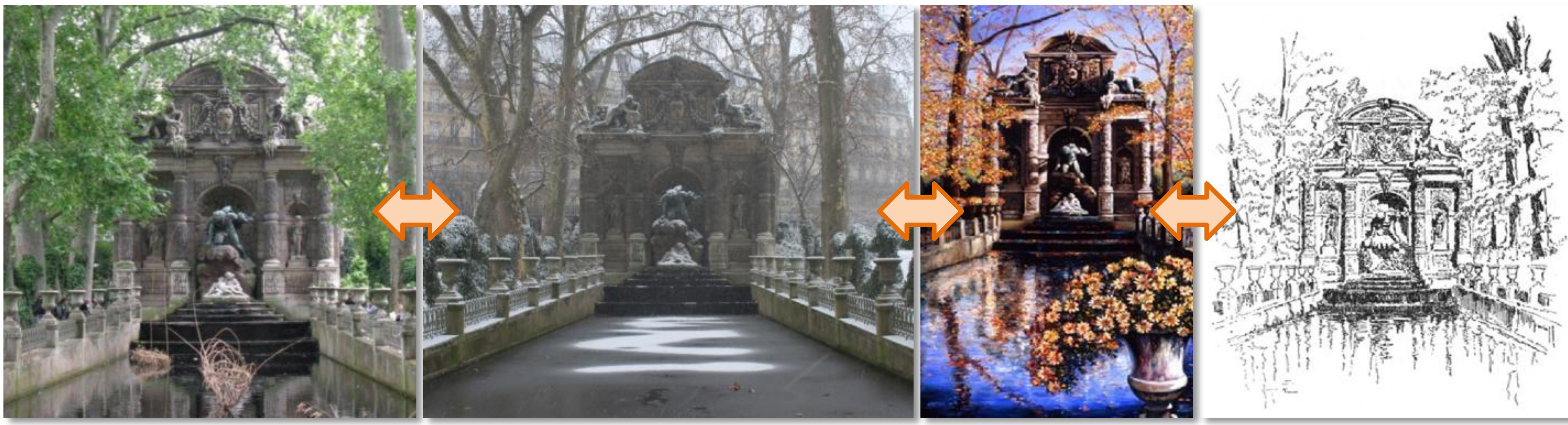
Image Matching

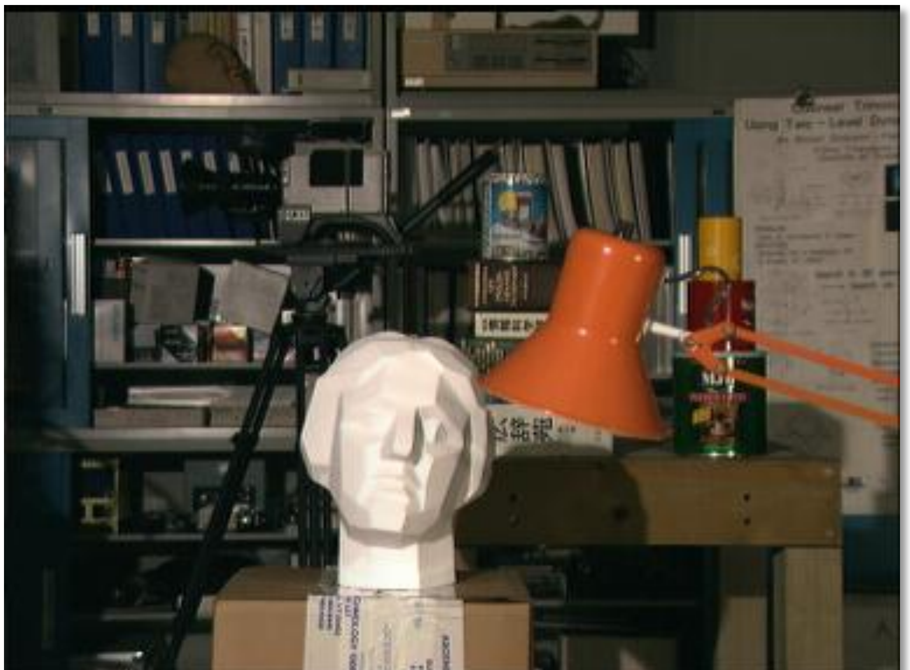


[S11]

[S11] Shrivastava et al. Data-driven visual similarity for cross-domain image matching, *SIGGRAPH ASIA* 2011

[C06] Cour et al. Balanced graph matching, *NIPS* 2006

Image Matching Problems



same object, same time, similar perspectives

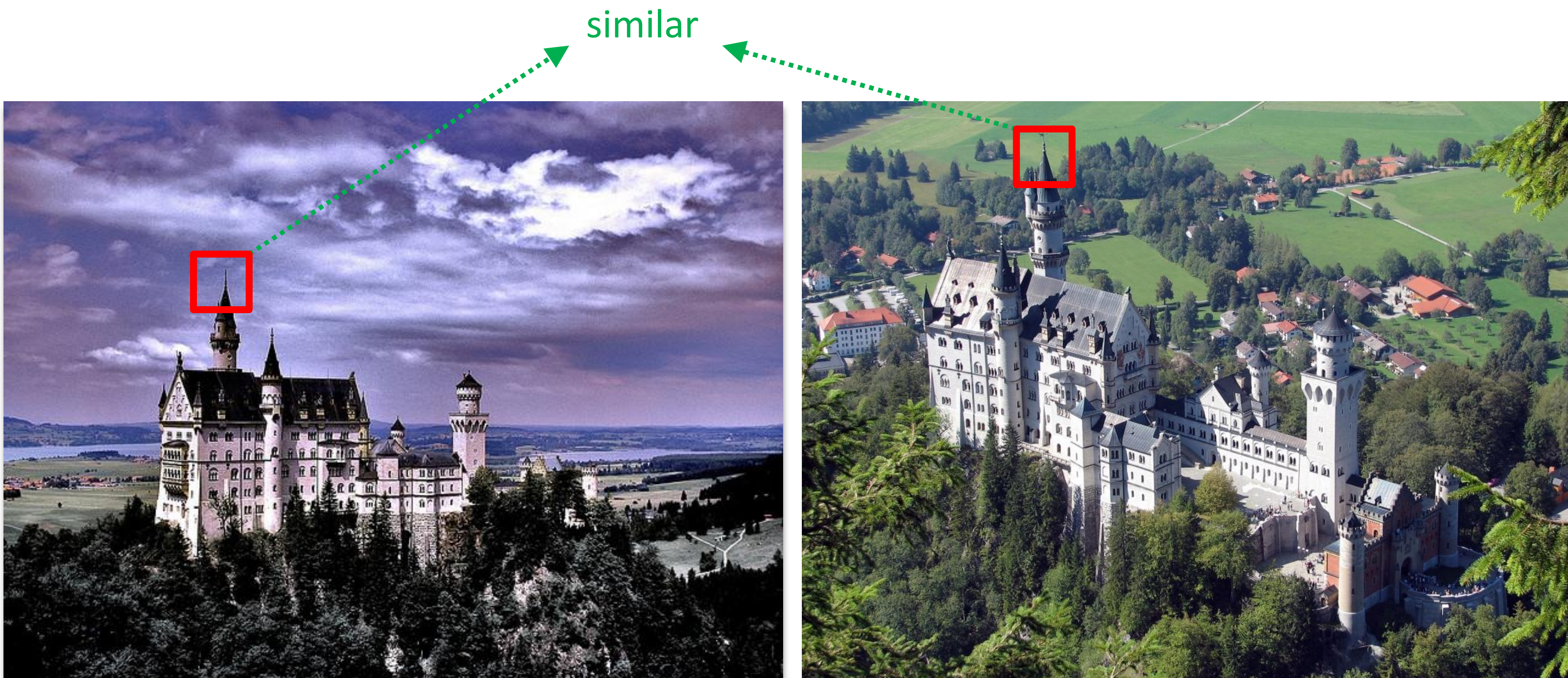
stereo, optic flow algorithms

Image Matching Problems

same object, changed appearance, similar perspectives

holistic image matching

Image Matching Problems



same object, changed appearance, different perspectives

geometry-based matching:
e.g., estimating fundamental matrix

Image Matching Problems

[SIFT flow]

different objects in the same class

SIFT flow

Image Matching Problems

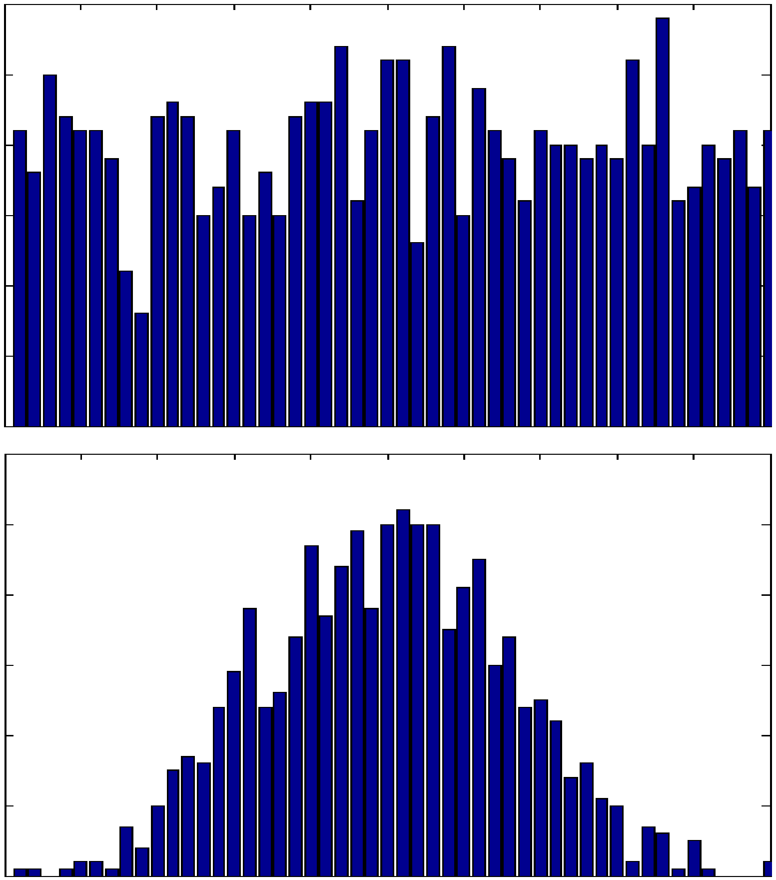
[S11]

different domain

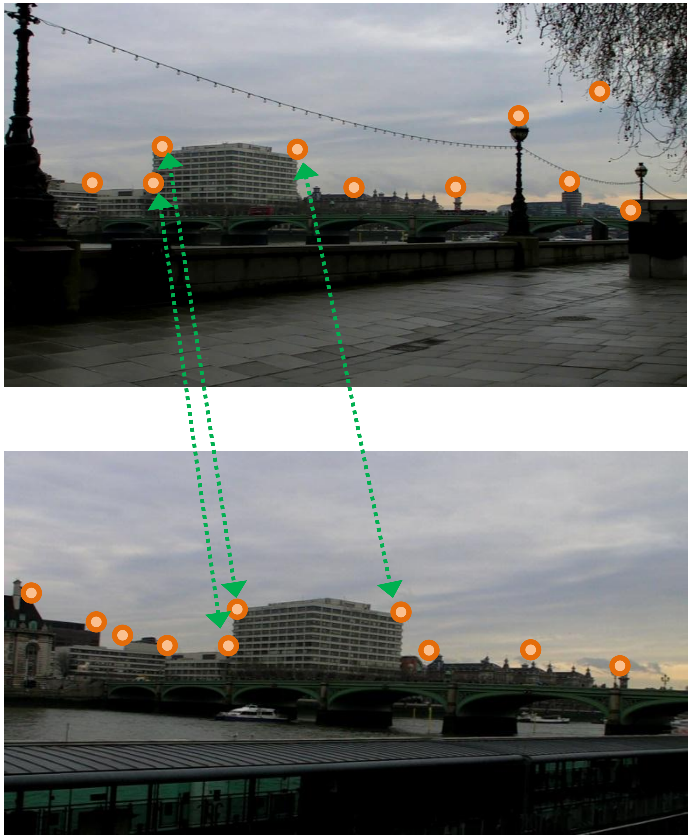
are they similar?

discrimination power vs. robustness

Image Matching Approaches



Holistic matching



Feature-based

output: real-valued score vs. feature-correspondence

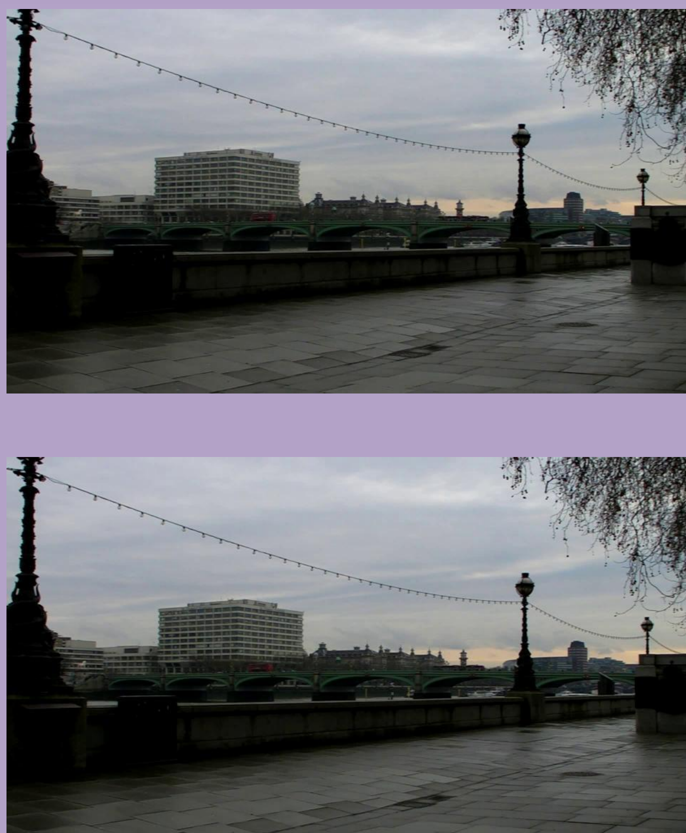
Holistic Matching

- Image represented as a vector
  - Dense raw data: color value, gradients, etc.
  - Compact geometry-preserving or independent representations
    - bag-of-words, GIST, etc.
- Similarity measure (for vector space)
  - Euclidean inner-product, histogram intersection, etc.
- Pros
  - Fast, robust against clutter
- Cons
  - Sensitive to scale, location, prespective, etc. variations

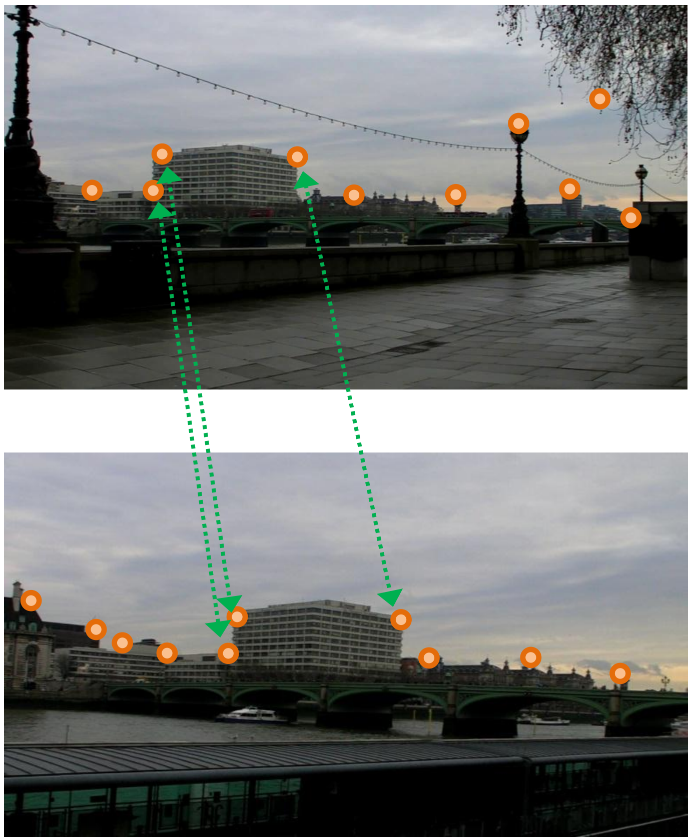
Feature-based Matching

- Matching performed based on detected features
- Pros
  - Robust against scale, location, prespective, etc. variations.
- Cons
  - Typically formulated as a non-trivial optimization problem (time consuming)
  - Bad for cluttered scene [S11]

Image Matching Approaches

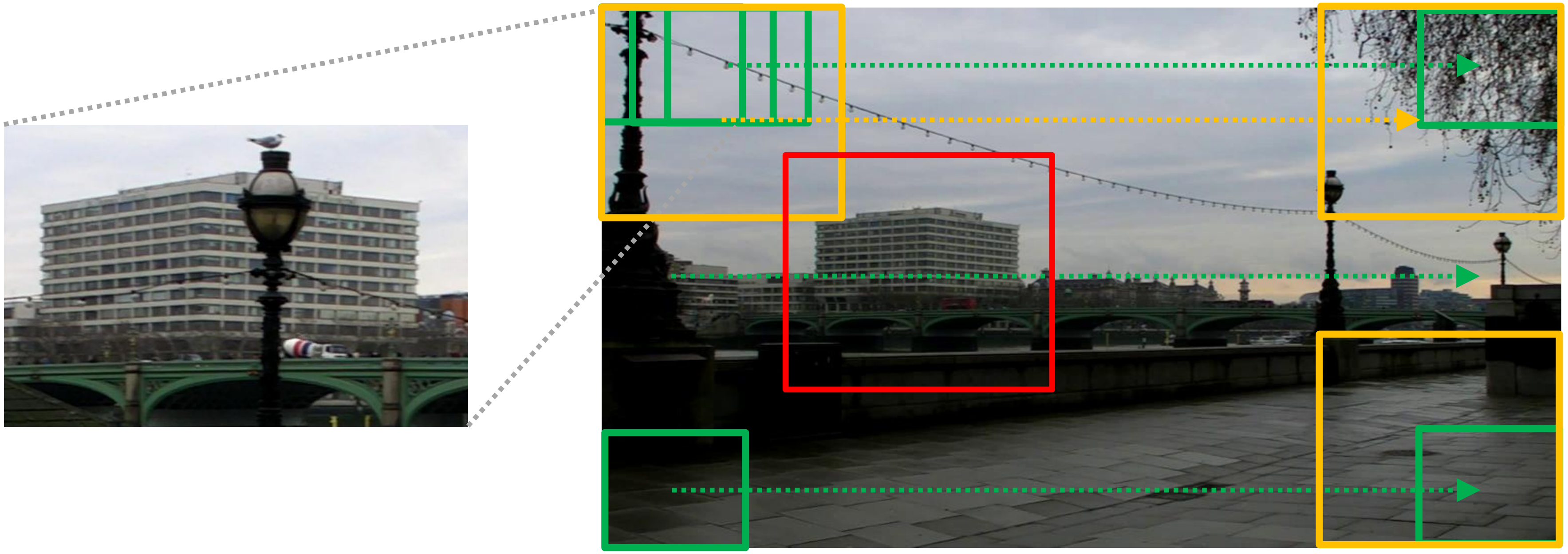


Holistic or dense matching



Feature-based

Sliding Window



helps bypassing problems of scale and location variations

Data-driven Visual Similarity [S11]

- Typical holistic image matching, e.g., Gist, bag-of-words:

$$s(x, y)$$

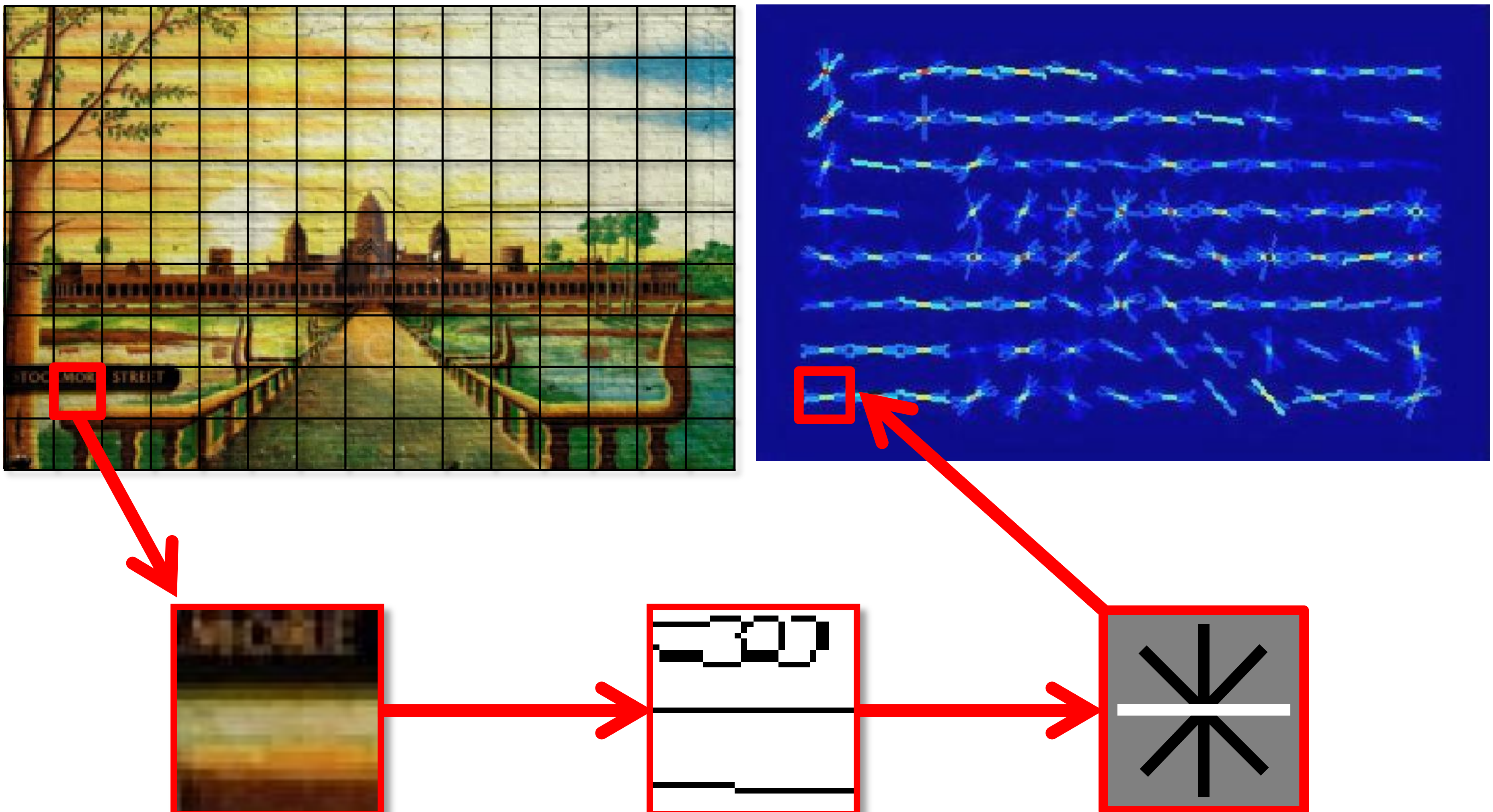
  - (dis-)similarity beteen x and y
  - function of two arguments, symmetric, general

- [S11]

$$s(y|x) := f_x(y)$$

  - (dis-)similarity of y given x
  - function of one argument, specific to x
  - exploits (huge) data set; unsupervised

Image Representation:
histogram of oriented gradients (HOG)



[HOG]

slide taken from the project website of [S11]

Similarity function $f_x(\cdot)$

- Constructed as a classifier for x against the rest
  - Training: x vs. many example images
    - Bootstrapping negative examples
  - Sliding windows
  - Linear support vector machine (SVM)

$$f_x(y) = {w_x}^T y$$

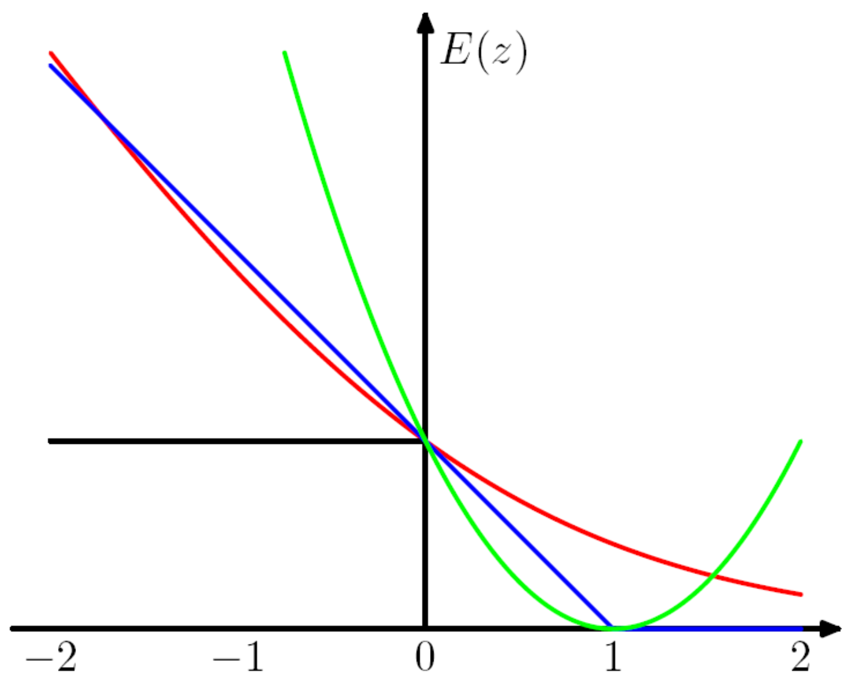
Linear SVM Classifier $f_g(y) = {w_g}^T y$

training error regularizer

$$w_g = \underset{w}{\operatorname{argmin}} \sum_i h(w^T x_i y_i) + \lambda \|w\|^2$$

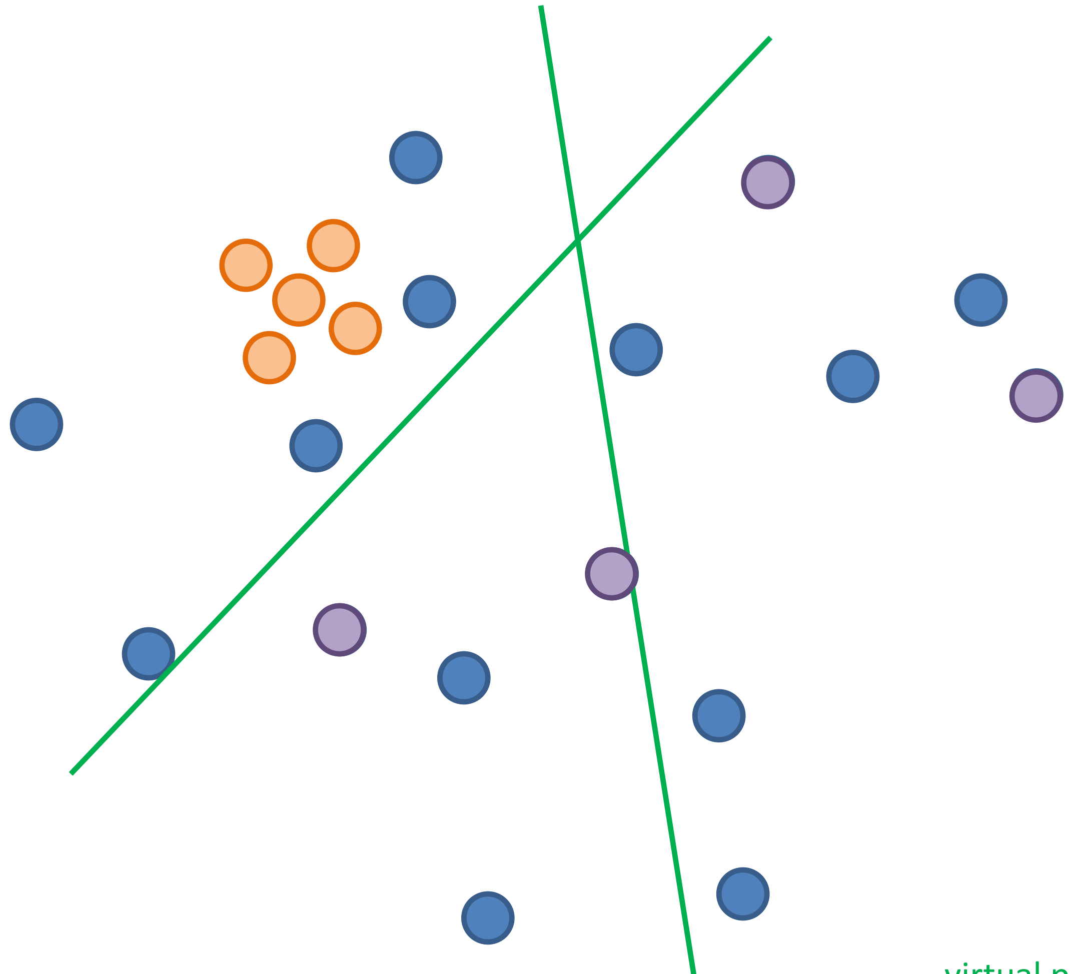
ω_i

- $y_i = \begin{cases} +1 \text{ if } x_i \text{ corresponds to } g \\ -1 \text{ otherwise} \end{cases}$
- h: hinge loss



[PRML]

Constructing f with only one positive example



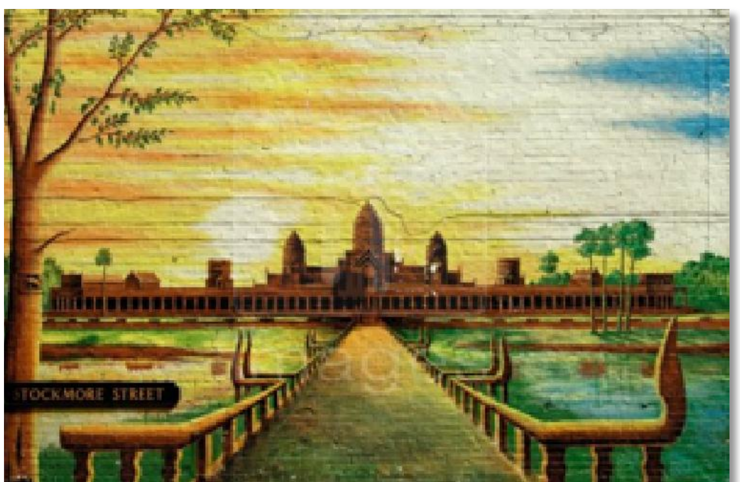
virtual positive examples +
bootstrapping for negative examples

Interpretation of w_g as Saliency Map

$$f_g(y) = {w_g}^T y = \sum_j [w_g]_j y_j$$

If the j-th element is not relevant, training SVM may result in small $\left|[w_g]_j\right|$

not discriminative

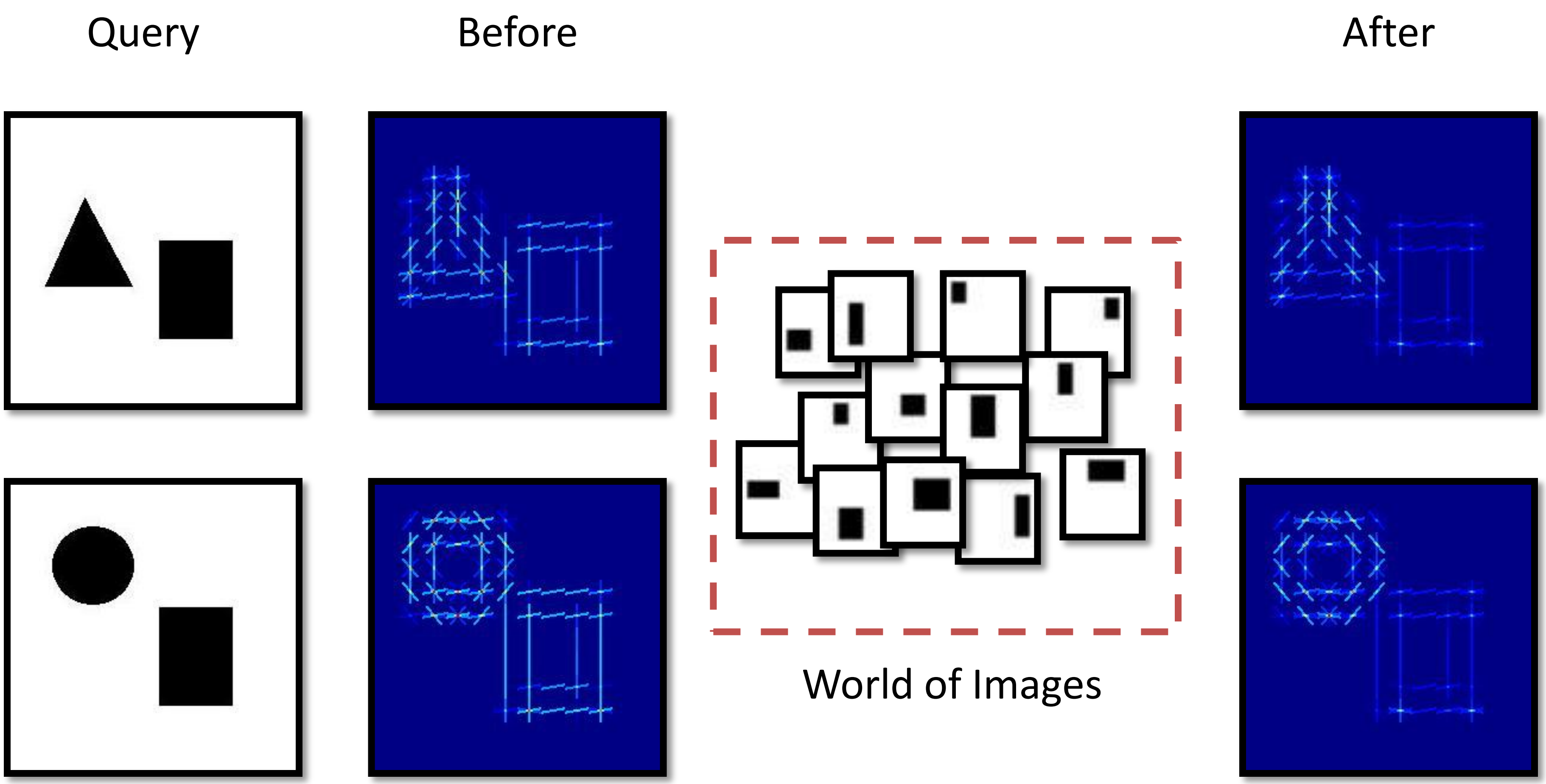


g

ideal w_g overlaid on g

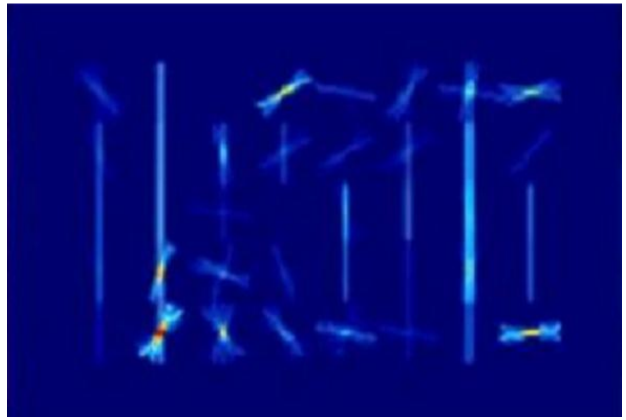
images taken from the project website of [S11]

Synthetic Example



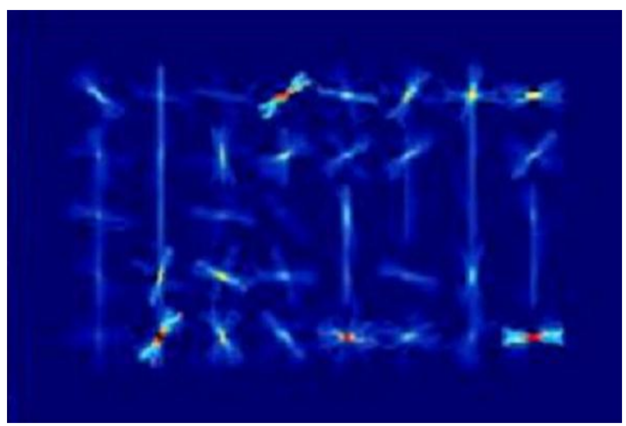
Query by Painting

Input Query



HOG

Top Match



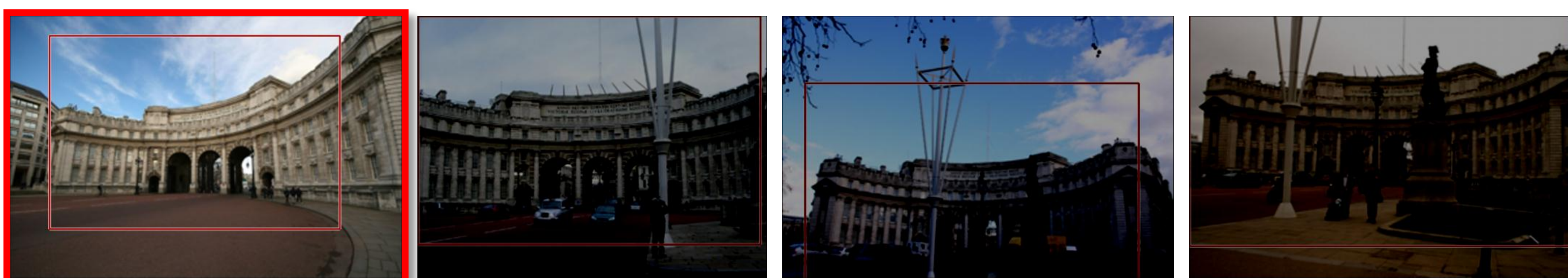
Learnt Weights

Top Match

slide taken from the project website of [S11]

Query by Image

Input Query



Top Matches

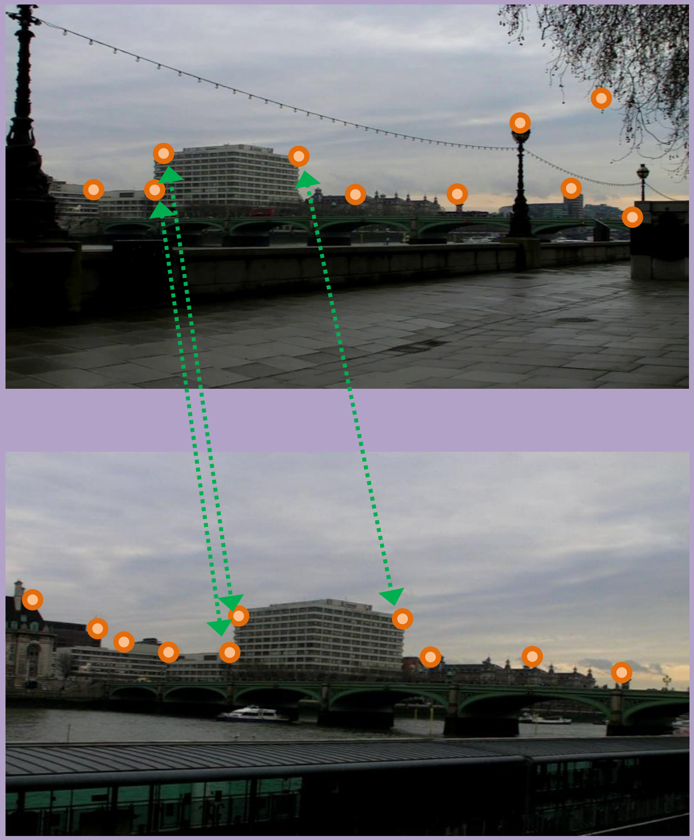
slide taken from the project website of [S11]

Applications

Video

Image Matching Approaches

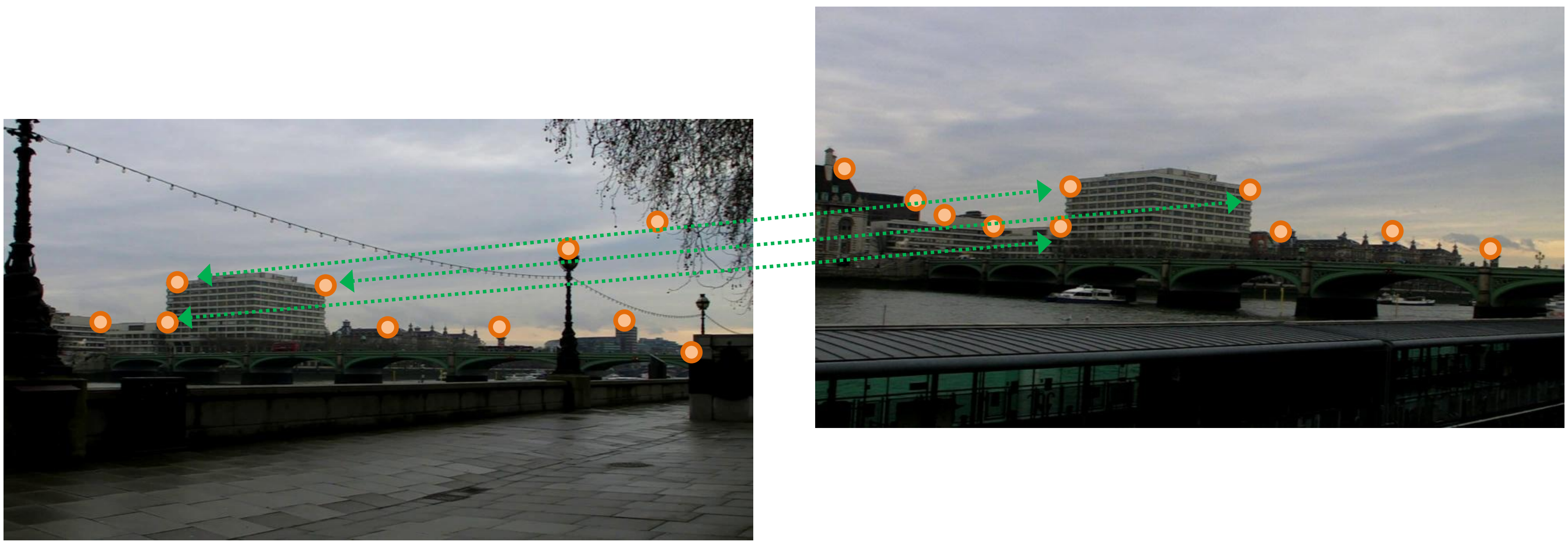
Holistic or dense matching



Feature-based

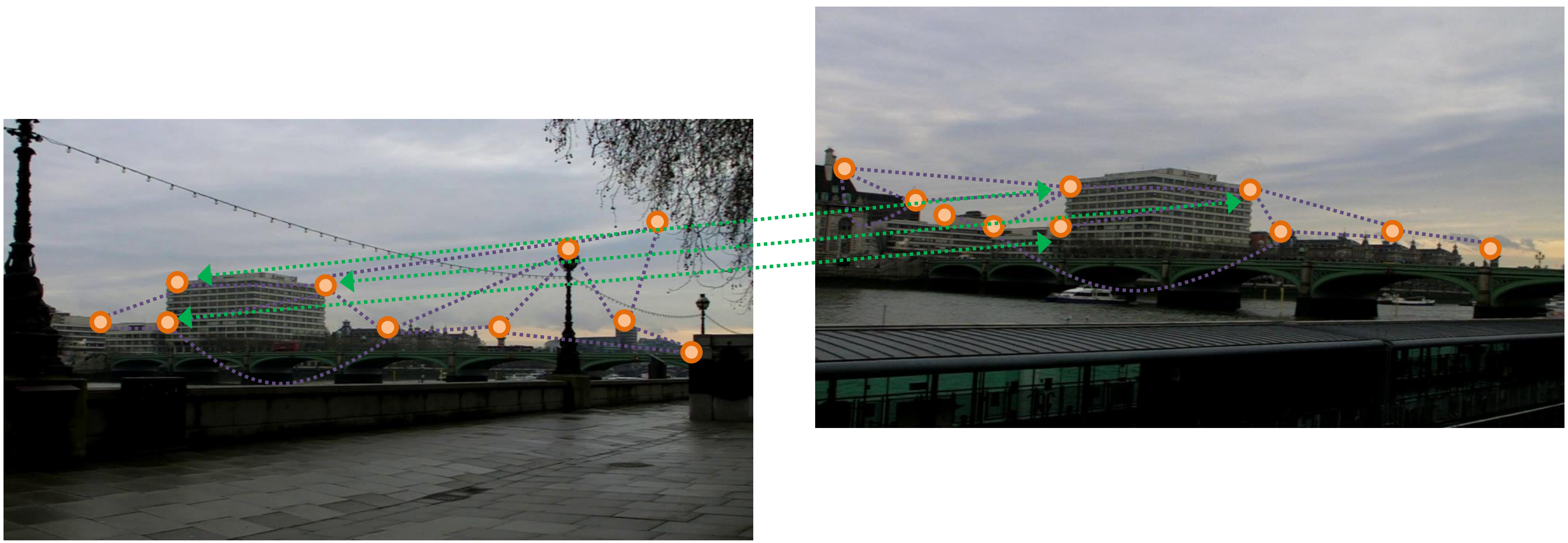
Balanced Graph Matching [C06]

- Image matching can be formulated as graph matching

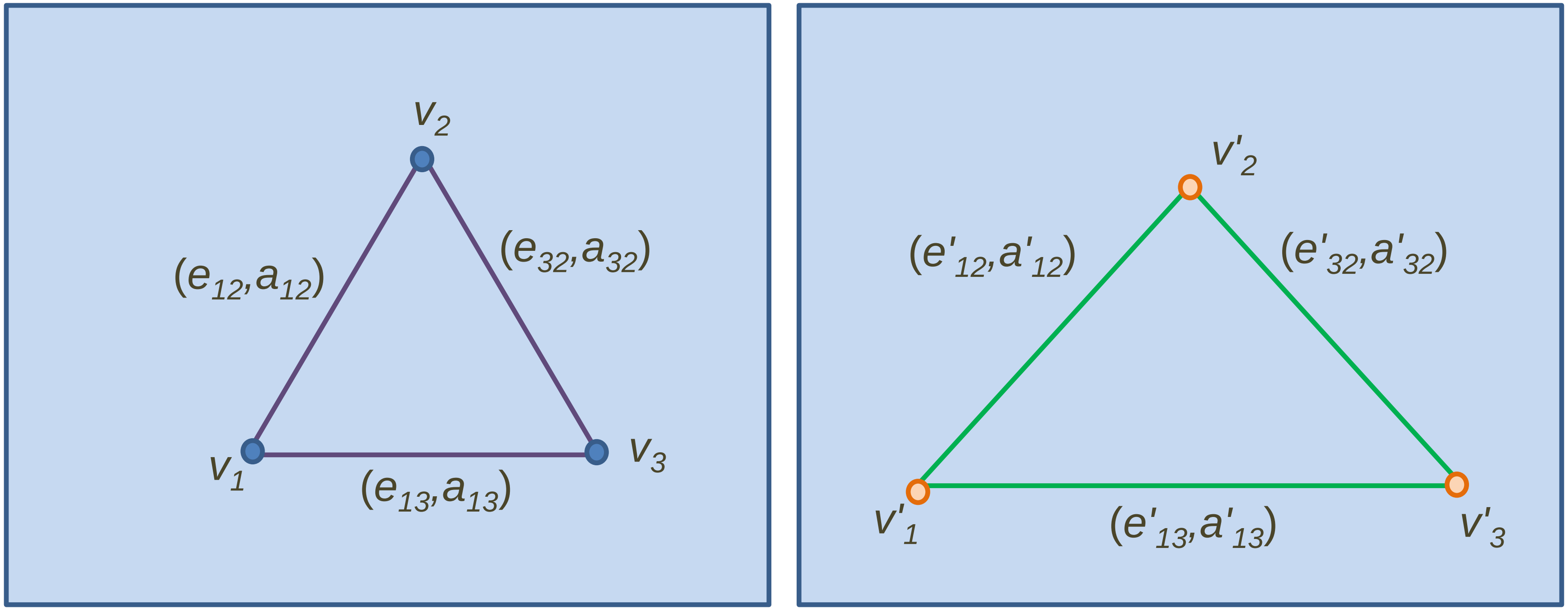


Balanced Graph Matching [C06]

- Image matching can be formulated as graph matching



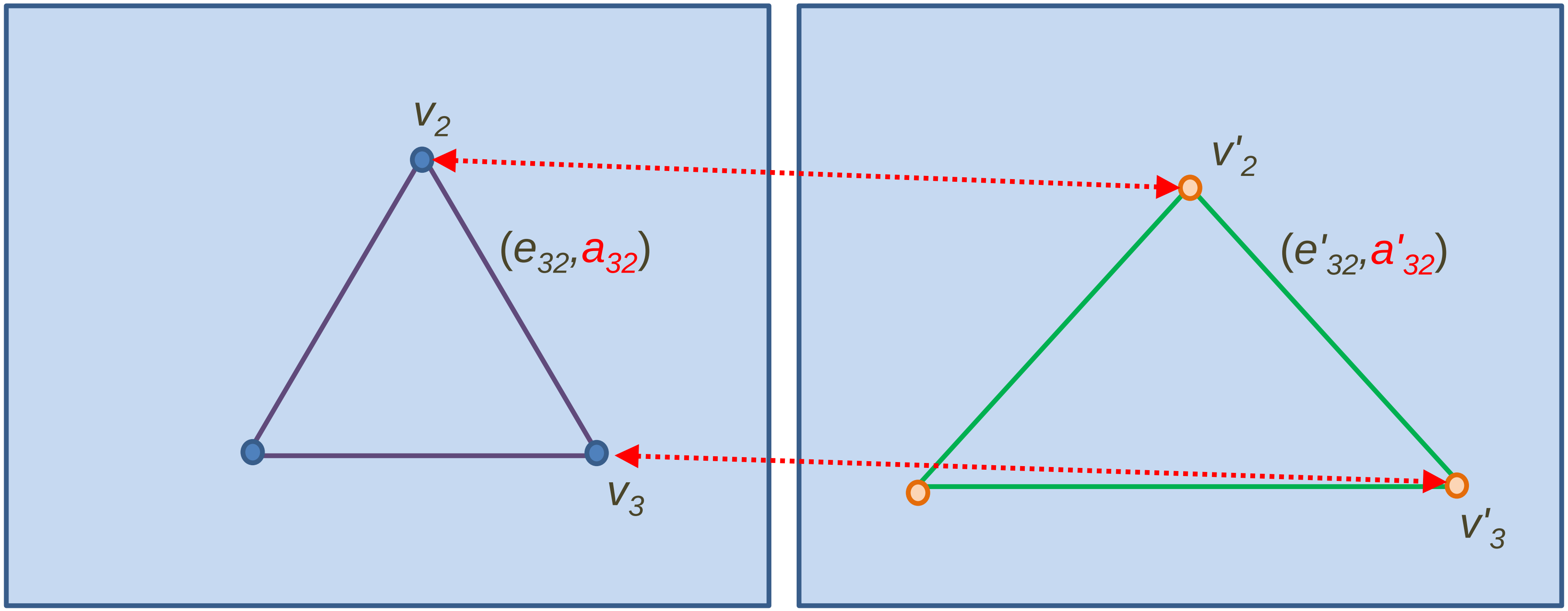
Graph Matching



$G=(V,E,A)$ $G'=(V',E',A')$

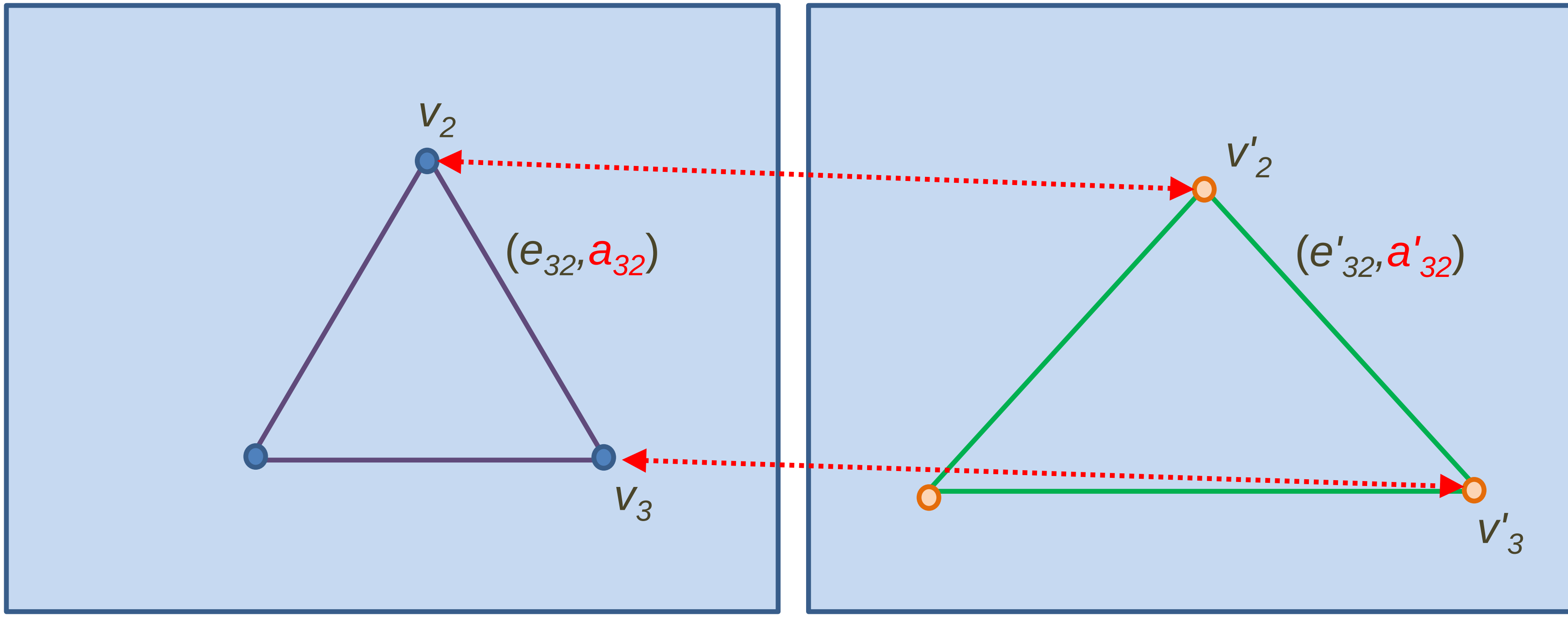
variable: matrix M containing vertice correspndences

Exploiting Regularity



If v_i and v_j matches v'_i and v'_j, respectively a_{ij} and a'_{ij} must be similar

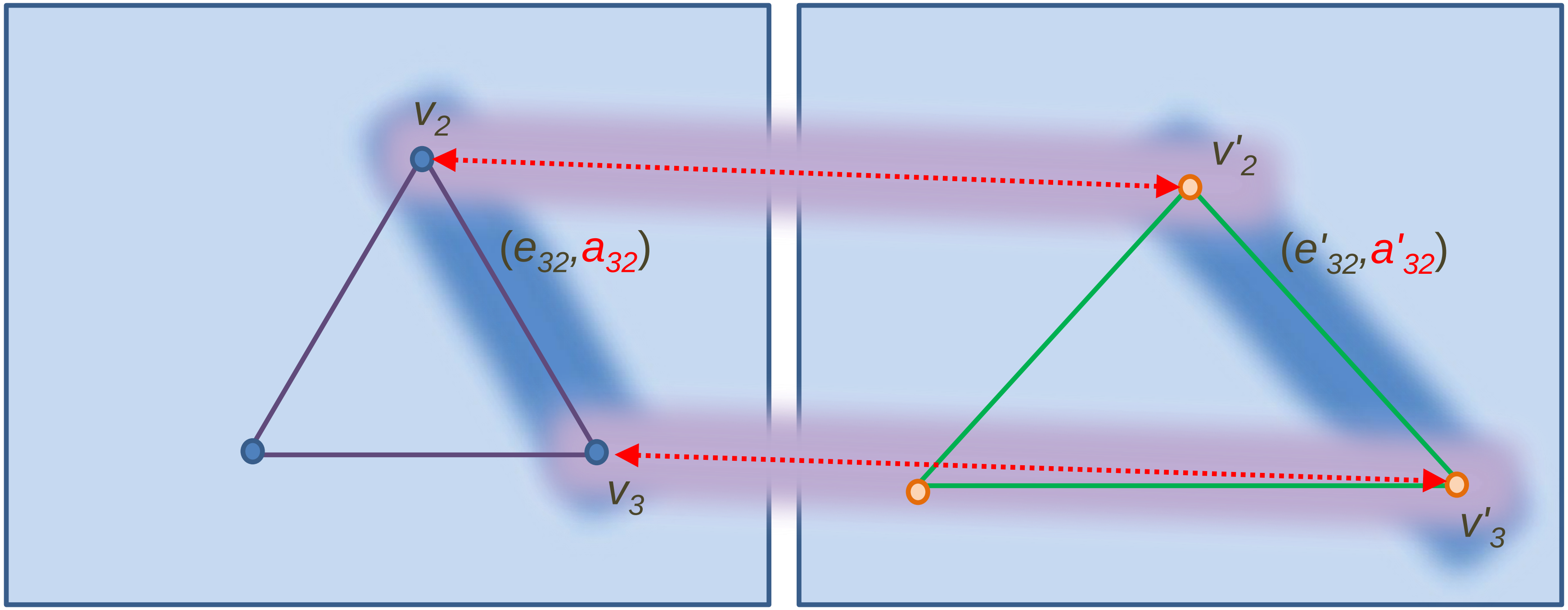
Graph Matching Score



$$E(M) = \sum_{e \sim e'} f(a_e, a_{e'}) \quad M = \{i, i'\}$$

f: similarity measure

(Original) Optimization Problem

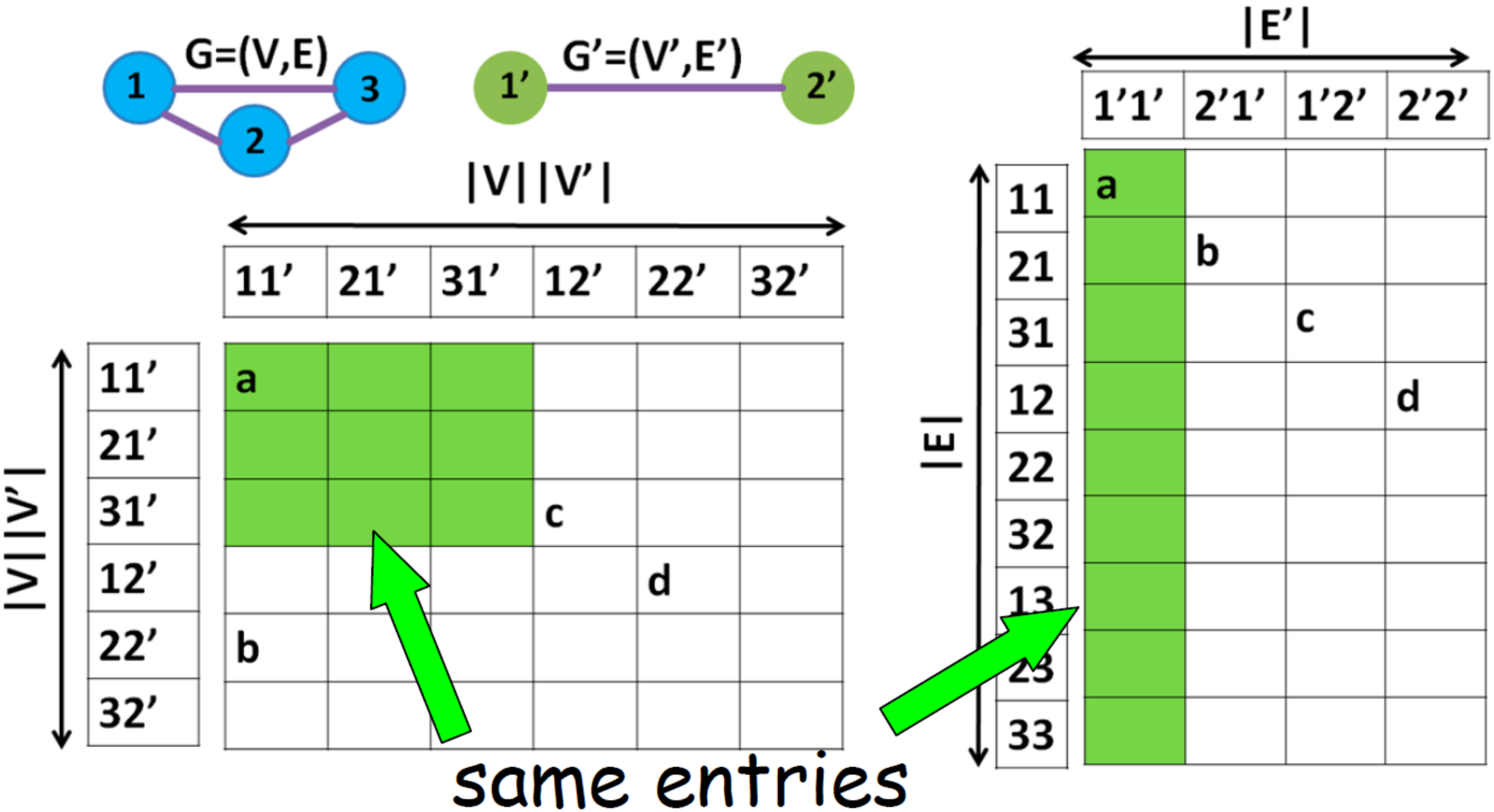


Maximize $E(M) = \sum_{e \sim e'} f(a_e, a_{e'}) \quad M = \{i, i'\}$

$$E(x) = x'Wx, W_{ii',jj'} = f\left(a_{ij}, a_{i'j'}\right) \quad x \in \{0,1\}^{nn'} \quad Cx \leq b$$

If v_i and v'_i is conntected to v_j and v'_j, respectively match vectors (i,i') and (j,j') must be similar

Dual Representations



W: matching compatibility matrix

S: edge similarity matrix

(Relxed) Optimization Problem

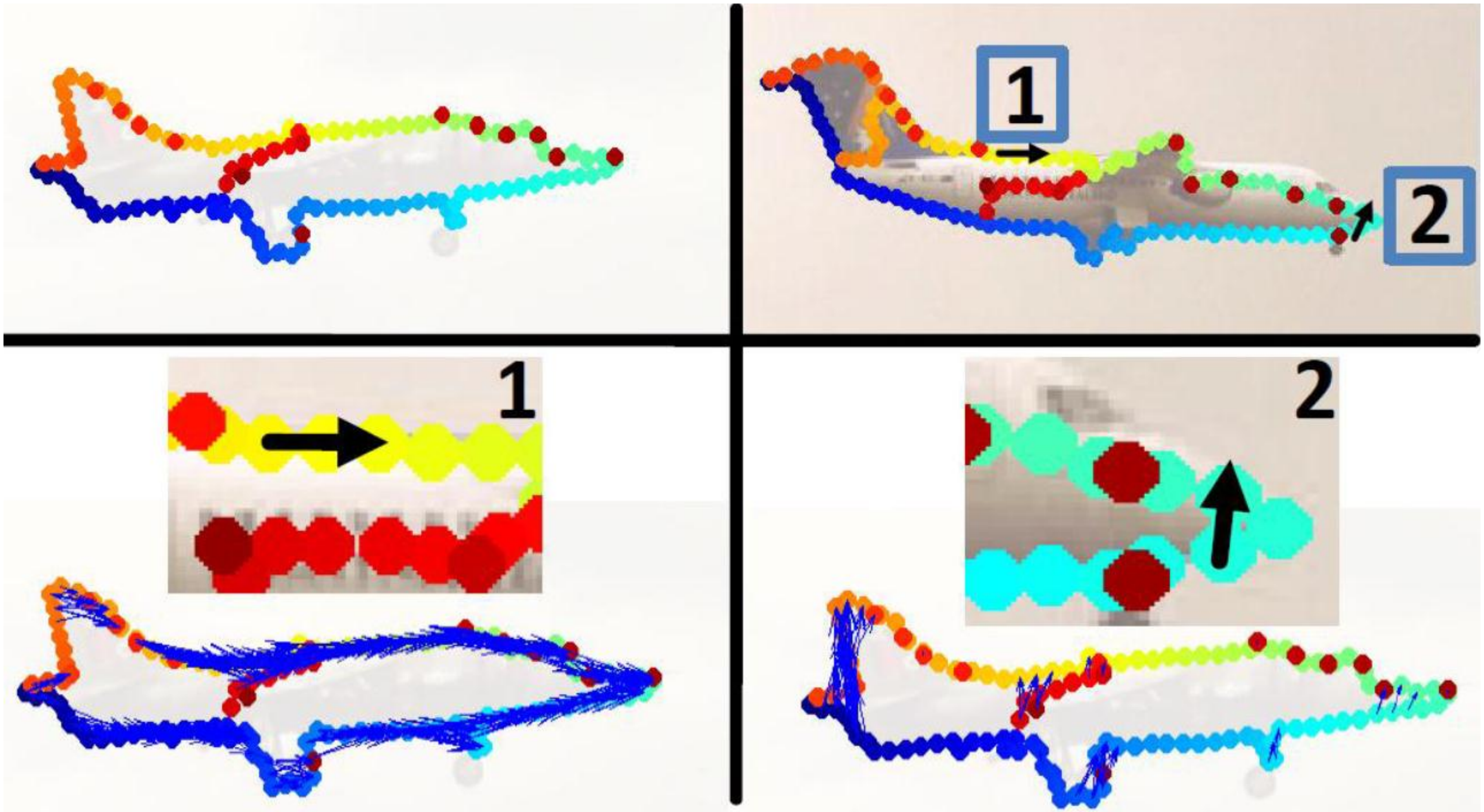
Maximize $\sum_{i'} x_{ii'}$=1, $\sum x_{ii'}$=1

Original: $E(x) = x'Wx$, $Cx \leq b$, $x \in \{0,1\}^{nn'}$

Relaxed: $E(x) = \frac{x'Wx}{x'x}$, $Cx = b$

x* obtained by solving an eigensystem

Normalization



edge 1 has many matching edges; individual matches are less informative

edge 2 has few matching edges; individual matches are more informative

Normalization

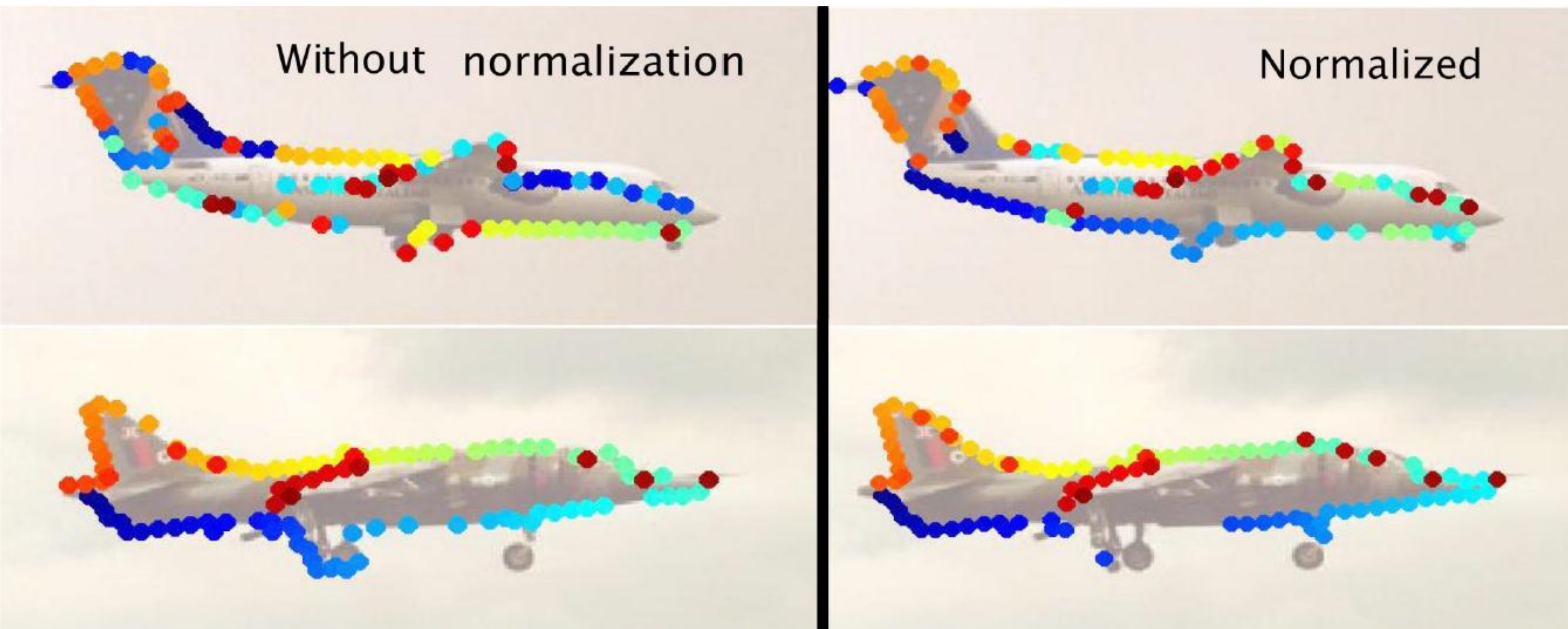
- Normalize W (equivalently S)
  s.t. each column / row of S sums to one

1. Input: compatibility matrix W, of size $nn' \times nn'$
2. Convert W to S: $S_{ij,i'j'} = W_{ii',jj'}$
3. repeat until convergence
   (a) normalize the rows of S: $S^{t+1}_{ij,i'j'} := S^{t}_{ij,i'j'} / \sum_{k'l'} S^{t}_{ij,k'l'}$
   (b) normalize the columns of S: $S^{t+2}_{ij,i'j'} := S^{t+1}_{ij,i'j'} / \sum_{kl} S^{t+1}_{kl,i'j'}$
4. Convert back S to W, output W

Image Matching

$S(e, e') = 1$ if $\cos(\angle e - \angle e') > \cos \pi/8$, $\frac{|l(e)-l(e')|}{\min(l(e),l(e'))}$ <0.5

$\angle(e)$: angle, $l(e)$: length $\quad e = ij$ within 30 pixels



Questions

- What do you like about those two algorithms?
- Why does the first algorithm work for cross-domain setting?
- What are the limitations?
- From where one could improve the algorithm?
- Time complexity?
- Cool application?

References

[SIFT flow] C. Liu, J. Yuen, and A. Torralba, SIFT flow: dense correspondence across scenes and its applications, *TPAMI* 2011

[HOG] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, *CVPR* 2005

[PRML] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006

Effect of Normalization