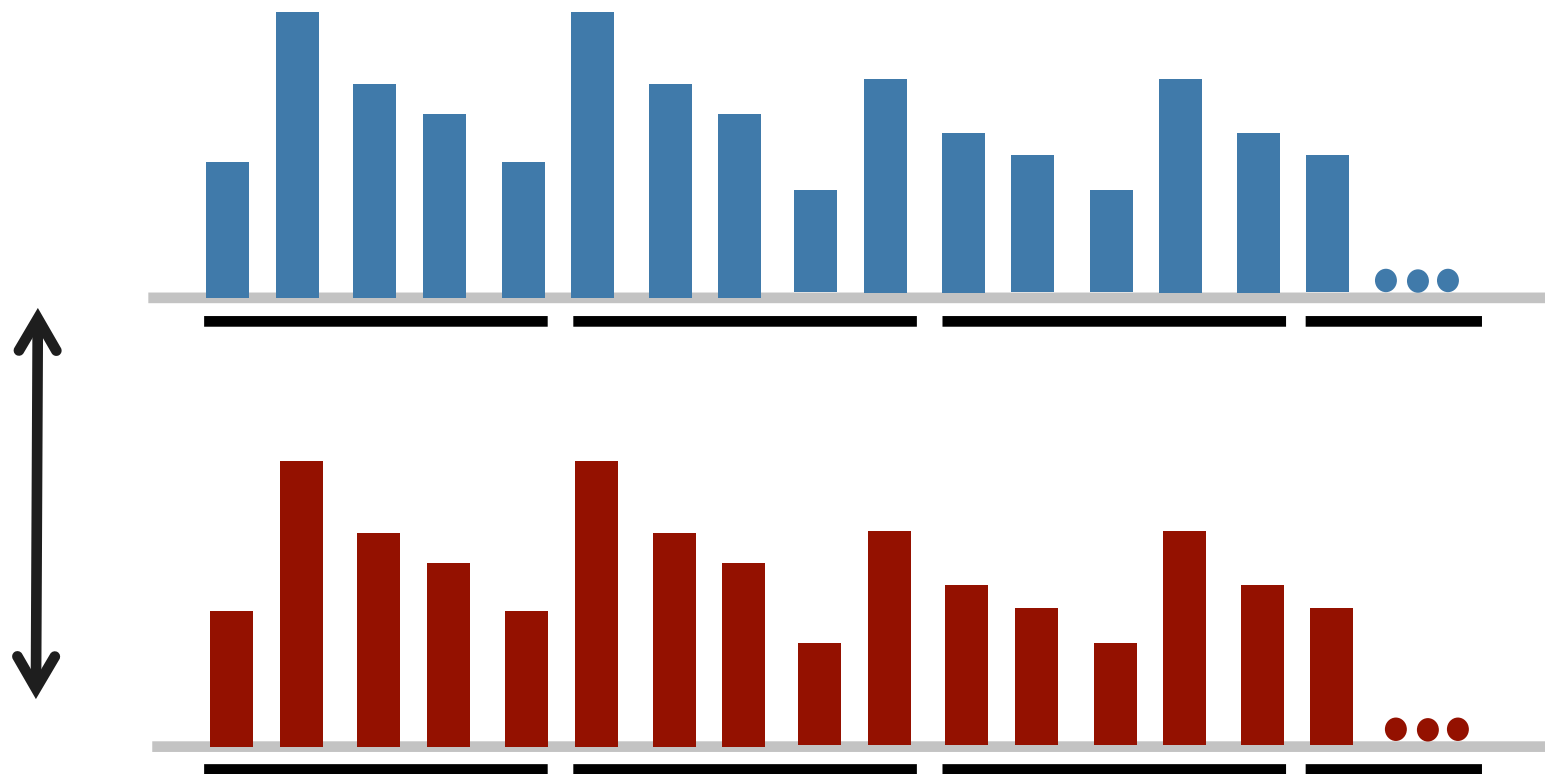
- Collects gradient orientation and magnitude in a cell
- Irrespectively from where exactly they are found in the cell

Provides a certain degree of robustness to view point changes

Comparing Two Images

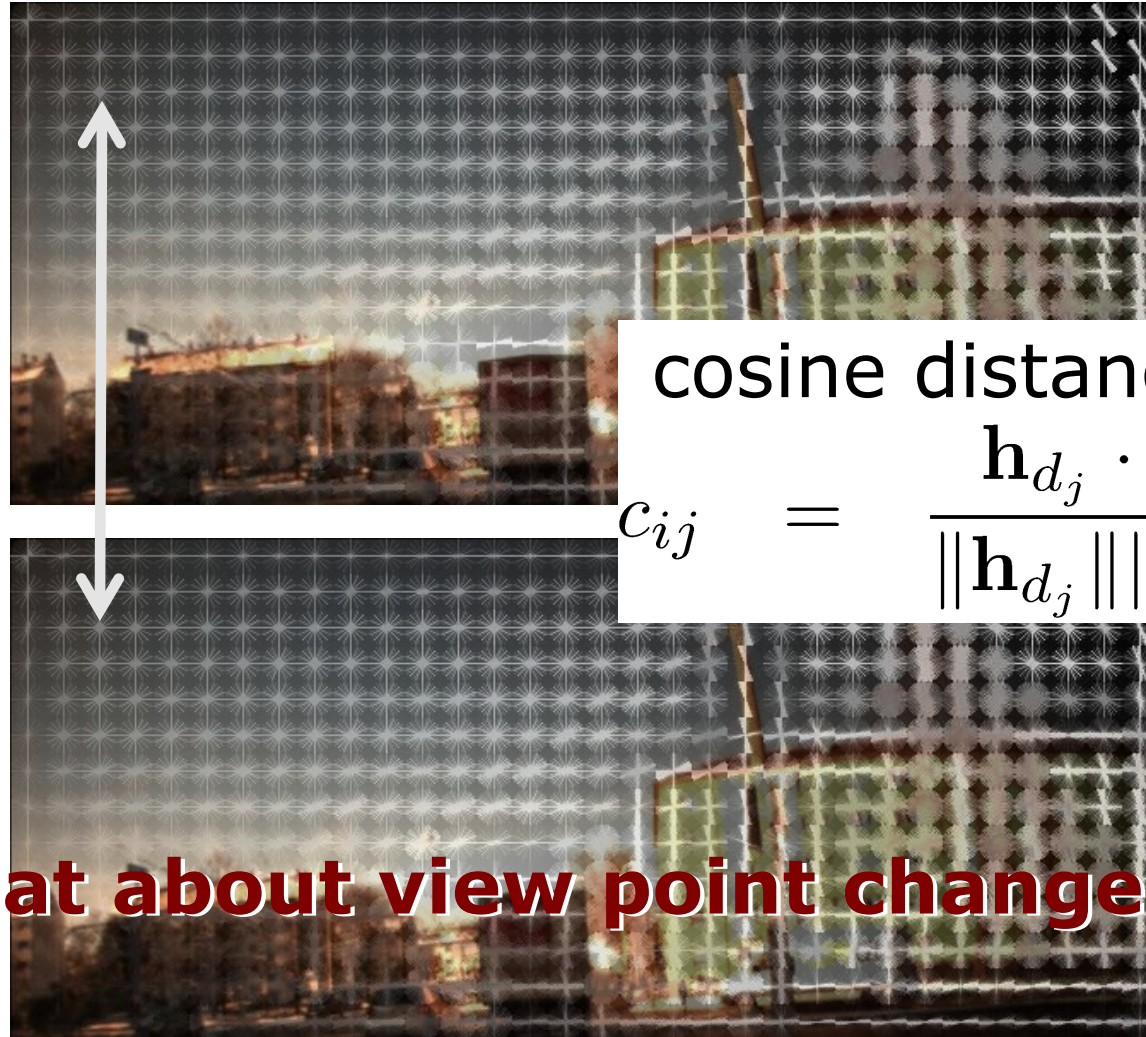


Comparing Two Images



$$c_{ij} = \frac{\mathbf{h}_{d_j} \cdot \mathbf{h}_{q_i}}{\|\mathbf{h}_{d_j}\| \|\mathbf{h}_{q_i}\|}$$

Comparing Two Images



cosine distance

$$c_{ij} = \frac{\mathbf{h}_{d_j} \cdot \mathbf{h}_{q_i}}{\|\mathbf{h}_{d_j}\| \|\mathbf{h}_{q_i}\|}$$

What about view point changes?

Image Shifts for Handling Varying Viewpoints

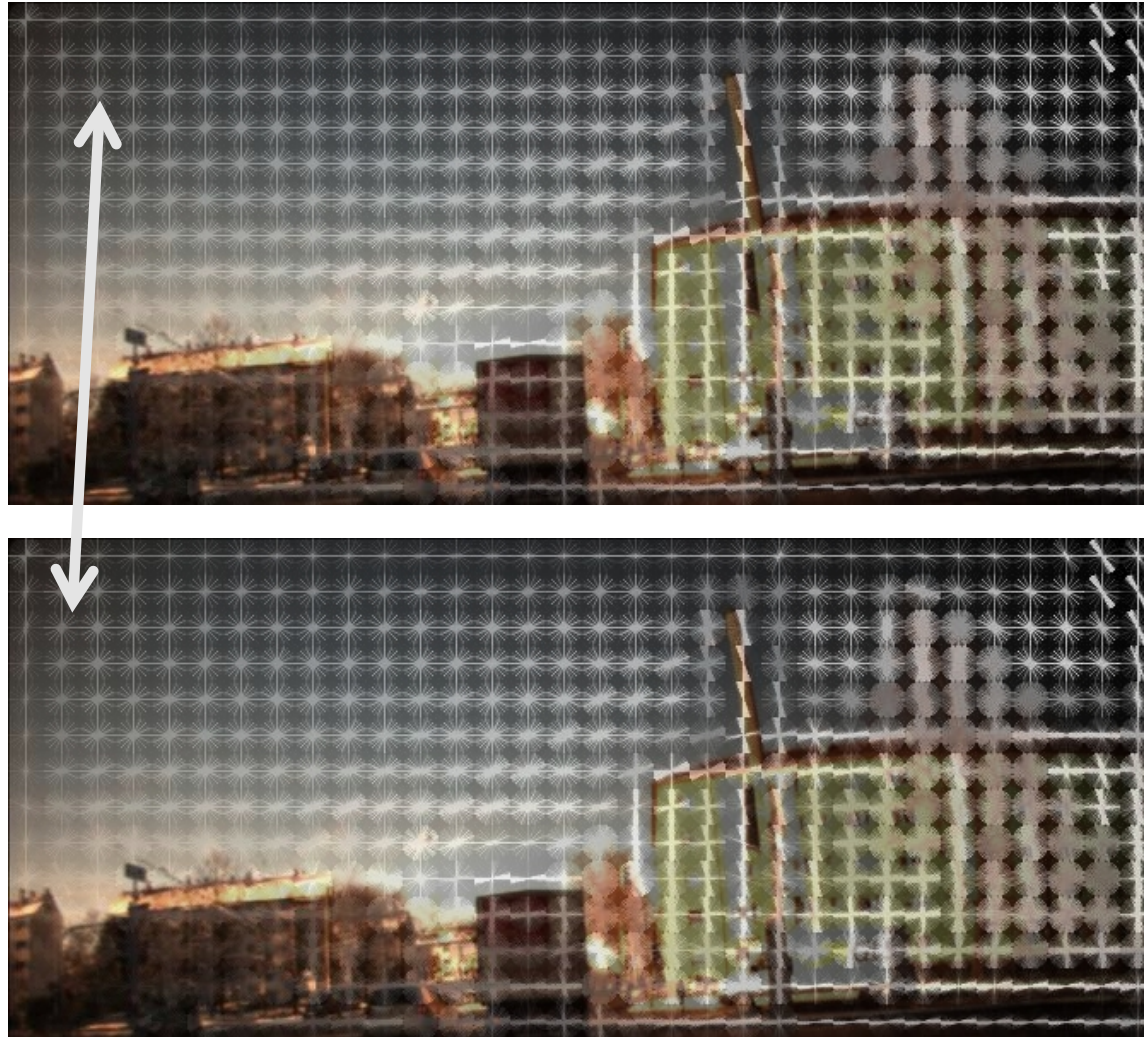


Image Shifts for Handling Varying Viewpoints

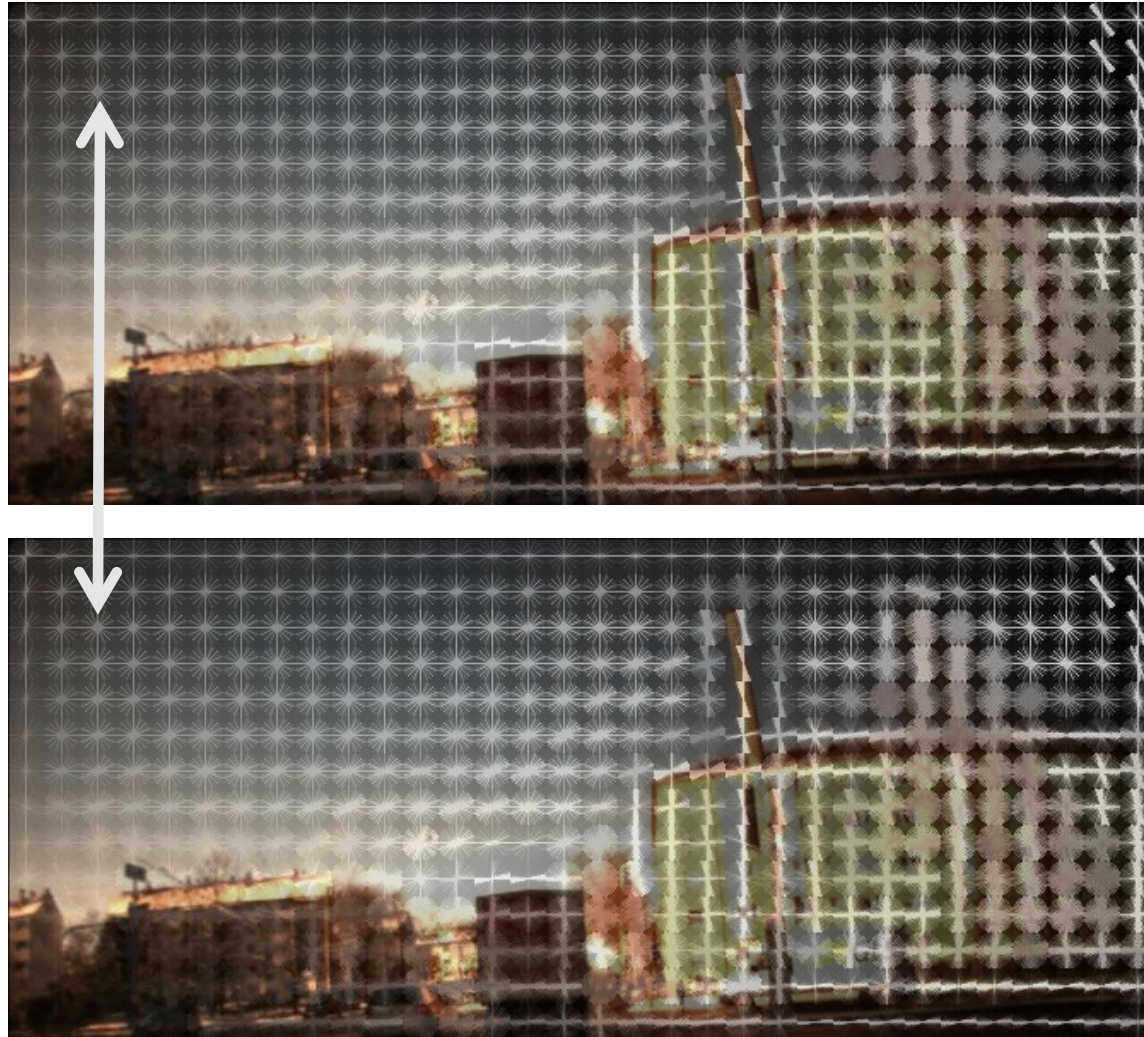


Image Shifts for Handling Varying Viewpoints

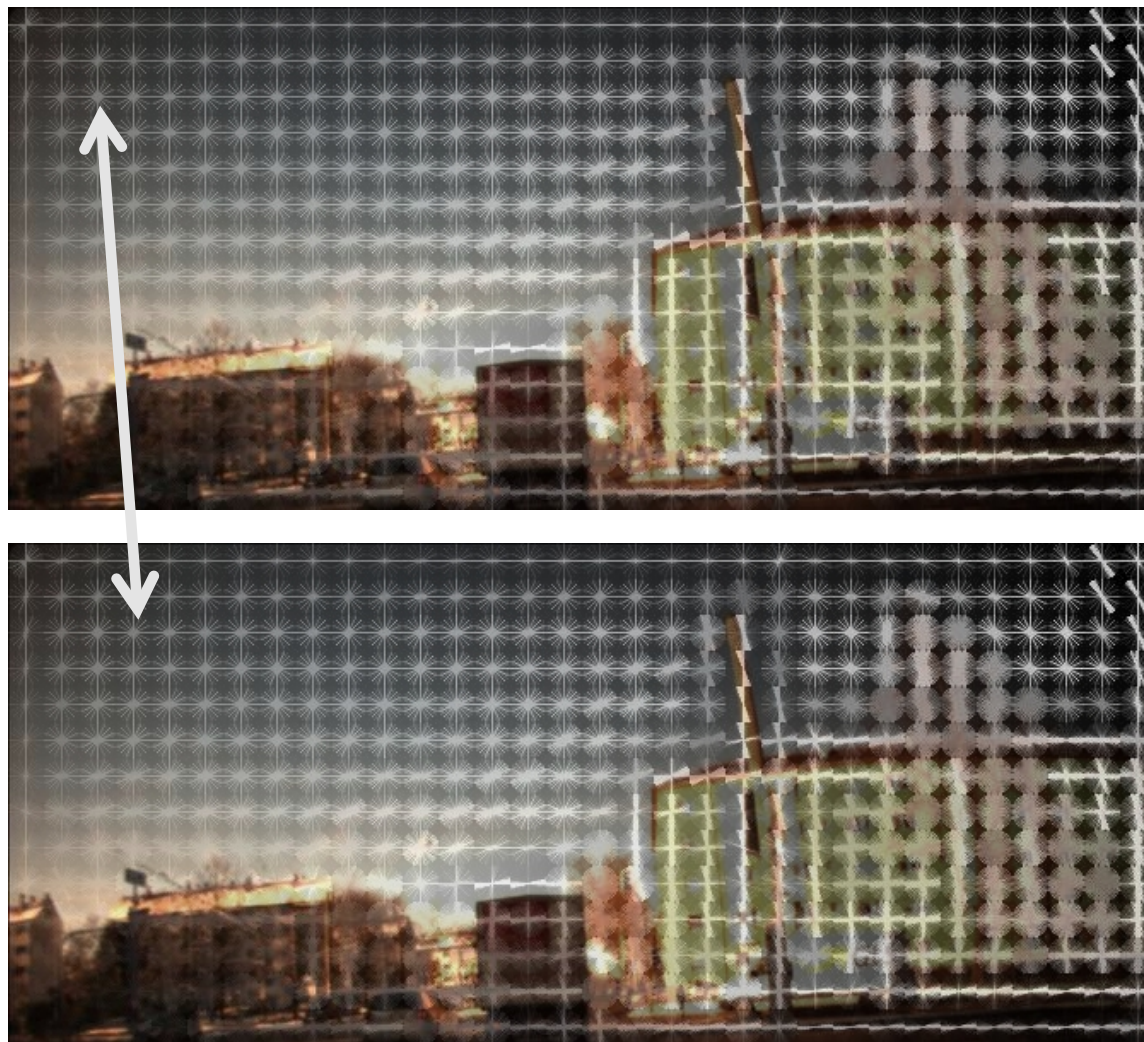


Image Shifts for Handling Varying Viewpoints

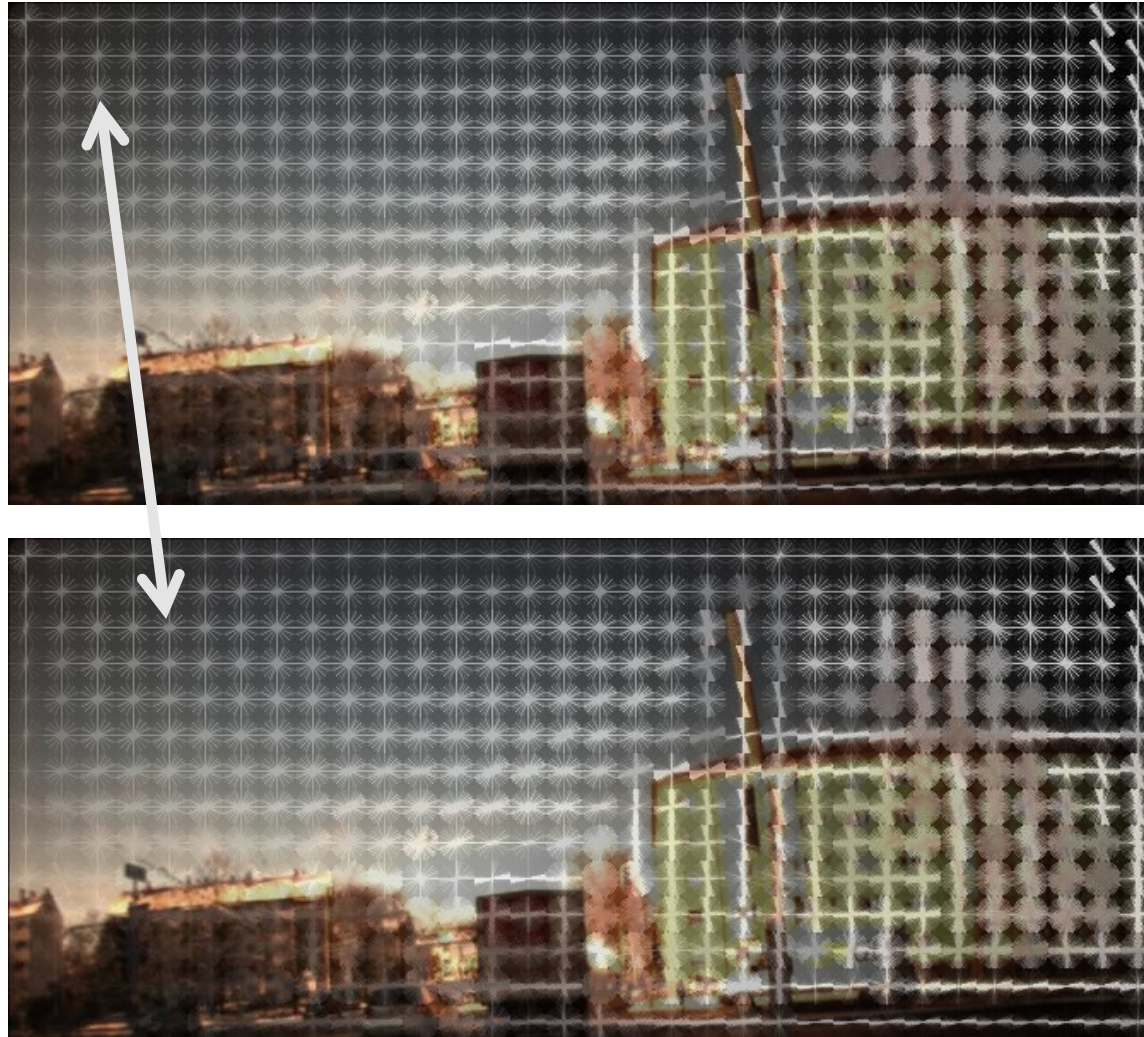


Image Shifts for Handling Varying Viewpoints

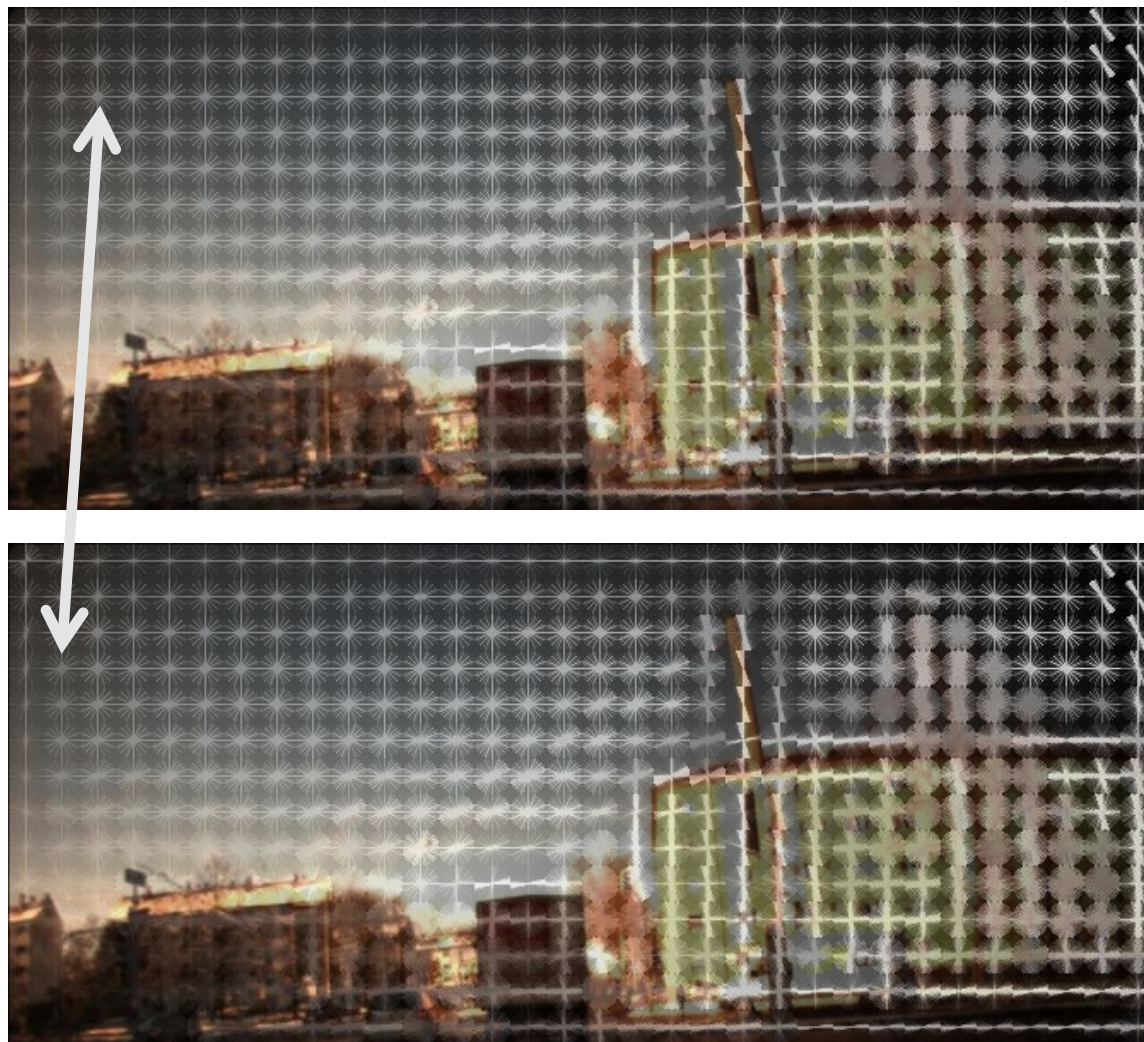


Image Shifts for Handling Varying Viewpoints

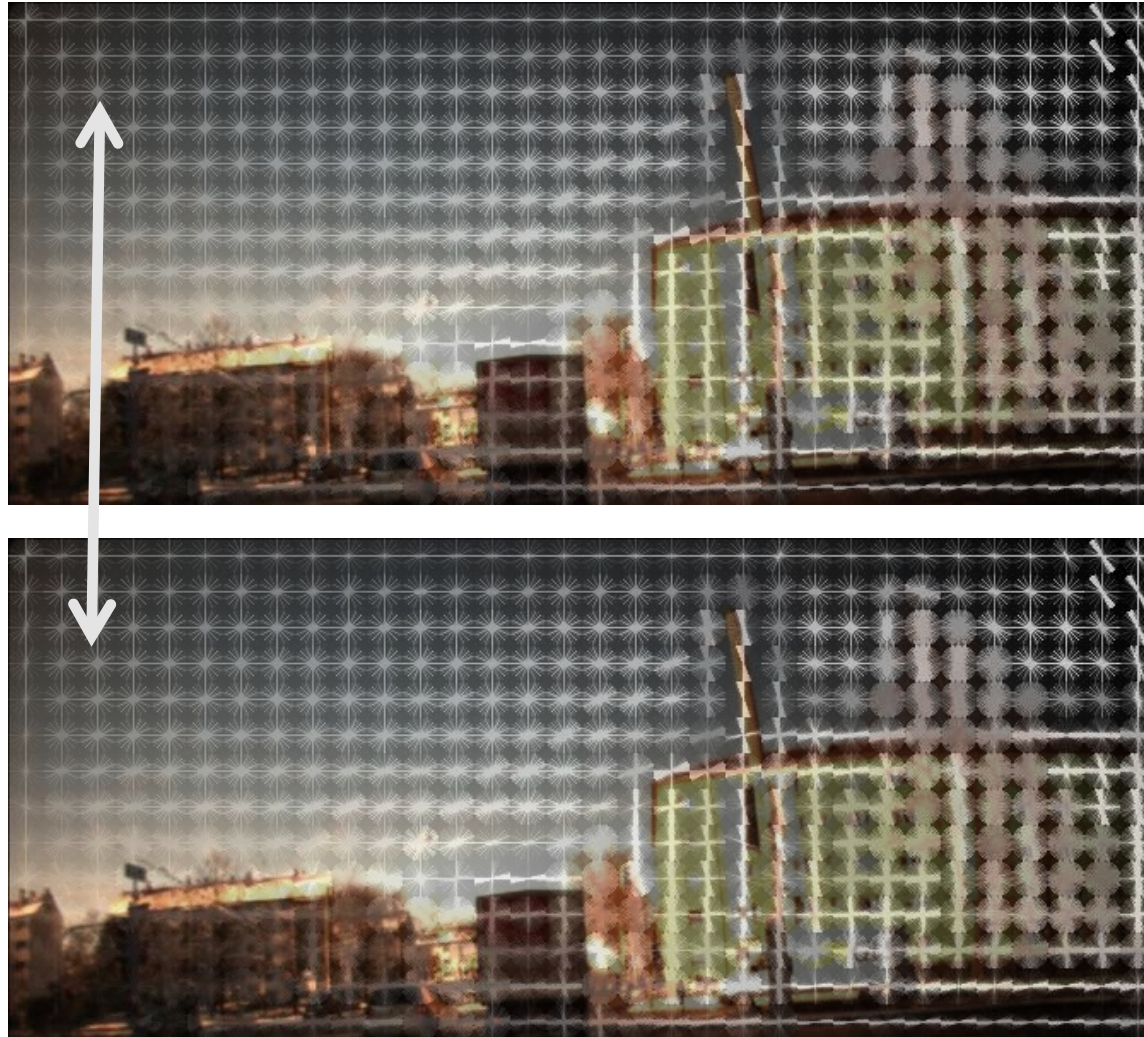


Image Shifts for Handling Varying Viewpoints

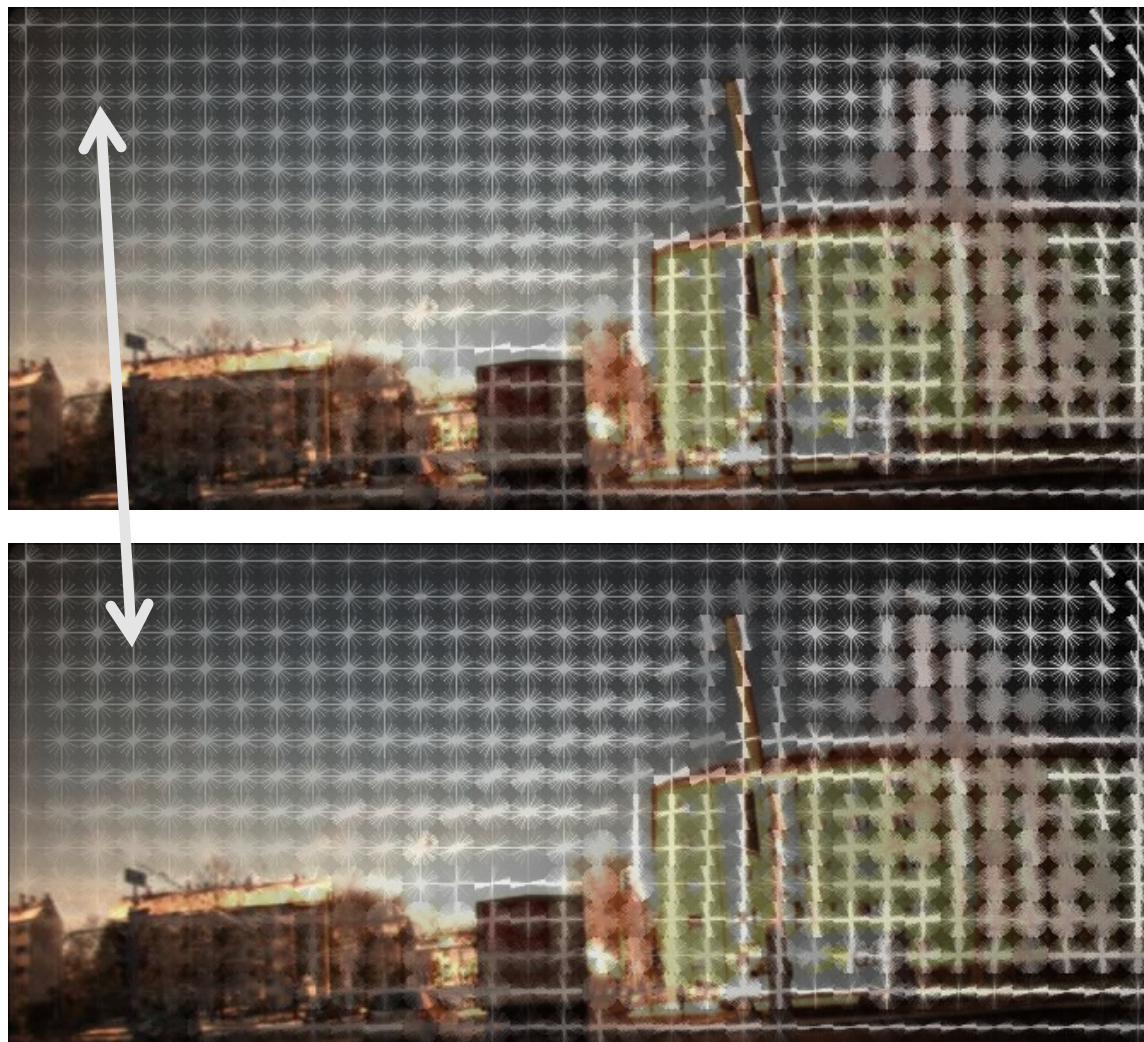


Image Shifts for Handling Varying Viewpoints

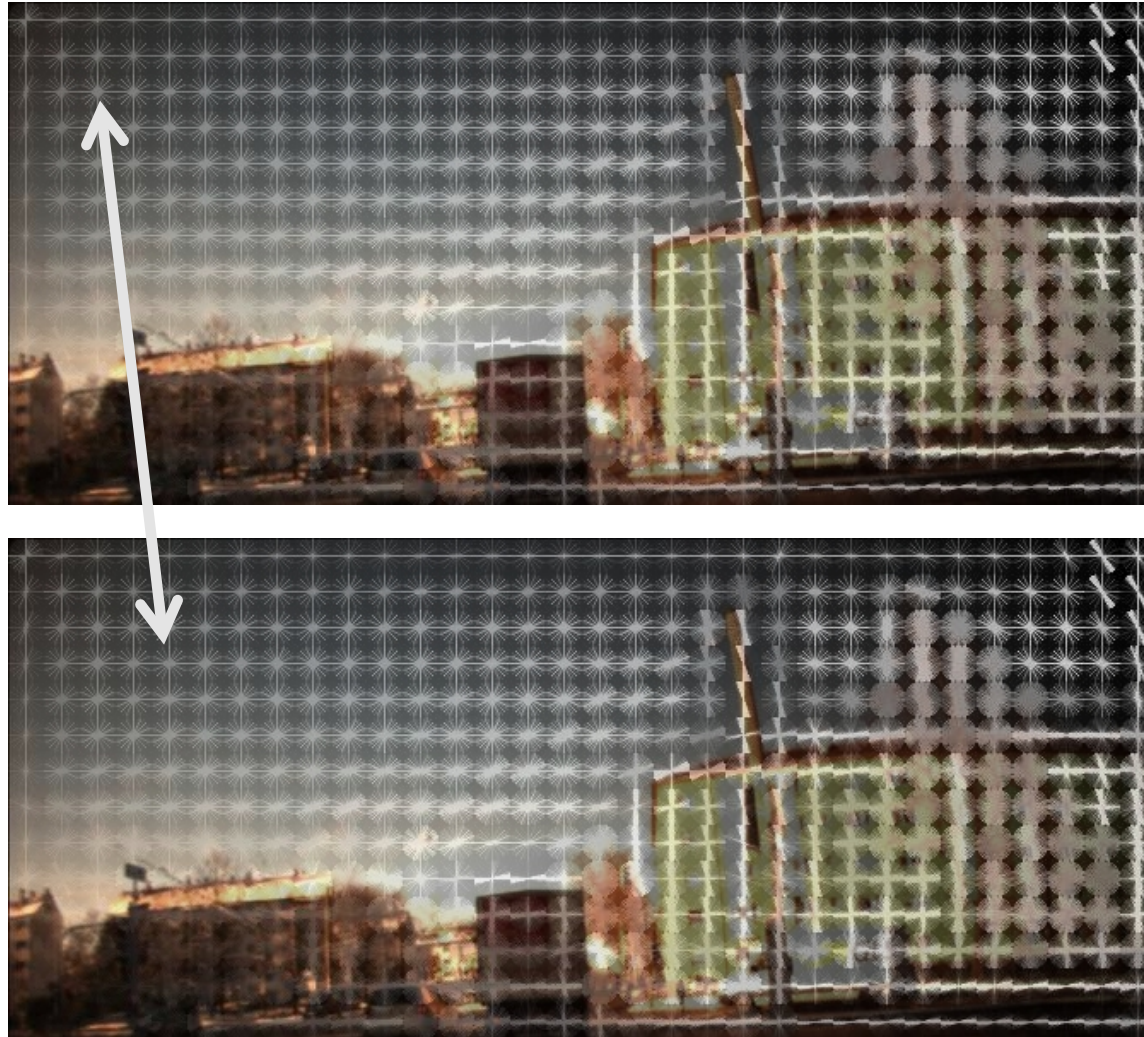
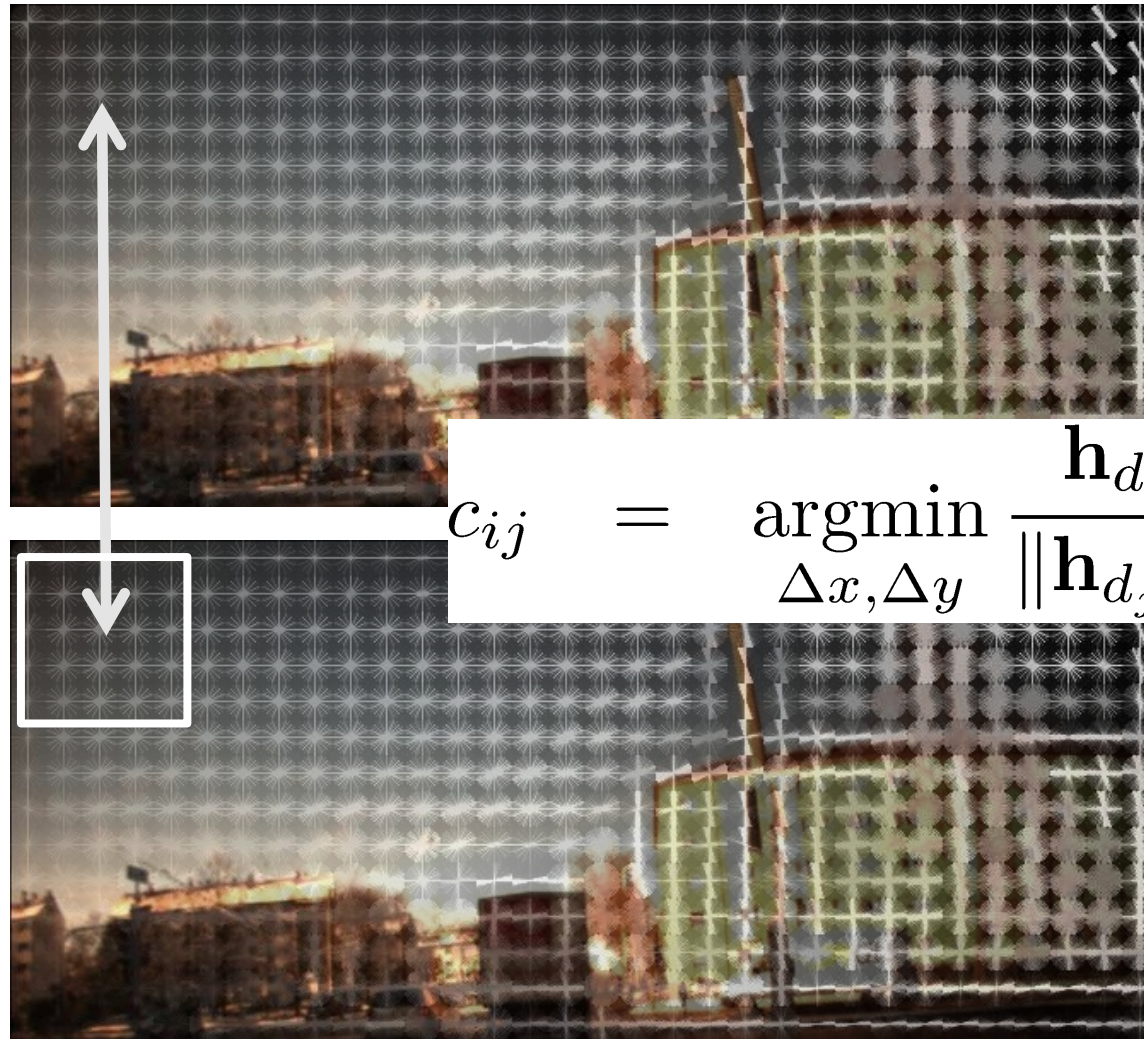


Image Shifts for Handling Varying Viewpoints



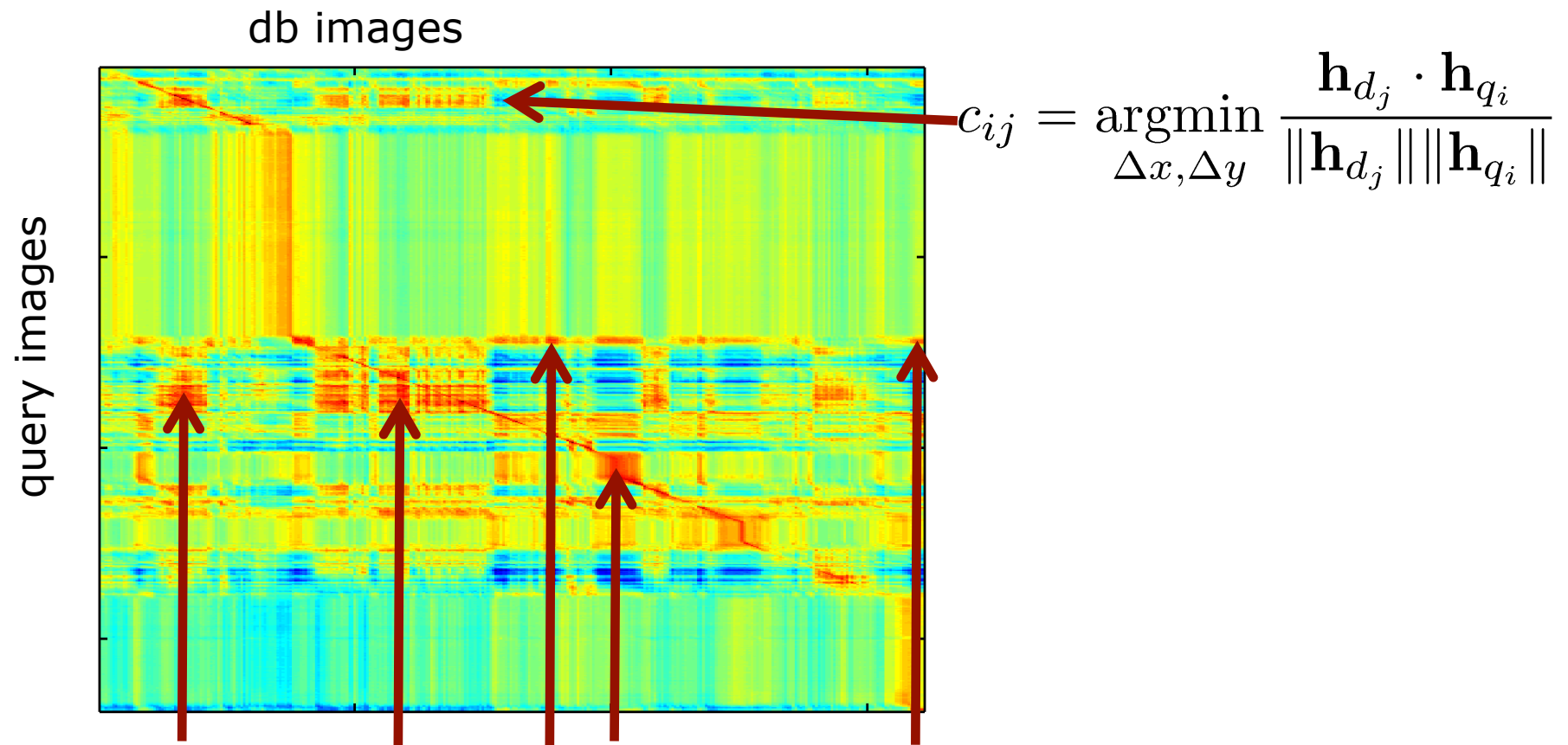
$$C_{ij} = \operatorname{argmin}_{\Delta x, \Delta y} \frac{\mathbf{h}_{d_j} \cdot \mathbf{h}_{q_i}}{\|\mathbf{h}_{d_j}\| \|\mathbf{h}_{q_i}\|}$$

Really, that's it?

No, We Are Not There Yet!

- Matching is only **improved, not solved**
- Individual images pairs often show a rather small similarity score
- Several images yield similar matching score

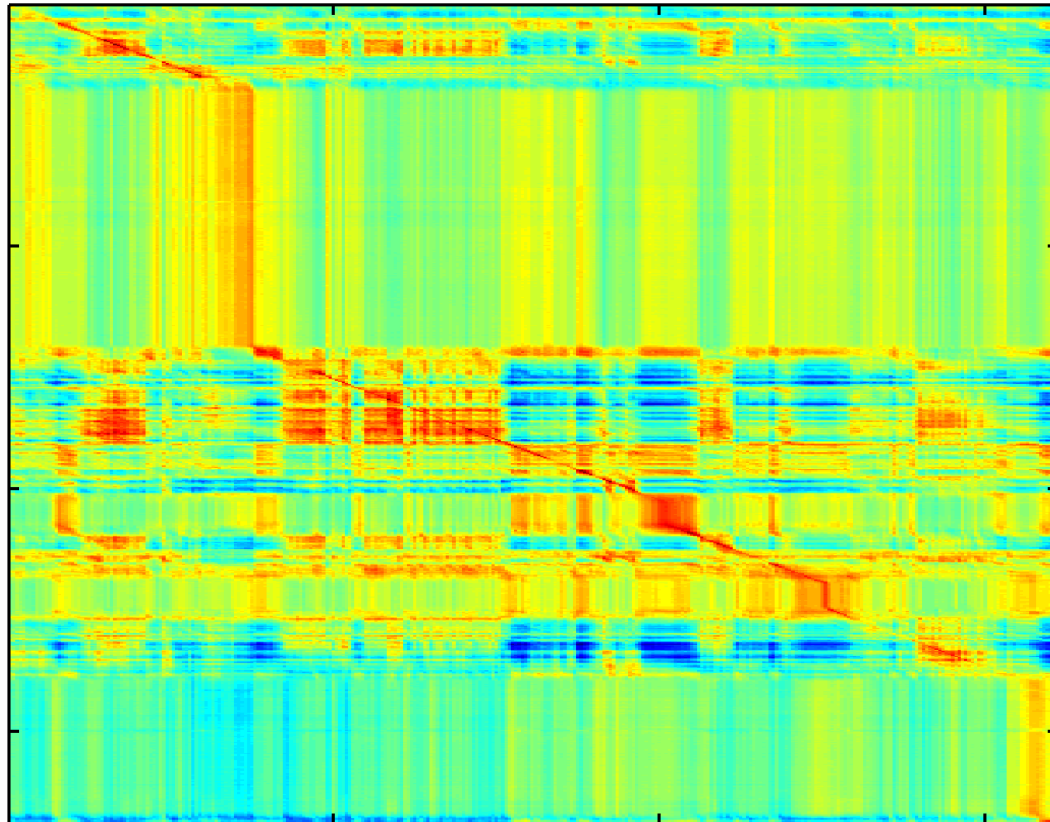
Pairwise Matching Matrix



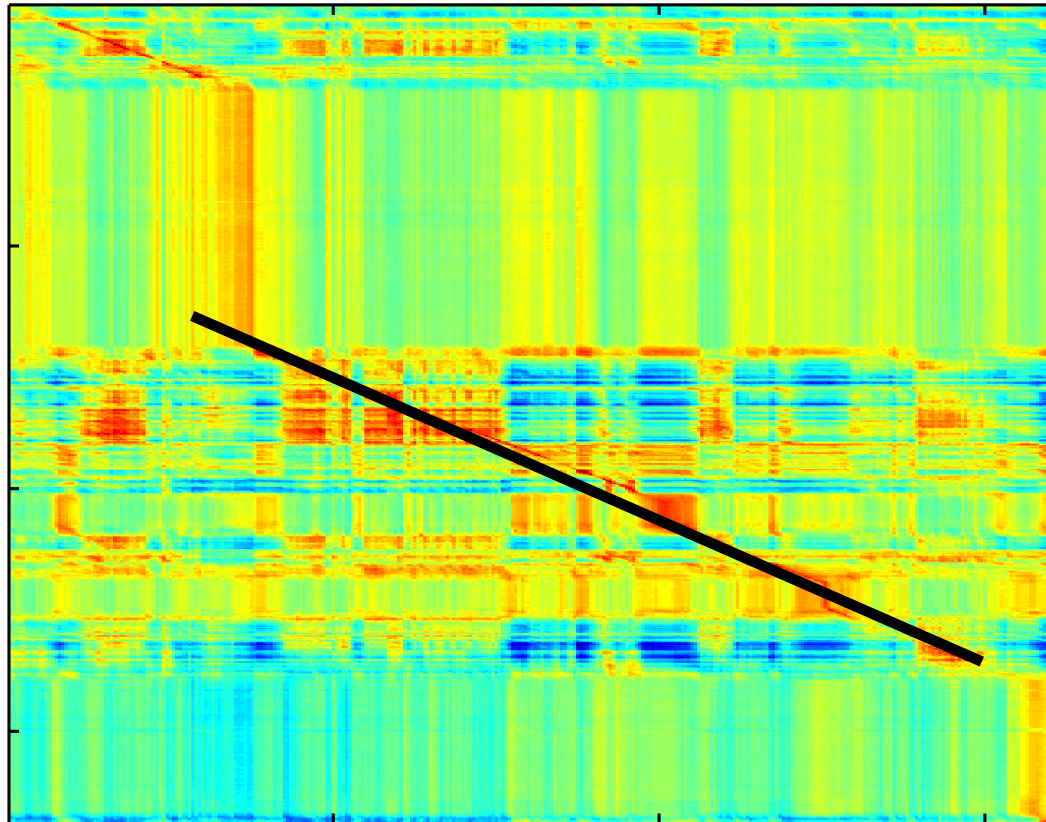
Just picking the best match does not work well

Matching Sequences

- We can improve matching if we consider it as a **sequence** of images



Sequence Matching by Fitting a Straight Line Through the Matrix

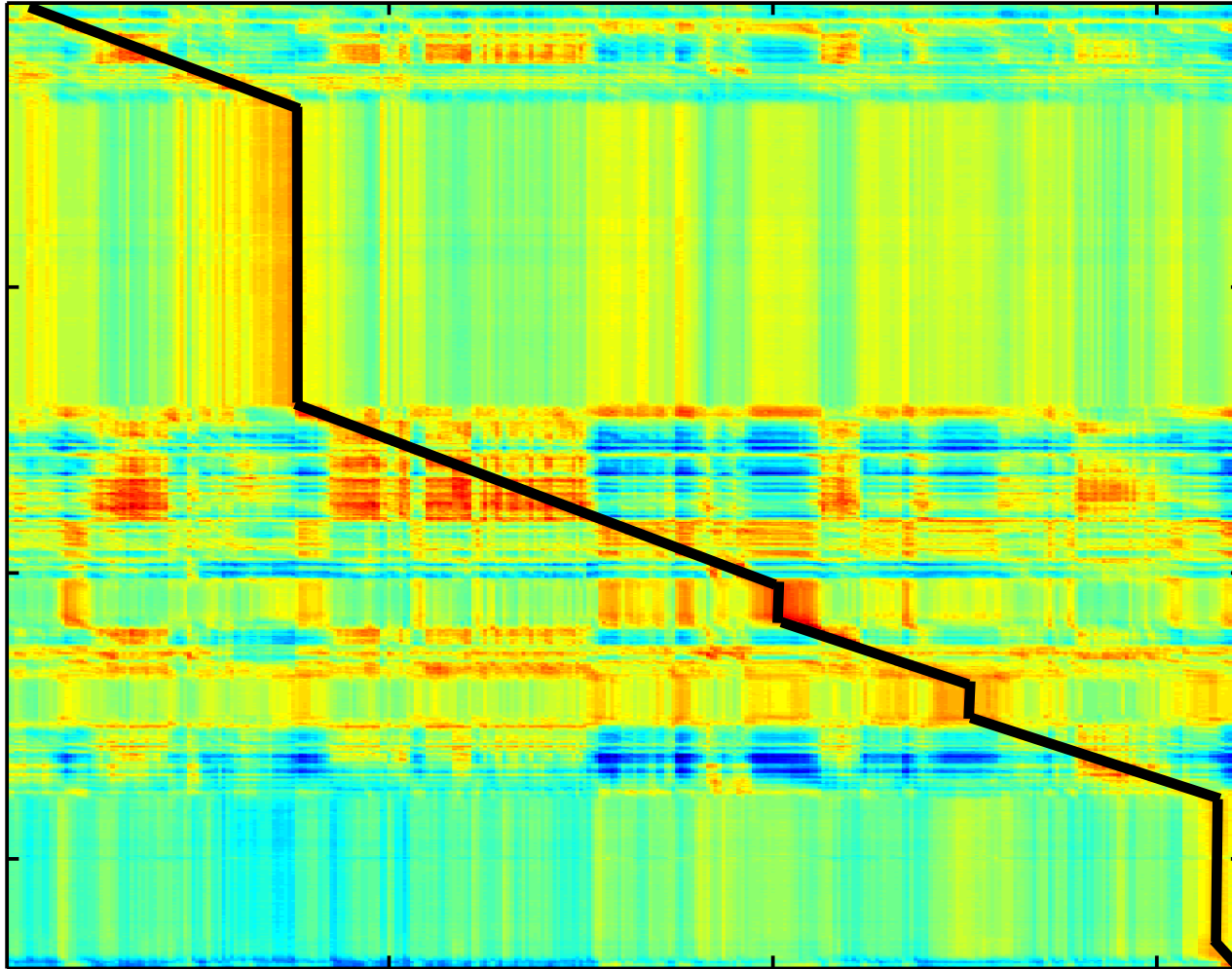


Can we do better?

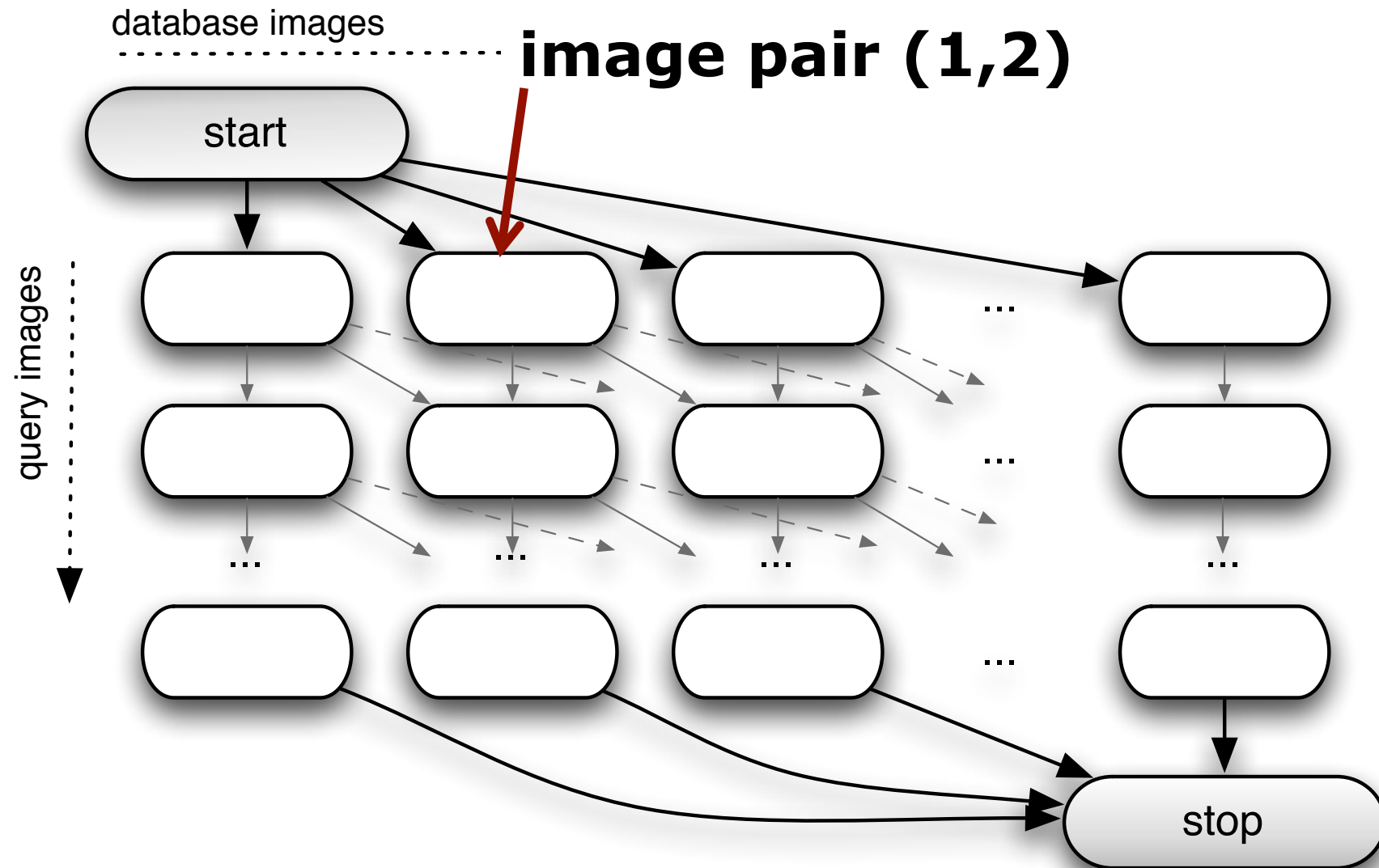
Can we do better?

Yes, we can!

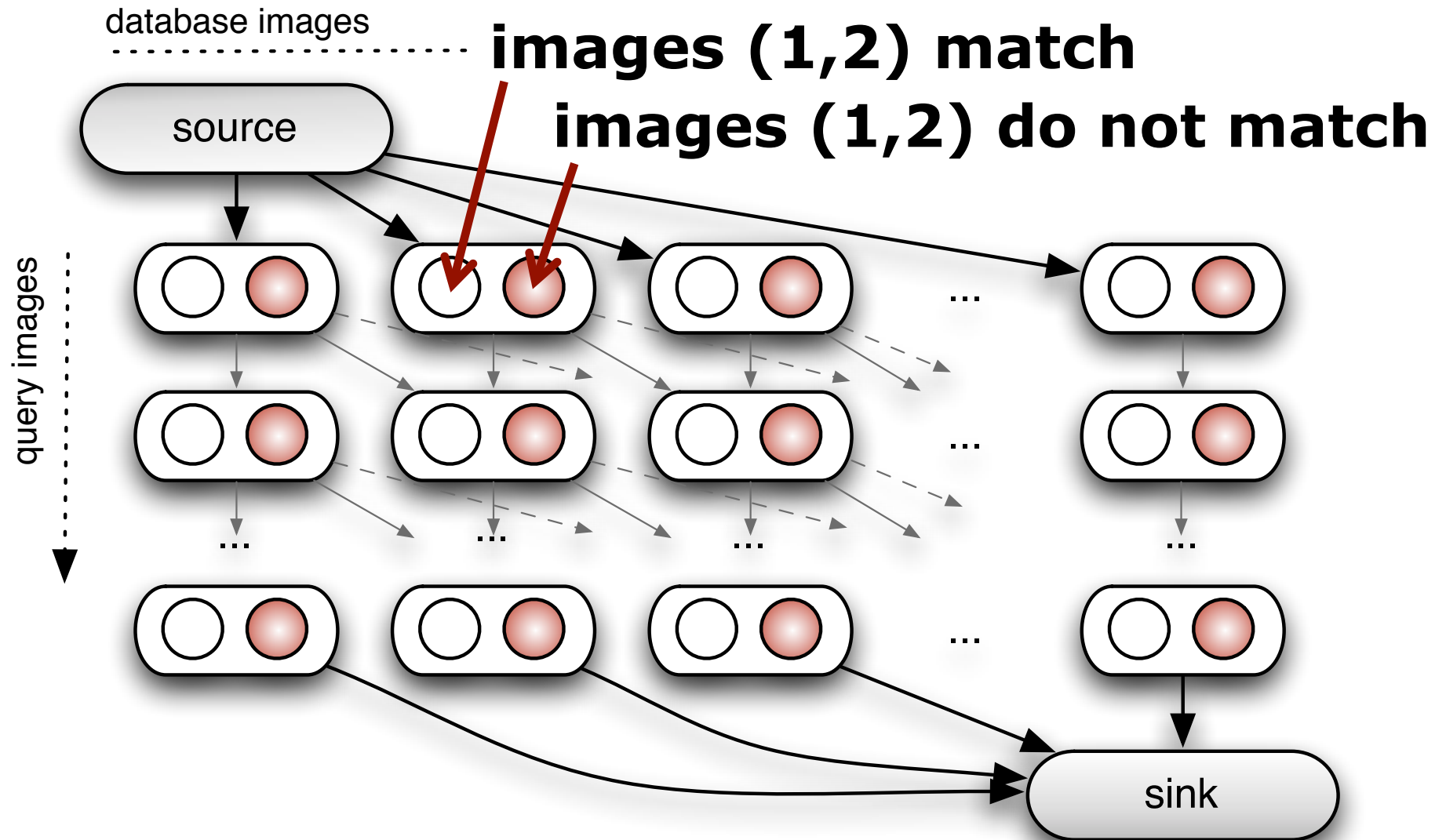
Yes, We Can



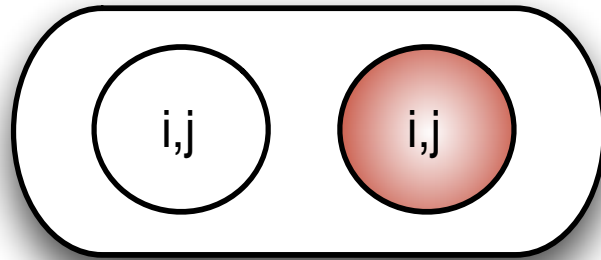
Data Association Graph



Data Association Graph

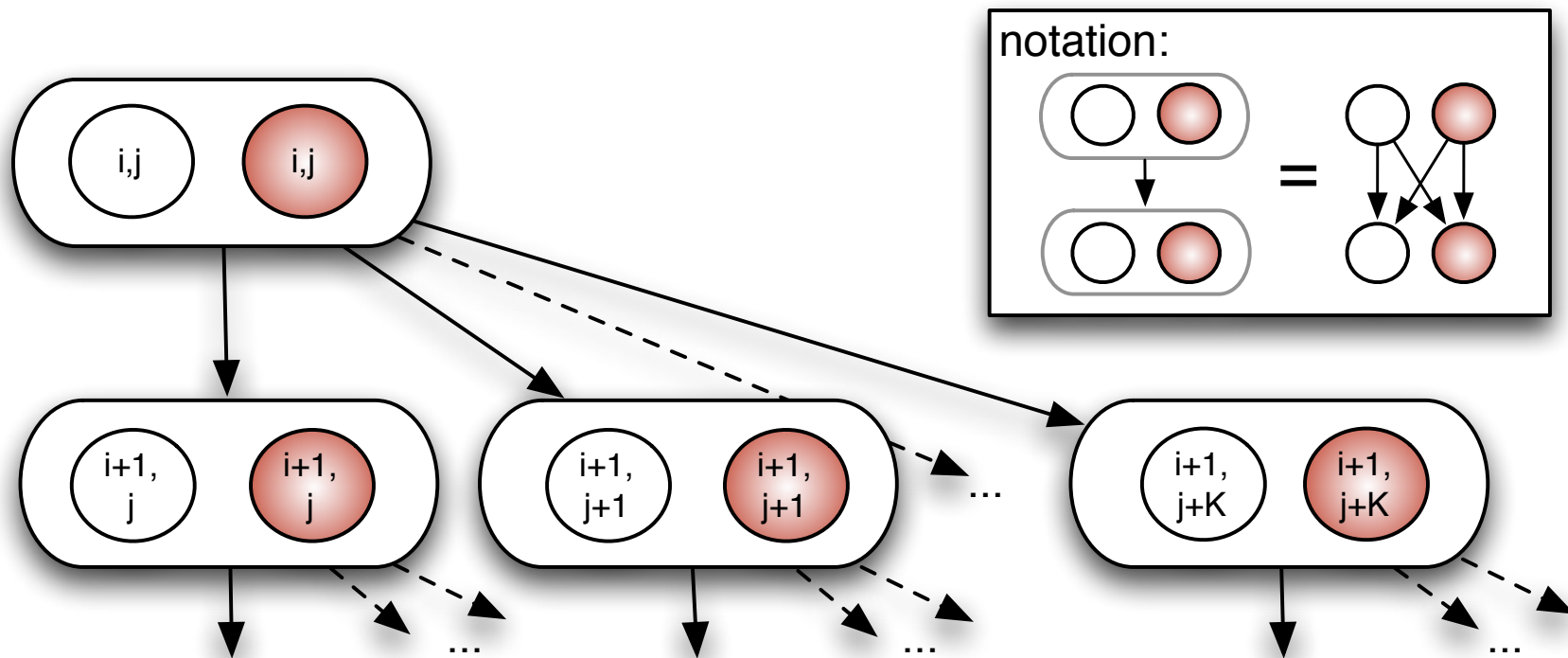


The Nodes



- **White matching node** models a successfully match of an image pair
- **Red hidden node** describes that the pair of images could not be matched
- Red nodes are needed to maintain the sequence

Data Association Graph (Magnified View)

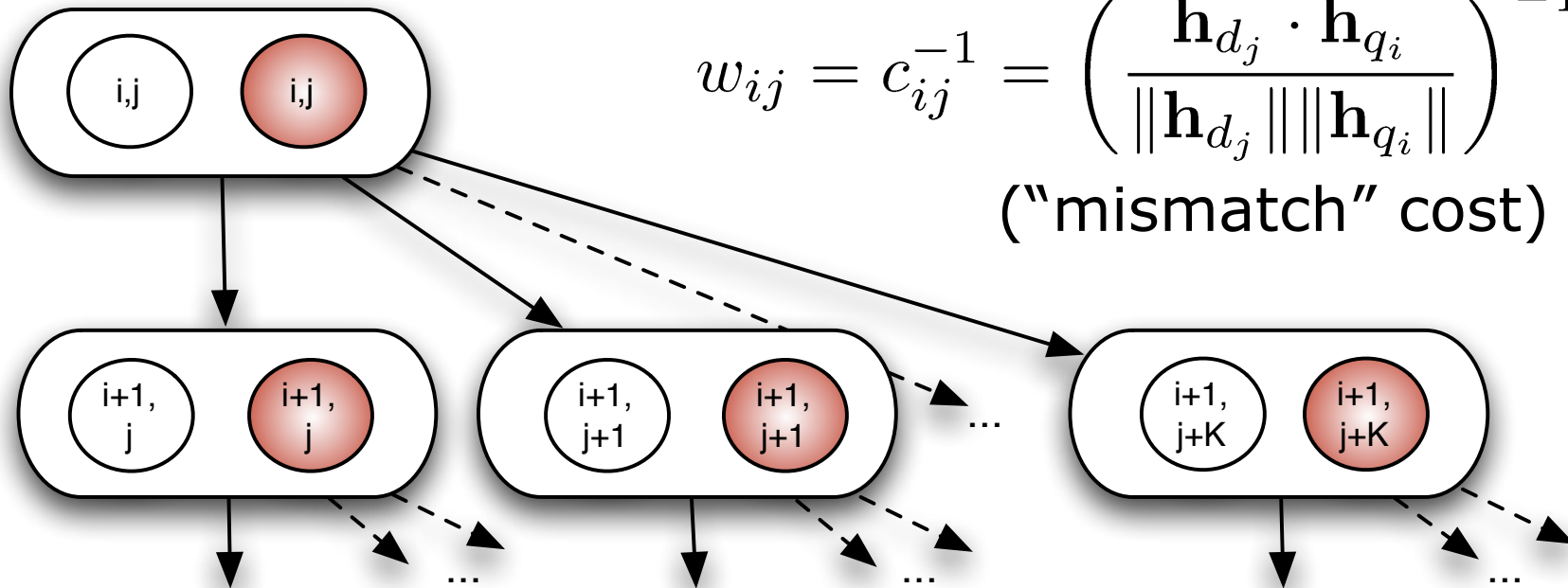


Edge Cost

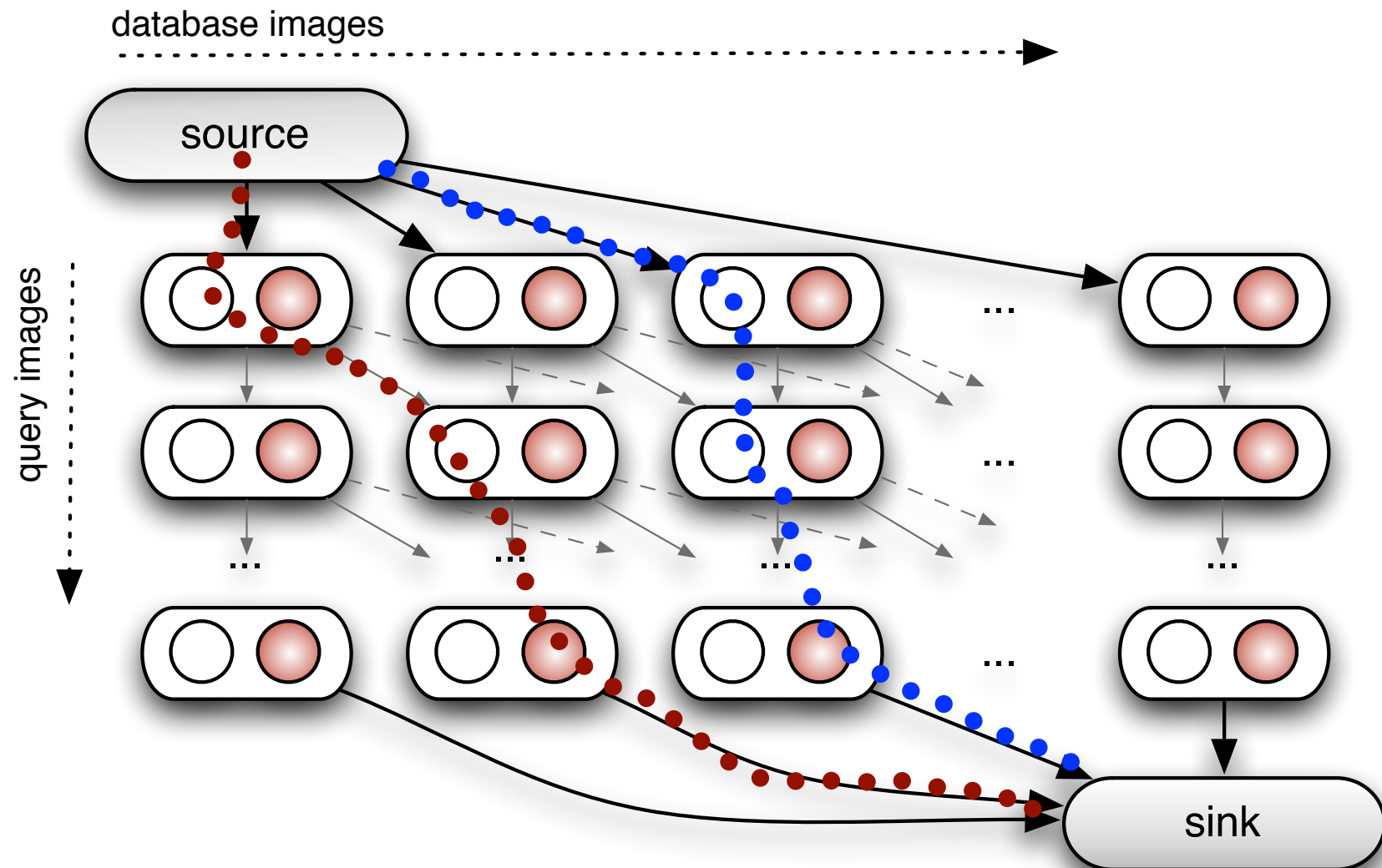
edge cost to node ij :

$$w_{ij} = c_{ij}^{-1} = \left(\frac{\mathbf{h}_{d_j} \cdot \mathbf{h}_{q_i}}{\|\mathbf{h}_{d_j}\| \|\mathbf{h}_{q_i}\|} \right)^{-1}$$

("mismatch" cost)



Flow Network



Network Flow

- Network flow parameters
 - Capacity=1 for edges to matching nodes
 - Flow $F > 1$

Network Flow

- Network flow parameters
 - Capacity=1 for edges to matching nodes
 - Flow $F > 1$

Why capacity=1?

Network Flow

- Network flow parameters
 - Capacity=1 for edges to matching nodes
 - Flow $F > 1$

Why capacity=1?

- It describes how often we allow for matching a pair of images

Network Flow

- Network flow parameters
 - Capacity=1 for edges to matching nodes
 - Flow $F > 1$

Why capacity=1?

- It describes how often we allow for matching a pair of images

Why $F > 1$?

Network Flow

- Network flow parameters
 - Capacity=1 for edges to matching nodes
 - Flow $F > 1$

Why capacity=1?

- It describes how often we allow for matching a pair of images

Why $F > 1$?

- The flow F describes the maximum number of hypotheses we can find
- Allows for handling loops

Complexity

- Solving the network flow problem in general has a complexity of

$$\mathcal{O}(|\mathcal{X}|^2 |\mathcal{E}| \log |\mathcal{X}|)$$

- Our data association graph is a DAG
- Solution computed as F shortest path problems in a DAG with a single source via topological sorting

$$\mathcal{O}(F(|\mathcal{X}| + |\mathcal{E}|)) = \mathcal{O}(F|\mathcal{E}|)$$

Probabilistic Interpretation

- The network flow solution can be seen as a MAP solution (with $F=1$)

$$\mathcal{A} = (x^s, x_{a_1}, \dots, x_{a_A}, x^t)$$

$$\mathcal{A}^* = \operatorname{argmax}_{\mathcal{A}} p(\mathcal{A} \mid \mathcal{G})$$

$$= \operatorname{argmax}_{\mathcal{A}} p(\mathcal{G} \mid \mathcal{A}) p(\mathcal{A})$$

$$= \operatorname{argmax}_{\mathcal{A}} \prod_i p(x_i \mid \mathcal{A}) p(\mathcal{A})$$

Probabilistic Interpretation

- The network flow solution can be seen as a MAP solution (with $F=1$)

$$\mathcal{A}^* = \operatorname{argmax}_{\mathcal{A}} \prod_i p(\mathcal{A}) p(x_i | \mathcal{A})$$

$$p(x_i | \mathcal{A}) = \begin{cases} 1/Q & \text{if } x_i \in \mathcal{A} \\ 0 & \text{otherwise} \end{cases}$$

$$p(\mathcal{A}) = p_s p(x_{a_2} | x_{a_1}) \dots p(x_{a_A} | x_{a_{A-1}}) p_t$$

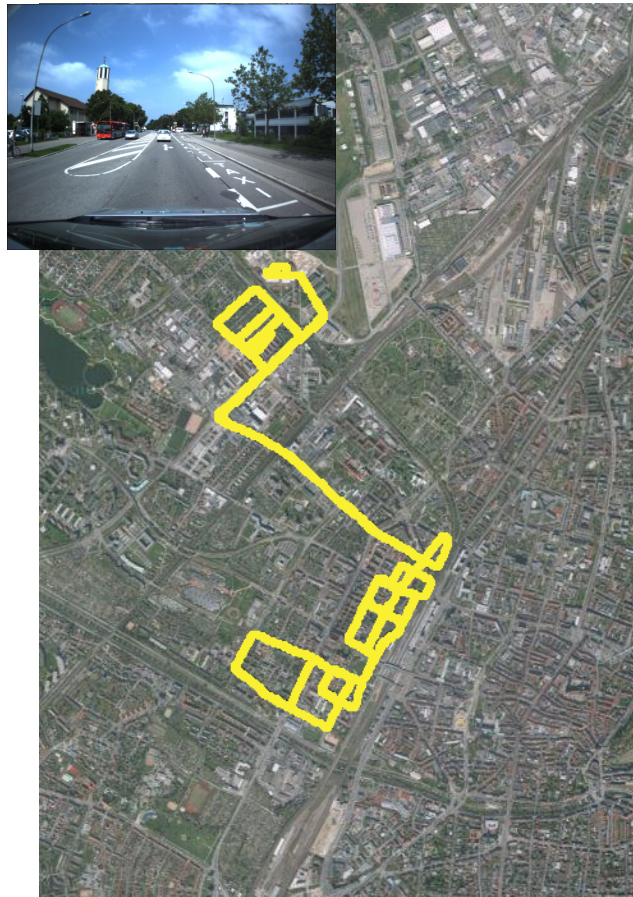


similarity c_{ij}

Summary of the Approach

- Use HOG descriptor on a tessellated image as the representation
- Compute similarity via cosine distance of the hog-based description
- Build the data association DAG
- Solve network flow problem through topological sorting

Summer & Winter Datasets Recorded in Freiburg



summer

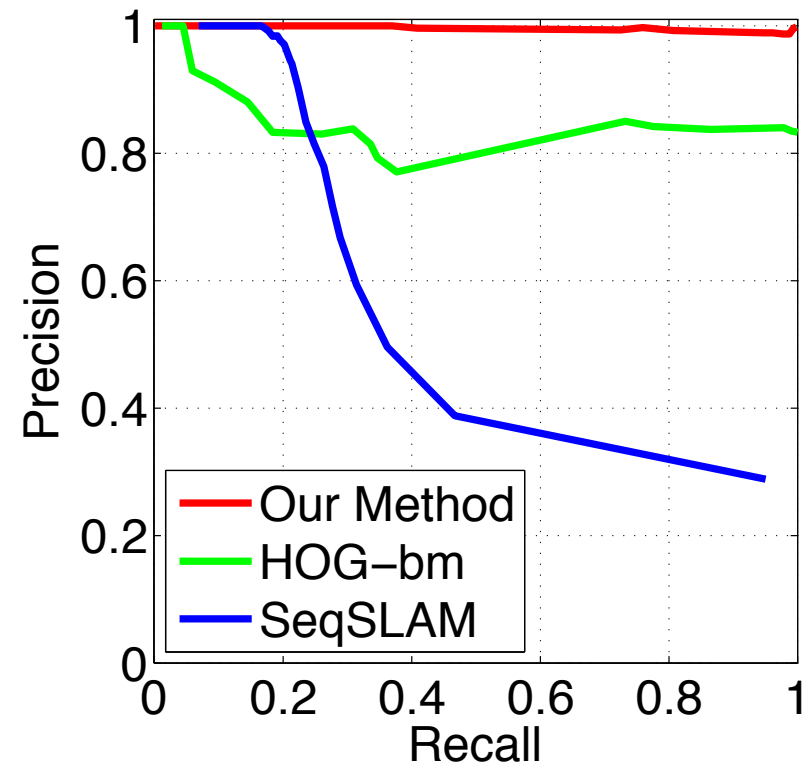
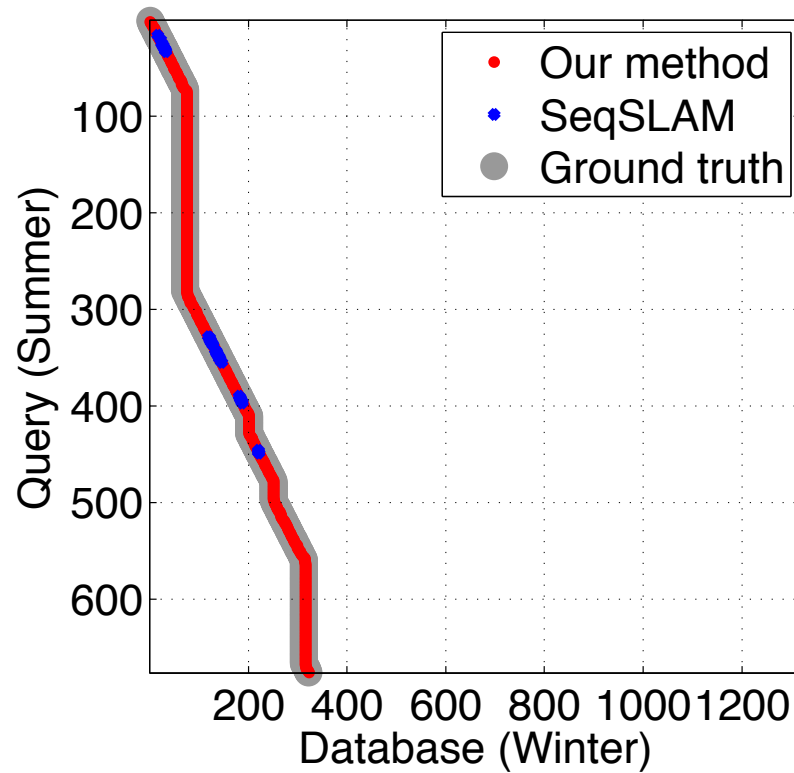


winter

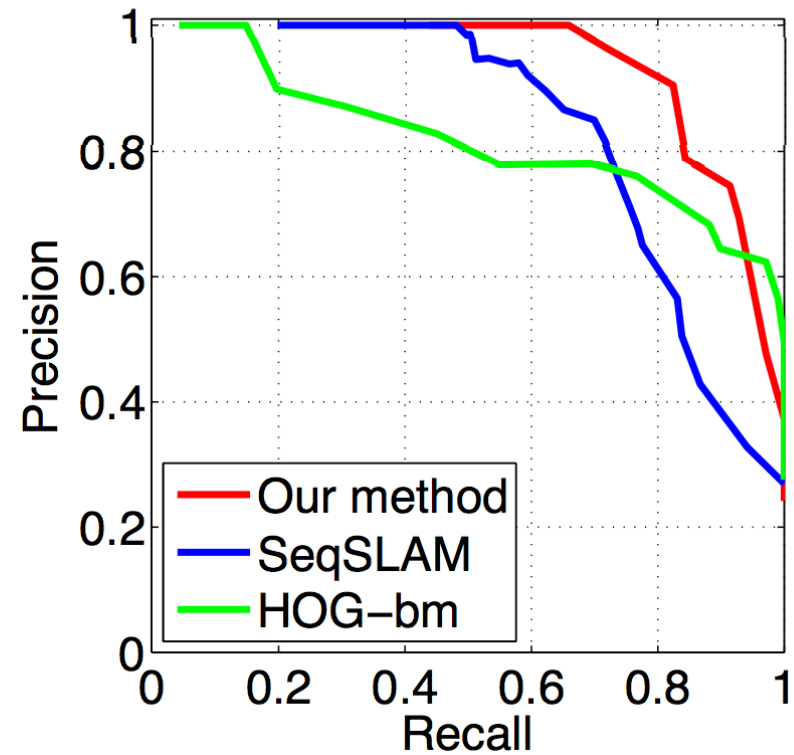
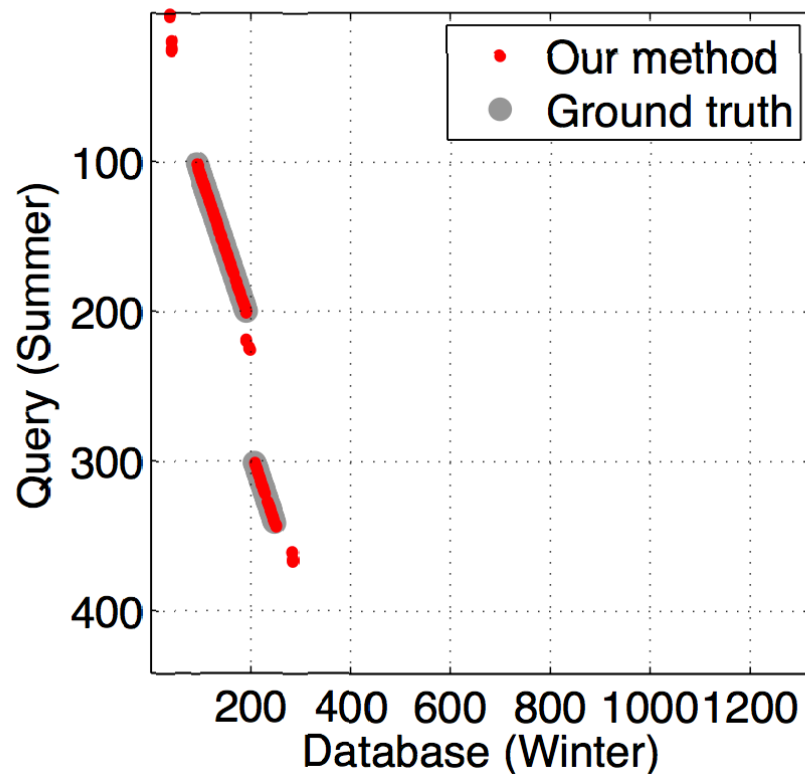
Example of a Matched Image Sequence Across Seasons



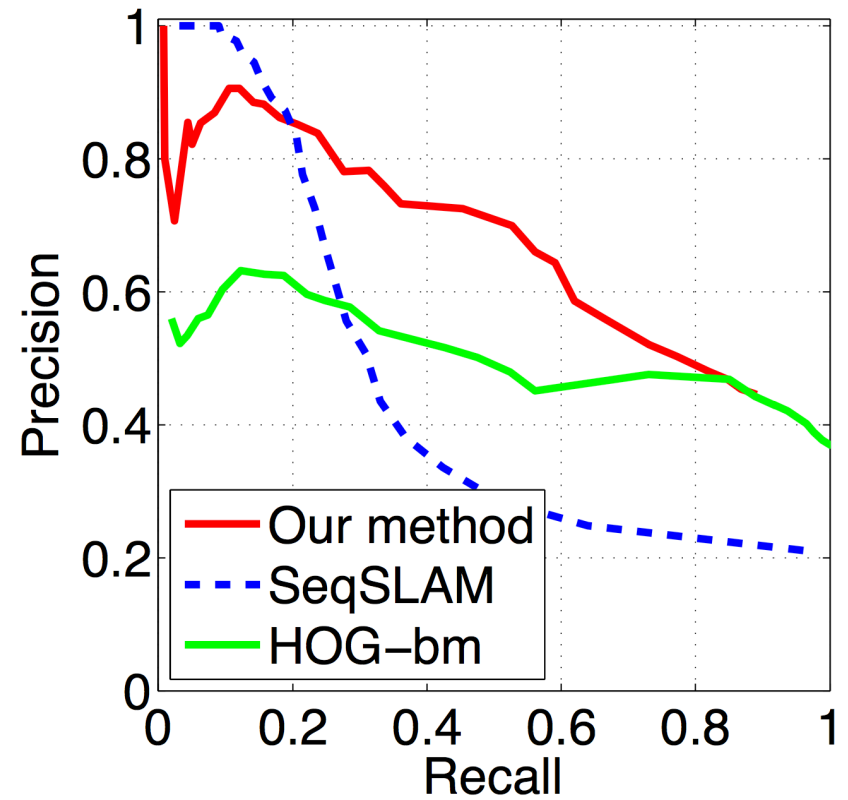
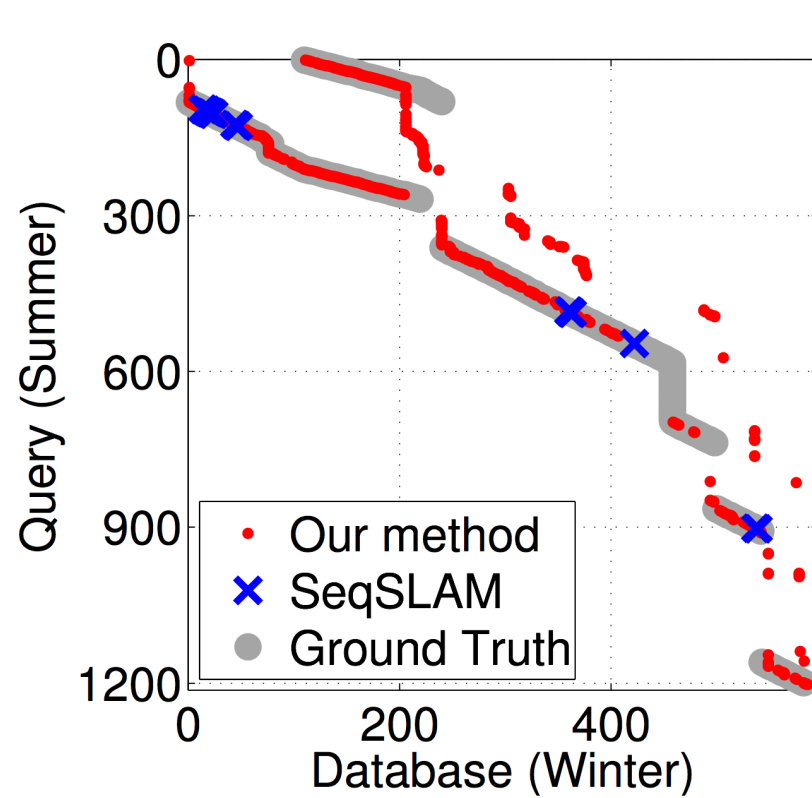
Results



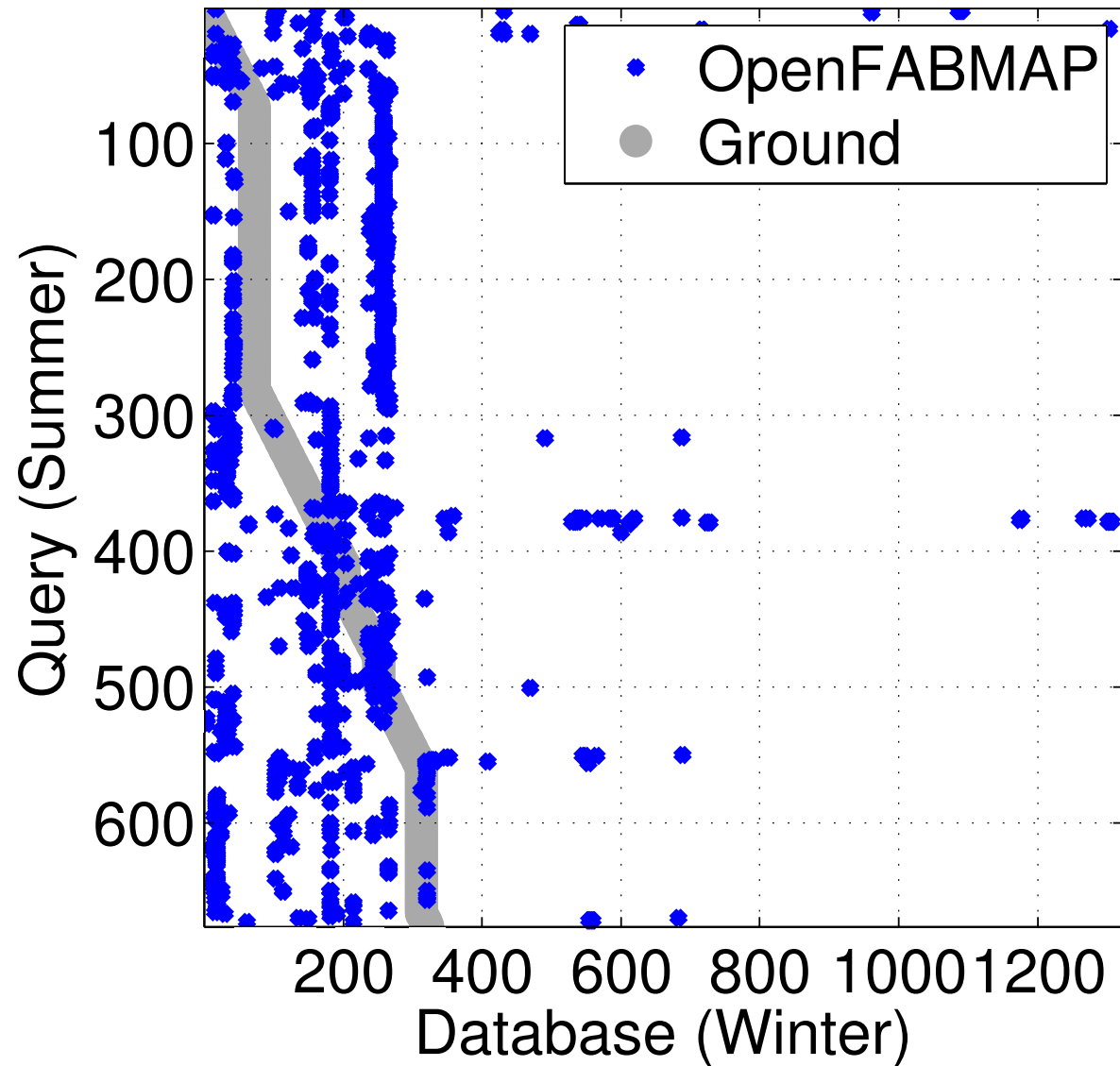
Results



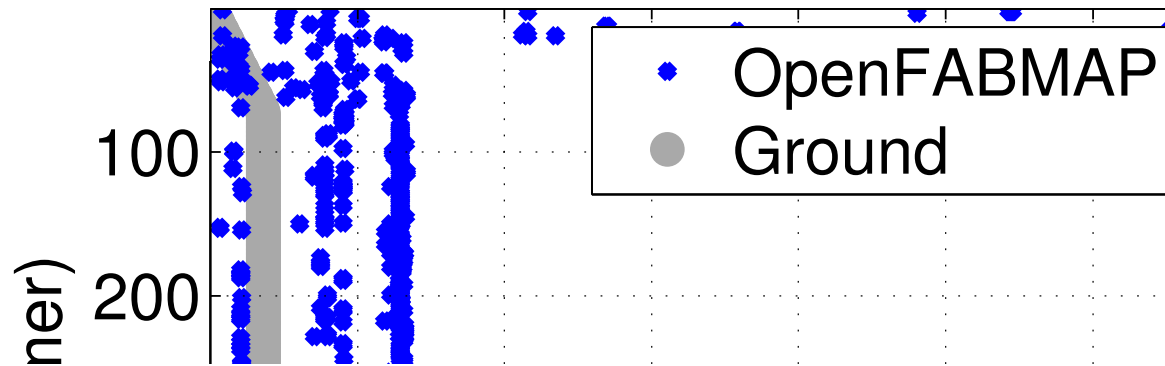
Loops



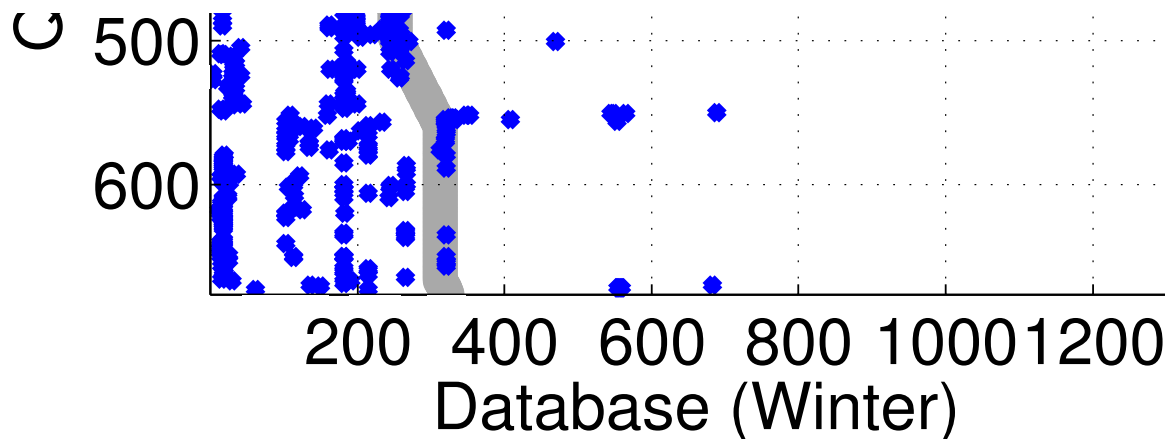
FABMAP2 Comparison



FABMAP2 Comparison



**Not a matter of FABMAP2
itself, but a result from
relying on SIFT/SURF**



Properties of the Approach

Visual localization along routes under strong seasonal changes

- without any GPS or odometry
- supports vehicles travelling at different speeds
- can handle revisits of places and loops
- can handle partial routes
- can operate on images taken at low frame rates
- does not require a pixel-accurate alignment
- does not require initialization
- and has no initial learning phase

Conclusion

- Across season place recognition
- Sequence information is valuable (this has been shown before)
- Flexible solution for finding sequences boost the performance
- Network flow in a DAG is effective and fast to compute
- Current bottleneck: computing the matching matrix

Thanks for Your Attention!

AAAI 2014

Robust Visual Robot Localization Across Seasons using Network Flows

Tayyab Naseer

University of Freiburg
naseer@cs.uni-freiburg.de

Luciano Spinello

University of Freiburg
spinello@cs.uni-freiburg.de

Wolfram Burgard

University of Freiburg
burgard@cs.uni-freiburg.de

Cyrill Stachniss

University of Bonn
cyrill.stachniss@igg.uni-bonn.de

Abstract

Image-based localization is an important problem in robotics and an integral part of visual mapping and navigation systems. An approach to robustly match images to previously recorded ones must be able to cope with seasonal changes especially when it is supposed to work reliably over long periods of time. In this paper, we present a novel approach to visual localization of mobile robots in outdoor environments, which is able to deal with substantial seasonal changes. We formulate image matching as a minimum cost flow problem in a data association graph to effectively exploit sequence information. This allows us to deal with non-matching image sequences that result from temporal occlusions or from visiting new places. We present extensive experimental evaluations under substantial seasonal changes. Our approach achieves accurate matching across seasons and outperforms existing state-of-the-art methods such as FABMAP2 and SeqSLAM.

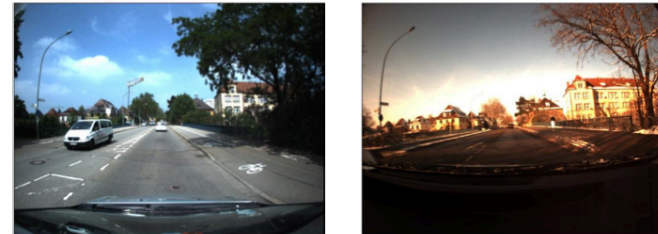


Figure 1: Two images of the same place in different seasons that are successfully matched by our approach.

route recognition over different seasons by exploiting the sequential nature of the data.

The main contribution of this paper is a novel approach that operates *without* using any GPS or odometry information, that does not require any season-based feature learning, that supports vehicles which travel at different speeds, that can handle revisits of places and loop closures and that