Deep Self-Correlation Descriptor for Dense Cross-Modal Correspondence

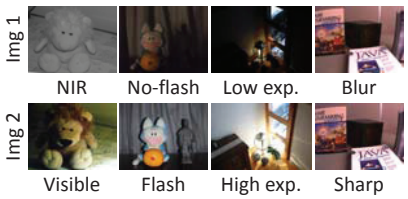
Seungryong Kim[1], Dongbo Min[2], Stephen Lin[3], and Kwanghoon Sohn[1]

[1]Yonsei University, [2]Chungnam National University, [3]Microsoft Research

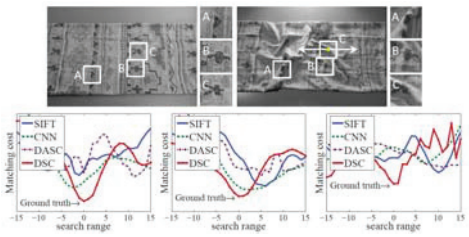
Introduction

❖ Goal

1) To develop dense descriptor for matching multi-modal and multi-spectral images
2) To develop efficient computation model



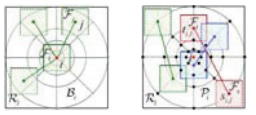
Motivation



cost func. in A | cost func. in B | cost func. in C

✓ **Existing descriptors cannot yield global minima**

Overview



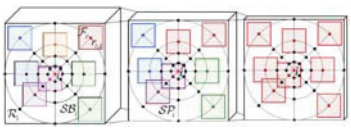
LSS descriptor | DASC descriptor

✓ LSS descriptor

$$\max_{j \in \mathcal{B}_i(l)} \{\exp(-\mathcal{S}(\mathcal{F}_i, \mathcal{F}_j)/\sigma_c)\}$$

✓ DASC descriptor

$$\exp(-(1 - |\mathcal{C}(\mathcal{F}_{s_{i,j}}, \mathcal{F}_{t_{i,j}})|)/\sigma_c)$$

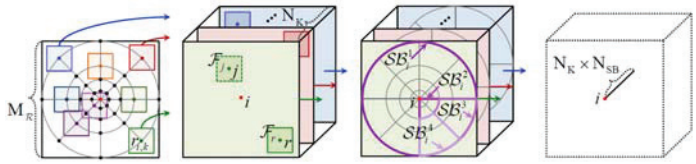


DSC descriptor

✓ hierarchical self-correlation
✓ circular spatial pyramid pooling

Deep Self-Correlation (DSC) Descriptor

❖ SSC: Single Self-Correlation Descriptor



local support window | self-correlation surface | circular spatial pyramid pooling (C-SPP) | SSC descriptor

1) Self-Correlation

- Computing adaptive self-correlation measure

$$\mathcal{C}(\mathcal{F}_r, \mathcal{F}_j) = \frac{\sum_{r',j'} \omega_{r,r'}(f_{r'} - \mathcal{G}_{r,r})(f_{j'} - \mathcal{G}_{r,j})}{\sqrt{\sum_{r'} \omega_{r,r'}(f_{r'} - \mathcal{G}_{r,r})^2} \sqrt{\sum_{r',j'} \omega_{r,r'}(f_{j'} - \mathcal{G}_{r,j})^2}}$$

2) Circular Spatial Pyramid Pooling (C-SPP)

circular spatial pyramid bins

3) Non-linear Gating/ Normalization

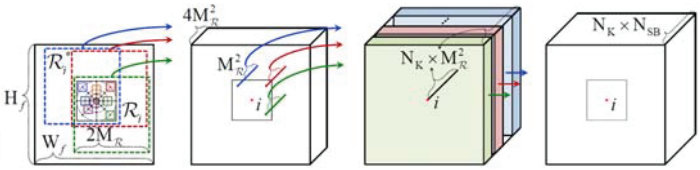
$$d_i^{\mathrm{SSC}}(l) = \exp(-(1 - |\hat{h}_i(l)|)/\sigma_c)$$

❖ Efficient Computation for Dense Description

- Cost reformulation and fast edge-aware filtering (EAF)

$$\mathcal{C}(\mathcal{F}_i, \mathcal{F}_{j_r}) = \frac{\mathcal{G}_{i,ij_r} - \mathcal{G}_{i,i} \cdot \mathcal{G}_{i,j_r}}{\sqrt{\mathcal{G}_{i,i^2} - (\mathcal{G}_{i,i})^2} \cdot \sqrt{\mathcal{G}_{i,j_r^2} - (\mathcal{G}_{i,j_r})^2}}$$

- Pre-computation scheme for self-correlation surface



$$O(IM_{\mathcal{F}}N_K M_{\mathcal{R}}^2) \longrightarrow O(I4M_{\mathcal{R}}^2)$$

$M_F \rightarrow 1$: Using edge-aware filtering (EAF)
$N_K {M_R}^2 \rightarrow 4{M_R}^2$: Using the pre-computation of self-correlation surfaces

❖ DSC: Deep Self-Correlation Descriptor

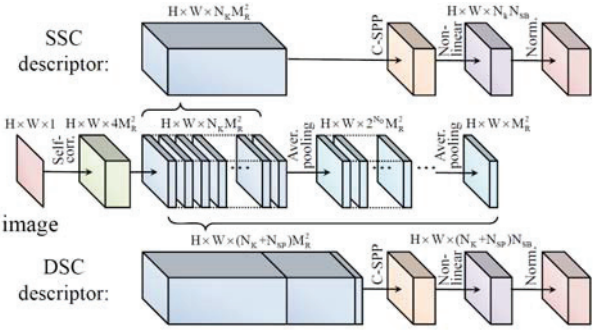
- Extending SSC descriptor by encoding self-similar structures at multiple levels

1) Hierarchical Self-Correlation

- Average Pooling

$$\sum_{r_{i,k} \in \mathcal{SP}_i(v)} \mathcal{C}(\mathcal{F}_{r_{i,k}}, \mathcal{F}_j)/N_v$$

2) Aggregating Hierarchical Self-Correlation



❖ DSC Description Summary

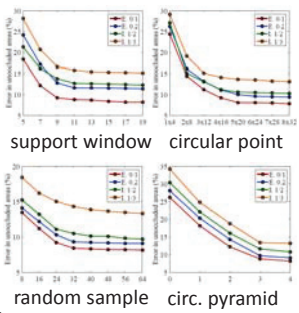
Algorithm 1: Deep Self-Correlation (DSC) Descriptor

Input : image f_i, random samples $r_{i,k}$.
Output : DSC descriptor $\mathcal{D}_i^{\mathrm{DSC}}$.
Parameters : The number of circular pyramidal bins (point sets) $N_{\mathcal{SB}}(N_{\mathcal{SP}})$.

Compute $\mathcal{C}(\mathcal{F}_i, \mathcal{F}_j)$ for a doubled support window $\mathcal{R}_i^*$ by using (6).
Estimate $\mathcal{C}(\mathcal{F}_{r_{i,k}}, \mathcal{F}_j)$ from $\mathcal{C}(\mathcal{F}_i, \mathcal{F}_j)$ according to the index mapping process.
for $v = 1 : N_{\mathcal{SP}}$ **do** /* **hierarchical aggregation using average pooling** */
  Determine a circular pyramidal point $\mathcal{SP}_i(v)$.
  Compute $\mathcal{C}(\mathcal{F}_v, \mathcal{F}_j)$ by using an average pooling for $\mathcal{SP}_i(v)$ on $\mathcal{C}(\mathcal{F}_{r_{i,k}}, \mathcal{F}_j)$.
end for
for $u = 1 : N_{\mathcal{SB}}$ **do** /* **hierarchical spatial aggregation using C-SPP** */
  Determine a circular pyramidal bin $\mathcal{SB}_i(u)$.
  Compute $h_i(k,u)$ and $h_i(v,u)$ by using C-SPP on each $\mathcal{SB}_i(u)$ from $\mathcal{C}(\mathcal{F}_{r_{i,k}}, \mathcal{F}_j)$ and $\mathcal{C}(\mathcal{F}_v, \mathcal{F}_j)$, respectively.
end for
Build hierarchical self-correlation responses $\hat{h}_i(l)$ from $h_i(k,u)$ and $h_i(v,u)$.
Compute a DSC descriptor $\mathcal{D}_i^{\mathrm{DSC}} = \bigcup_l d_i^{\mathrm{DSC}}(l)$, followed by L-2 normalization.

Experimental Results and Discussion

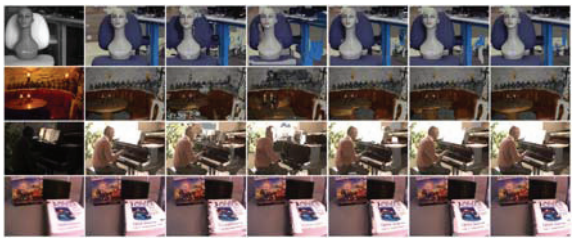
❖ Parameter Evaluation



support window | circular point | random sample | circ. pyramid

❖ Cross-modal and Cross-spectral Benchmark

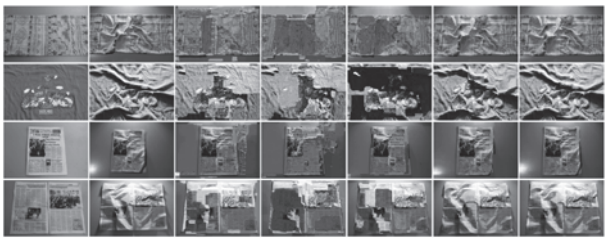
- RGB-NIR, different exposure, flash-noflash, blurring



img1 | img2 | BRIEF | LSS | DASC | SSC | DSC

❖ DaLI Benchmark

- non-rigid deformation and severe illumination changes



img1 | img2 | DAISY | BRIEF | LSS | DaLI | DSC

Conclusion

❖ Deep Self-Correlation (DSC) Descriptor

- For establishing dense correspondences between images taken under different imaging modalities

Contact

- e- mail: srkim89@yonsei.ac.kr
- homepage: http://diml.yonsei.ac.kr/~srkim/

European Conference on Computer Vision (ECCV) 2016