Joint Data Alignment Up To (Lossy) Transformations

Andrea Vedaldi Gregorio Guidi Stefano Soatto - UCLA Vision Lab

Joint alignment

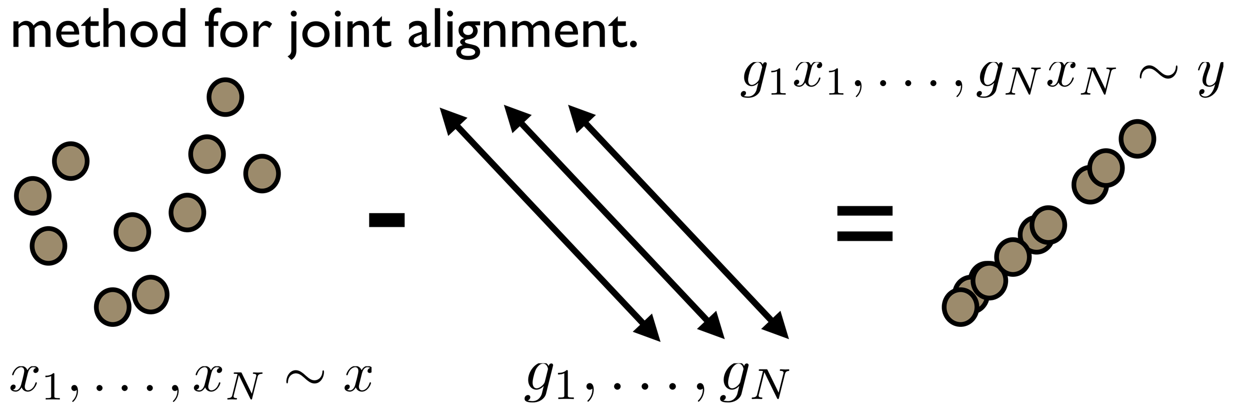
Visual data is often affected by nuisance transformations (e.g. viewpoint, illumination, calibration of sensors, etc.).

Removing the irrelevant variability makes analysis (e.g. recognition) much easier.

> **Goal:** ***remove*** systematic nuisance ***transformations*** from a ***collection*** of data in order to ***simplify*** further analysis.

Image congealing

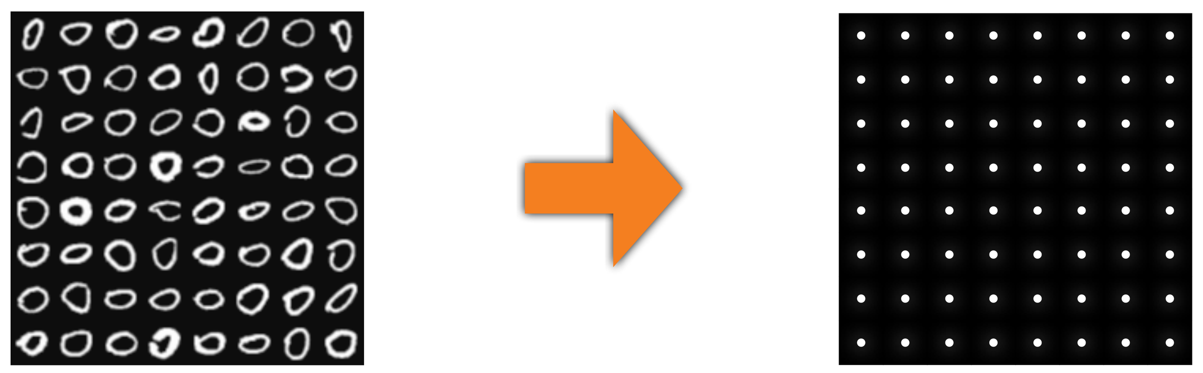
Image congealing (IC) [CONG] is a powerful method for joint alignment.



- Find a transformed version of the data which is "as simple as possible".
- Complexity: (differential) entropy $\mathcal{H}(y)$
- Formulation:
  $$\min_{g_1,\dots,g_N} \mathcal{H}\{g_1 x_1, \dots, g_N x_N\}$$

What if transformations g_i are lossy?

Example: Affine warps of digital images



Must regularize!

$$\min_{g_1,\dots,g_N} \mathcal{H}\{g_1 x_1, \dots, g_N x_N\} + \lambda \mathcal{R}\{g_1, \dots, g_N\}$$

But what regularizer?

Lossy Compression

Do we really need to regularize?

> **Idea 1:** Obtaining simple data is not enough. We want a ***simple representation of the original data***.

Complexity-distortion formulation

- Invariant distortion
  $D(x,y) = E[\min_{g \in G} d_0(x, gy)]$
- Complexity $\mathcal{C}(x,y)$
- Search for optimal trade-off
  $$\min_{p(g,y|x)} \mathcal{D}(x,y) + \lambda \mathcal{C}(x,y)$$

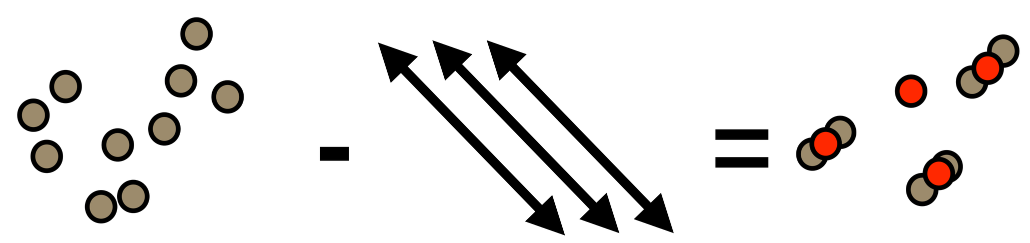
The formulation is reminiscent of rate-distortion, vector quantization, entropy constrained vector quantization. Advantages:

- Finds an actual representation
- Handles naturally lossy transformations
- Similarly to IC, scales better than [TCA]

"Structural" Complexity

Differential entropy ≈ # of prototypes to approximate data with ε accuracy.

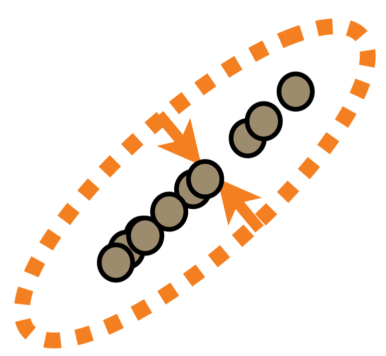
Differential entropy may not characterize well data alignment:



> **Idea 2:** Find ***measure of complexity*** that:
> - encourages ***global alignment***;
> - captures ***meaningful properties***.

Example: Affine complexity

$$\mathcal{C}(x,y) = \frac{1}{2} \log\det\left(I + \frac{\Sigma_p}{\epsilon^2}\right)$$

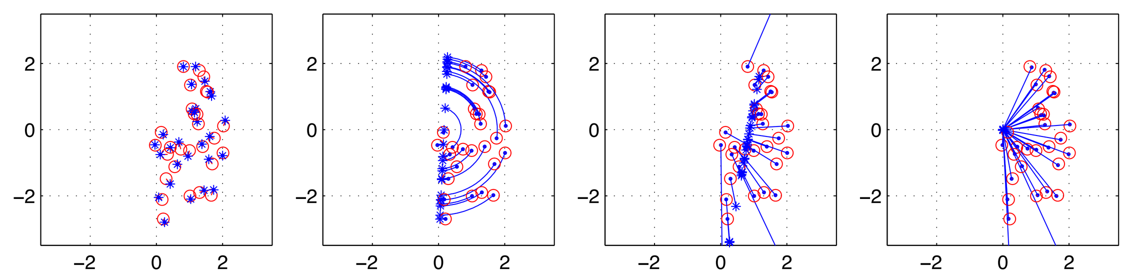


Interpretation and Scaling

- Continuous data ⇒ differential entropy.
- Differential entropy is meaningful only up to a quantization error, which is relative to the scale of the data.
- If transformations include data scalings, minimizing differential entropy may become meaningless.

> **Idea 3:** Differential entropy can be made ***meaningful by*** fixing the ***scale*** of the data.

$$\mathcal{C}(x,y) = \frac{1}{2} \log\det\left(I + \frac{\Sigma_p}{\epsilon^2 \operatorname{tr} \Sigma_p}\right)$$



Algorithms

How do we align very large dataset?

$$\frac{1}{K}\sum_{k=1}^{K} \|x_k - g_k y_k\|^2 + \frac{\lambda}{2} \log\det\left(I + \frac{YY^\top}{\epsilon^2 YY^\top}\right)$$

> **Observation:** It is easy to compute the approximate variation of the energy when a single point is moved.

Three algorithms (all optimize one point per time):

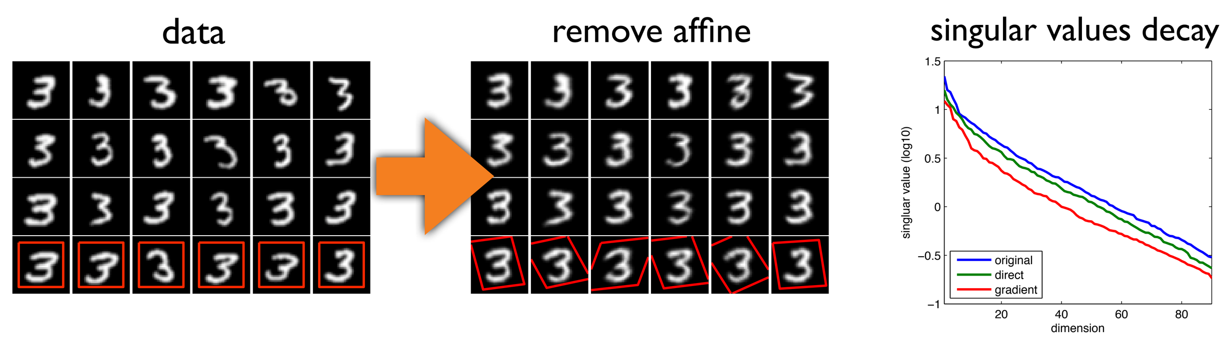
1. Coordinate descent.
2. Gradient descent.
3. Efficient gradient descent by approximating the reconstruction error.

$$\mathcal{D}(x,y) \approx \frac{1}{K}\sum_{k=1}^{K} \frac{\beta_k}{\det A_k} - \frac{1}{\gamma}\sum_{t=1}^{16} \log(-e_l^\top (M\alpha_k + b))$$

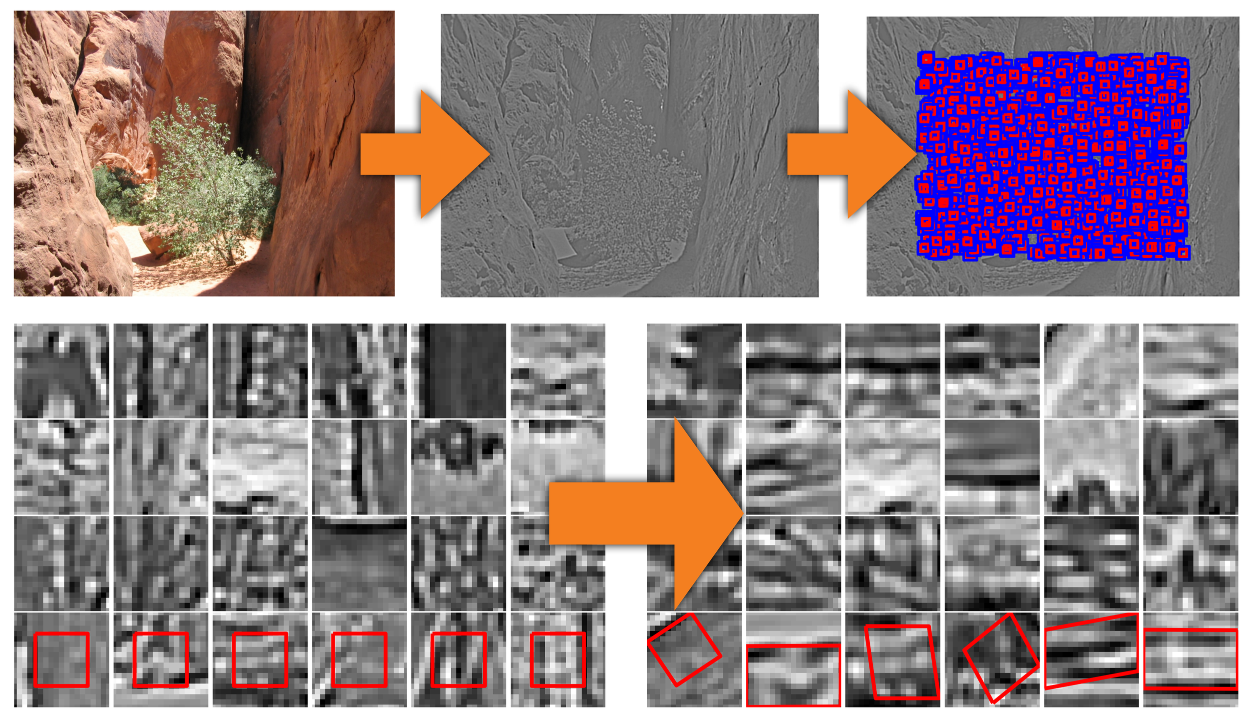
Images and their boundaries. Often neglected, boundaries are an important problem. Solved by padding or by natural extension for image patche.

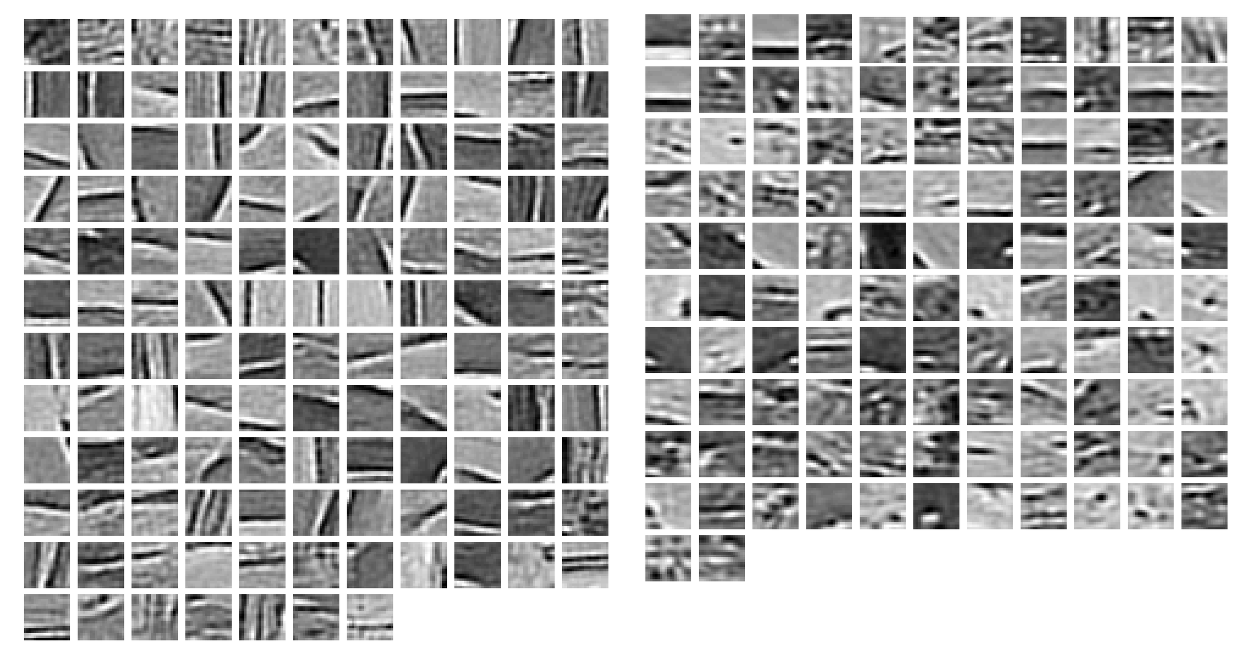
Experiment 1: NIST digits

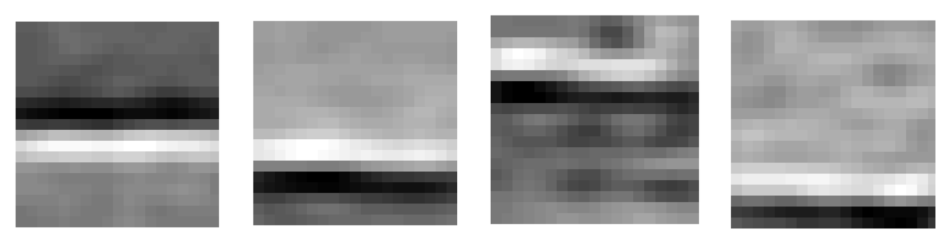
NIST digits (hand-written digits)



Experiment 2: Natural Patches







Conclusions

- Complexity-distortion regularizes IC automatically.
- Complexity can encode and encourage meaningful properties of the data.
- Algorithms can align large dataset efficiently, even if the the data structure is subtle.

References

[IC] E. G. Learned-Miller, "Data driven image models through continuous joint alignment," *PAMI*, vol. 28, no. 2, 2006

[TCA] B. J. Frey and N. Jojic. Transformed component analysis: Joint estimation of spatial transformations and image components. In *Proc. ICCV*, 1999.