Scale-less Dense Correspondences

Tal Hassner
The Open University of Israel

ICCV’13 Tutorial on Dense Image Correspondences for Computer Vision
Matching Pixels

In different views, scales, scenes, etc.

Invariant detectors + robust descriptors + matching
Observation:

Invariant detectors require dominant scales

BUT

Most pixels do not have such scales
Observation:

Invariant detectors require dominant scales

BUT

Most pixels do not have such scales

But what happens if we want dense matches with scale differences?

(c) ANMS 250, $r = 24$

(d) ANMS 500, $r = 16$

[Szeliski’s book]
Why is this useful?

[Hassner&Basri ’06a, ’06b,’13]
Why is this useful?

[Label transfer / scene parsing]

[Liu, Yuen & Torralba '11; Rubinstein, Liu & Freeman' 12 ]
Why is this useful?

[Hassner & Basri '06]

[Depth transfer]

[Karsch, Liu & Kang '12]

[Liu, Yuen & Torralba '11; Rubinstein, Liu & Freeman '12]

Shape by-example

[Tal Hassner
Scale-less dense correspondences]
Why is this useful?
Why is this useful?
Why is this useful?
Dense matching with scale differences

Solution 1:

Ignore scale differences – Dense-SIFT
Dense SIFT (DSIFT)

Arbitrary scale selection

VLFeat.org

Tutorials - DSIFT/PHOW

Home
Download
Documentation
Tutorials
SIFT
DSIFT/PHOW
MSER
IKM
HIKM
AIB
Quick shift
kd-tree
Distance transf.
Utils

VLFeat implements a fast dense version of SIFT, called \texttt{vl_dsift}. The function is roughly equivalent to running SIFT on a dense grid of locations at a fixed scale and orientation. This type of feature descriptors is often used for object categorization.

**Dense SIFT as a faster SIFT**

The main advantage of using \texttt{vl_dsift} over \texttt{vl_sift} is speed. To see this, load a test image:

\begin{verbatim}
I = imread(fullfile(vl_root,'data','a.jpg'));
I = single(vl_imdown(rgb2gray(I)));
\end{verbatim}

To check the equivalence of \texttt{vl_dsift} and \texttt{vl_sift} it is necessary to understand in detail how the parameters of the two descriptors are related.

- **Bin size vs keypoint scale.** DSIFT specifies the descriptor size by a single parameter, \texttt{size}, which controls the size of a SIFT spatial bin in pixels. In the standard SIFT descriptor, the bin size is related to the SIFT keypoint scale by a multiplier, denoted \texttt{magint} below, which defaults to 3. As a consequence, a DSIFT descriptor with bin size equal to 5 corresponds to a SIFT keypoint of scale $5/3 \approx 1.66$.

[Vedaldi and Fulkerson’10]

Tal Hassner
Scale-less dense correspondences
SIFT-Flow

[Liu et al. ECCV’08, PAMI’11]

“The good”: Dense flow between different scenes!
SIFT-Flow

[Liu et al. ECCV’08, PAMI’11]

Left photo  Right photo  Left warped onto Right

“The bad”: Fails when matching different scales
Dense matching with scale differences

Solution 2:

Scale Invariant Descriptors (SID)*

* Kokkinos and Yuille, *Scale Invariance without Scale Selection*, IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2008
Log-Polar sampling

\[ r, \theta \]
From Rotation + Scale to translation
Translation invariance

Absolute of the Discrete-Time Fourier Transform
SID-Flow

Left

Right

DSIFT

SID
SID-Flow

Left

Right

DSIFT

SID
Dense matching with scale differences

Solution 3:

Scale-Less SIFT (SLS)*

Joint work with
Viki Mayzels and Lihi Zelnik-Manor and

SIFTs and Multiple Scales
SIFTs and Multiple Scales

\[ \sigma_1, \ldots, \sigma_k \]

\[
\begin{bmatrix}
    h_{\sigma_1}, \ldots, h_{\sigma_k}
\end{bmatrix}
\]

\[ \sigma_1, \ldots, \sigma_k \]

\[
\begin{bmatrix}
    h'_{\sigma_1}, \ldots, h'_{\sigma_k}
\end{bmatrix}
\]
Observation 1

Corresponding points have multiple SIFT matches at multiple scales
Scale-less dense correspondences

Left image SIFTs

Right image SIFTs
Matching ver. 1

Use set-to-set distance:

\[ dist(p, p') = \min \quad \text{dist}(h_{\sigma_i}, h'_{\sigma_j}) \]
To Illustrate

...if SIFTs were 2D
Observation 2

SIFT changes gradually across scales
Suggests they reside on manifold
Main Assumption

SIFTs in multi-scales lie close to a linear subspace

Fixed local statistics:

$$h_{\sigma_i} = h_{\sigma_j}$$

Gradual changes across scales:

$$h_{\sigma_i} = \sum_j w_{ij} h_{\sigma_j}$$

$$H = \begin{bmatrix} h_{\sigma_1}, \ldots, h_{\sigma_k} \end{bmatrix} = \begin{bmatrix} h_1, \ldots, h_b \end{bmatrix} W = \hat{H} W$$
So, for each pixel...

Extract SIFTs at multi-scales

$$\begin{bmatrix} h_{\sigma_1}, \ldots, h_{\sigma_k} \end{bmatrix}$$

Compute basis (e.g., PCA)

$$\hat{H} = \begin{bmatrix} h_1, \ldots, h_b \end{bmatrix}$$

This low-dim subspace reflects SIFT behavior through scales
Matching ver. 2

Use subspace to subspace distance:

$$\text{dist}(p, q) = \text{dist} (\hat{H}_p, \hat{H}_q) = \| \sin \theta \|^2_2$$
To Illustrate
The Scale-Less SIFT (SLS)

Map these subspaces to points!

[Basri, Hassner, Zelnik-Manor, CVPR’07, ICCVw’09, TPAMI’11]

For each pixel $p$

$$ A_p = \hat{H}_p \hat{H}_p^T $$

$$ SLS(p) = Vec(A_p) = \left[ \frac{a_{11}}{\sqrt{2}}, a_{12}, \ldots, a_{1D}, \frac{a_{22}}{\sqrt{2}}, a_{23}, \ldots, \frac{a_{DD}}{\sqrt{2}} \right] $$

$$ \| SLS(p) - SLS(q) \|^2 = \mu \cdot dist^2 \left( \hat{H}_p, \hat{H}_q \right) $$
The Scale-Less SIFT (SLS)

Map these subspaces to points!
[Basri, Hassner, Zelnik-Manor, CVPR’07, ICCVw’09, TPAMI’11]

A point representation for the subspace spanning SIFT’s behavior in scales!!!

\[
SLS(\mathbf{p}) = \text{Vec}(\mathbf{A}_p) = \begin{bmatrix}
\frac{a_{11}}{\sqrt{2}}, a_{12}, \ldots, a_{1D}, \frac{a_{22}}{\sqrt{2}}, a_{23}, \ldots, \frac{a_{DD}}{\sqrt{2}}
\end{bmatrix}
\]

\[
\|SLS(\mathbf{p}) - SLS(\mathbf{q})\|^2 = \mu \cdot \text{dist}^2(\hat{\mathbf{H}}_p, \hat{\mathbf{H}}_q)
\]
SLS-Flow

Left Photo

DSIFT

Right Photo

SID [Kokkinos & Yuille, CVPR'08]

Our SLS
Dense-Flow with SLS
Using SIFT-Flow to compute the flow

Left Photo

Right Photo

DSIFT

SID [Kokkinos & Yuille, CVPR’08]

Our SLS
Dense-Flow with SLS
Using SIFT-Flow to compute the flow

Left Photo

Right Photo

DSIFT

SID [Kokkinos & Yuille, CVPR’08]

Our SLS
What we saw

Dense matching, even when scenes and scales are different
Thank you!

hassner@openu.ac.il

www.openu.ac.il/home/hassner
References


- **[Hassner, Saban & Wolf]** Hassner, Tal, , Gilad Saban, and Lior Wolf, In submission


Resources

- **SIFT-Flow**

- **DSIFT (vlfeat)**
  - [http://www.vlfeat.org/](http://www.vlfeat.org/)

- **SID**
  - [http://vision.mas.ecp.fr/Personnel/iasonas/code.html](http://vision.mas.ecp.fr/Personnel/iasonas/code.html)

- **SLS**
  - [http://www.openu.ac.il/home/hassner/projects/siftscales/](http://www.openu.ac.il/home/hassner/projects/siftscales/)

- **Me**
  - [http://www.openu.ac.il/home/hassner](http://www.openu.ac.il/home/hassner)
  - [hassner@openu.ac.il](mailto:hassner@openu.ac.il)
Tal Hassner
Scale-less dense correspondences