

Supplementary Material for “2D/3D Rotation-Invariant Detection using Equivariant Filters and Kernel Weighted Mapping”

Kun Liu, Qing Wang, Wolfgang Driever, and Olaf Ronneberger

University of Freiburg, Germany

This supplementary material contains more detail about the implementation of the techniques used in the main paper.

Appendix A: 2D covariant HOG-based descriptions Although a histogram-of-oriented-gradient is often shown in a discrete manner, the original information it encodes is a continuous distribution. This distribution is a function of the angle, so we can also use Fourier series to encode the information. In practice, we start the projection on Fourier basis from the gradient field. Let the gradient field of an image be $\mathbf{d}(\mathbf{x})$, with its orientation $\varphi_d(\mathbf{x})$. The “histogram” of the gradient on *a single pixel* is like a Dirac function $\delta(\varphi - \varphi_d(\mathbf{x}))$ of height $\|\mathbf{d}(\mathbf{x})\|$. Projecting it on the Fourier basis $e^{im\varphi}$, we get

$$\hat{d}_m(\mathbf{x}) = \langle \|\mathbf{d}(\mathbf{x})\| \delta(\varphi - \varphi_d(\mathbf{x})), e^{im\varphi} \rangle = \|\mathbf{d}(\mathbf{x})\| e^{-im\varphi_d(\mathbf{x})} .$$

We can convolve a triangular kernel (of width = b) with $\hat{d}_m(\mathbf{x})$ to create $\tilde{d}_m(\mathbf{x})$, which actually represent the HOG with spatial soft-binning. Then, as a matter of choice, we construct the local description from the Fourier HOG fields by sampling $\tilde{d}_{m=\{1,\dots,M\}}(\mathbf{x})$ on two circles around each position, with the radii $r_1 = 0.5b$ and $r_2 = 1.5b$. The sampling will be done by convolving with templates of the form $\delta(r - r_p)e^{in\varphi}$ (the template are shown in Fig.1. The Dirac function is smoothed), as

$$f_{mnp} = \tilde{d}_m * \delta(r - r_p)e^{in\varphi} .$$

What we get is a group of covariant features, encoding the information of HOG in $2b$ range around each point. And **the feature f_{mnp} has the rotation order $n - m$** . So in our framework, this description can be used as the same as the features coming from self-steerable filters.

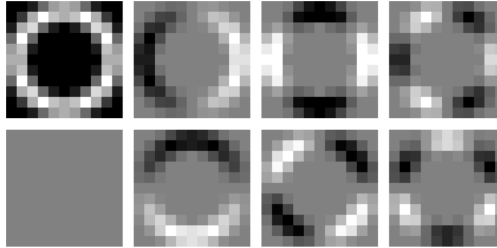


Figure 1: The template $\delta(r - r')e^{in\varphi}$ for sampling the HOG features projected on Fourier basis. Top row: real part. Bottom row: imaginary part

Appendix B: Implementation detail for the 3D Filter Based on the explanation in Sec.3.4 and Eq.(6) in the main paper, the general form of the filter can be rewritten for the 3D setting as

$$S = \sum_j \left(\sum_k \sum_{m_d=m_j} \tilde{\mathcal{K}}_k w_{jkd} F_d \right) \tilde{\bullet}_0 u_j , \quad (1)$$

where the symbol $\tilde{\bullet}_0$ means coupling two spherical tensor fields into a zero-order tensor field by the spherical tensor convolution, which combines the convolution and spherical tensor product. Here, we explicitly write out the terms

which respect equivariance. The constraint $m_d = m_j$ is decided by the nature of the spherical tensor product, *i.e.*, only two spherical tensor fields of the same order can be coupled into a zero-order tensor field by the tensor product (convolution) [1].

Using the SGD basis for description ($\nabla_q^p G_{\sigma_d} : \mathbb{R}^3 \rightarrow \mathbb{C}^{2(p-q)+1}$) and voting ($\nabla^j G_{\sigma_v} : \mathbb{R}^3 \rightarrow \mathbb{C}^{2j+1}$), and taking advantage of the commutativity between convolution and differentiation, we can write the computation procedure on an image V as

$$S = G_{\sigma_v} * \sum_j \nabla_j \left[\sum_{\substack{p,q \\ p-q=j}} \left(\sum_k \tilde{\mathcal{K}}_k w_{jkpq} \right) \nabla_q^p (G_{\sigma_d} * V) \right]. \quad (2)$$

The detection process is shown in Algorithm 1. Refer to the main paper about how to adapt the process for training.

Algorithm 1 The Landmark Detection Scheme

Input: image $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ trained model $\{\hat{\mathbf{f}}_k, h, w_{jkpq}\}$

Output: probability map $y : \mathbb{R}^3 \rightarrow \mathbb{R}(\mathcal{T}_0)$, $y := H(V)$

//compute SGD filtering by taking advantage of the the commutativity between convolution and differentiation

The first convolution $F^0 := G_{\sigma_d} * V$

for $p = 1 : p_{\max}$ **do**

$$F^p = \nabla^1 F^{p-1}$$

for $q = 1 : \min(p_{\max} - p, p)$ **do**

$$F_q^p := \nabla_1 F_{q-1}^p$$

end for

end for

//create rotation-invariant features

$$\mathcal{I}(\mathbf{F}(\mathbf{x})) := [\|F^0(\mathbf{x})\|, \|F^1(\mathbf{x})\|, \|F_1^1(\mathbf{x})\|, \|F^2(\mathbf{x})\|, \|F_1^2(\mathbf{x})\|, \dots]$$

//evaluate the weighting kernel values

$$\tilde{\mathcal{K}}_k(\mathbf{x}) = \frac{\exp(-\|\mathcal{I}(\mathbf{F}(\mathbf{x})) - \hat{\mathbf{f}}_k\|^2/2h^2)}{\sum_{k'} \exp(-\|\mathcal{I}(\mathbf{F}(\mathbf{x})) - \hat{\mathbf{f}}_{k'}\|^2/2h^2)}$$

//assign voting coefficients to each voxel

for $j = 0 : p_{\max}$ **do**

$$\tilde{A}_j(\mathbf{x}) = \sum_k \sum_{\substack{p,q \\ p-q=j}} \tilde{\mathcal{K}}_k(\mathbf{x}) w_{jkpq} F_q^p(\mathbf{x})$$

end for

//carry out the voting (using the commutativity again)

$$\text{Initialize } \mathbf{y}_{p_{\max}} = \underbrace{[0, \dots, 0]}_{p_{\max}} \in \mathcal{T}_{p_{\max}}$$

for $j = p_{\max} : -1 : 1$ **do**

$$\mathbf{y}_{j-1} = \nabla_1(\mathbf{y}_j + \tilde{A}_j)$$

end for

$$y = \mathbf{y}_0 + \tilde{A}_0$$

The second convolution $y = G_{\sigma_v} * y$

Here we further show the computational detail of **spherical (tensor) derivatives** [1], which are used **to create basis functions with spherical harmonics as the angular part**. The spherical up-derivate is computed as $\mathbf{V} = \nabla^1 \mathbf{V}'$, where $\mathbf{V}' : \mathbb{R}^3 \rightarrow \mathbb{C}^{2(\ell-1)+1}$, $\mathbf{V} : \mathbb{R}^3 \rightarrow \mathbb{C}^{2\ell+1}$ are the input and output tensor fields. It is defined as a tensor product between spherical gradient operator $\nabla = [\frac{1}{\sqrt{2}}(\partial x - \mathbf{i}\partial y), \partial z, -\frac{1}{\sqrt{2}}(\partial x + \mathbf{i}\partial y)]$ and a spherical tensor field. The computation of a tensor product needs some real coefficients (called *Clebsch-Gordan* coefficients [2], which depend on the orders of the coupled tensor fields). By indexing the elements of \mathbf{V} and \mathbf{V}' as $\{V_{-\ell}, \dots, V_{\ell}\}$ and $\{V'_{-\ell+1}, \dots, V'_{\ell-1}\}$, the computation rule of ∇^1 is:

$$\begin{aligned} V_m &= c_{(m,m+1,-1)} \frac{1}{\sqrt{2}} (\partial x - \mathbf{i}\partial y) V'_{m+1} \\ &+ c_{(m,m,0)} \partial z V'_m \\ &- c_{(m,m-1,1)} \frac{1}{\sqrt{2}} (\partial x + \mathbf{i}\partial y) V'_{m-1} \quad , \end{aligned} \quad (3)$$

where $c_{(m,m',a)}$ is computed from two *Clebsch-Gordan* coefficients $\frac{\langle \ell m | (\ell-1) m', 1 a \rangle}{\langle \ell 0 | (\ell-1) 0, 1 0 \rangle}$ [2]. When m' goes out of the range $[-\ell + 1, \ell - 1]$, take the value $V'_{m'}$ as zero. Thus, the computation of the spherical derivatives is a group of weighted finite differences.

Those coefficients compute as:

$$c_{(m,m+1,-1)} = \sqrt{\frac{(\ell-m)(\ell-m-1)}{2\ell^2}} \quad , \quad (4)$$

$$c_{(m,m,0)} = \sqrt{\frac{(\ell+m)(\ell-m)}{\ell^2}} \quad , \quad (5)$$

$$c_{(m,m-1,1)} = \sqrt{\frac{(\ell+m)(\ell+m-1)}{2\ell^2}} \quad . \quad (6)$$

Equation (3) also fits the spherical down-derivative $\mathbf{V} = \nabla_1 \mathbf{V}'$, by taking $\mathbf{V} : \mathbb{R}^3 \rightarrow \mathbb{C}^{2(\ell+1)+1}$ and $\mathbf{V}' : \mathbb{R}^3 \rightarrow \mathbb{C}^{2\ell+1}$. Then we need different *Clebsch-Gordan* coefficients for the different coupling orders, $c_{(m,m',a)} = \frac{\langle (\ell+1) m | \ell m', 1 a \rangle}{\langle (\ell+1) 0 | \ell 0, 1 0 \rangle}$:

$$c_{(m,m+1,-1)} = \sqrt{\frac{(\ell+m)(\ell+m+1)}{2\ell^2}} \quad , \quad (7)$$

$$c_{(m,m,0)} = \sqrt{\frac{(\ell+m)(\ell-m)}{\ell^2}} \quad , \quad (8)$$

$$c_{(m,m-1,1)} = \sqrt{\frac{(\ell-m)(\ell-m+1)}{2\ell^2}} \quad . \quad (9)$$

References

- [1] M. Reiser and H. Burkhardt. Spherical tensor calculus for local adaptive filtering. In S. Aja-Fernández, R. de Luis García, D. Tao, and X. Li, editors, *Tensors in Image Processing and Computer Vision*, Advances in Pattern Recognition. Springer, 2009. 2
- [2] M. Rose. *Elementary theory of angular momentum*. Dover Publications, 1995. 2, 3