3D Rotation-Invariant Description from Tensor Operation on Spherical HOG Field

Kun Liu, Henrik Skibbe, Thorsten Schmidt, Olaf Ronneberger
{liu,skibbe,tschmidt,ronneber}@informatik.uni-freiburg.de
Department of Computer Science,
University of Freiburg

Thomas Blein, Klaus Palme
{thomas.blein,klaus.palme}@biologie.uni-freiburg.de
Institute of Biology II - Botany,
University of Freiburg

We present a method to compute rotation-invariant descriptions from 3D volumetric data sets. They are based on local histograms of oriented gradients (HOG). The main contributions of this paper are the continuous representation of these histograms (Fig.2(b)(c)) by spherical harmonics (Fig.2(a)), and their regional description by spherical tensor operations (Fig.2(d)). Existing methods for fast tensor filtering makes the method applicable for large biomedical data sets.

Motivation Most of the existing methods to extend SIFT or HOG feature to 3D data [2] are based on local or global pose normalization. In contrast, there exist analytic methods which guarantee the rotation invariance directly, from Fourier analysis in spherical coordinates [4], where spherical harmonics (*SH*) are used as the angular basis. However, these *SH* based methods have no special ability to deal with object deformation and image distortion, which proves to be important for robust description. So we are motivated to take advantage of some HOG-like feature to enhance the performance of *SH* based descriptions.

Spherical Harmonics and 3D Rotations In 2D, an angular signal on a certain radius (a circle) could be well described by a one-dimensional Fourier transform [4]. In 3D, to describe a signal on a sphere with 3D rotations, the tools we need are the spherical harmonics (*SH*) and the rotation group $SO(3)$ [1, 4]. Spherical harmonics $Y_m^\ell : S^2 \to \mathbb{C}$ form an orthonormal basis for the 2-sphere. Any square-integrable scalar function on a sphere, can be expanded into a linear combination as: $f = \sum_{\ell=0}^{\infty} \hat{\mathbf{f}}^{\ell^\top} \mathbf{Y}^\ell(\theta,\phi)$ where $\mathbf{Y}^\ell(\theta,\phi)$ and $\hat{\mathbf{f}}^\ell \in \mathbb{C}^{2\ell+1}$, ℓ denotes the band of expansion. To analyse the rotation behaviour of expansion coefficients, we also need the so called Wigner D-matrices [1], which are the irreducible representation of a rotation $\mathfrak{g} \in SO(3)$. They are determined by the 3D rotation angles, and denoted by unitary matrices $\mathbf{D}^\ell \in \mathbb{U}^{2\ell+1}$, for the rotation in the ℓ^{th} band. The *SH* expansion coefficients for a rotated function just look like $\mathbf{D}_\mathfrak{g}^\ell \hat{\mathbf{f}}^\ell$.

Spherical HOG Representation A small trick which connects HOG to *SH* framework comes from a simple observation: although a histogram is often shown in a discrete manner, the original information it encodes is a continuous distribution. A typical image of 2D gradient histograms is shown in the left of Fig. 1. In comparison, the representation as circular signals are equal and even more accurate, while only low-order Fourier series are enough to encode them. We will construct a 3D HOG as a

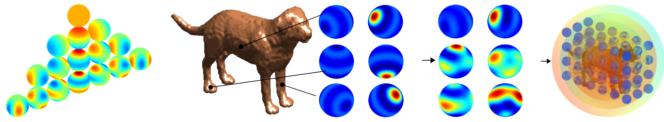
Figure 1: HOG and the equivalent circular signals

signal on the 2-sphere, and rotations could be easily addressed by using *SH* coefficients to represent the signal.

A raw *spherical HOG*, which only describes the gradient distribution at one voxel, should be an oriented impulse signal on the 2-sphere. Let $\theta_d(\mathbf{r}), \phi_d(\mathbf{r}))$ and $d(\mathbf{r})$ be the direction and magnitude of the gradient at the position $\mathbf{r}$, its *SH* coefficients compute to

$$\hat{\mathbf{d}}^\ell = \frac{2\ell+1}{4\pi} < \mathbf{Y}^\ell, d\cdot\delta(\theta_d,\phi_d) > = \frac{2\ell+1}{4\pi} d\mathbf{Y}^\ell(\theta_d,\phi_d) \qquad (1)$$

For the spatial aggregation with "soft binning", we just need to convolve a Gaussian kernel g_σ with all *SH* coefficients component-wisely, as $\tilde{d}_m^\ell = \hat{d}_m^\ell * g_\sigma$. The construction of spherical HOG is illustrated in Fig. 2(b)(c).



(a) (b) (c) (d)

Figure 2: The main ideas: (a) *Spherical harmonics* for representing HOG. (b) Individual gradient as function on a sphere. (c) The spatially aggregated *spherical HOG*. (d) *Spherical tensor* operations create rotation-invariant descriptions from the *spherical HOG field*.

Regional Description by Tensor Operation It can be proved that the spherical HOG coefficient $\tilde{\mathbf{d}}^\ell(\mathbf{r})$ form a rank-ℓ spherical tensor field ($\mathscr{T}_l$) [1, 3]. In SIFT and sliding window techniques, histograms at neighbouring grids are concatenated into a region description. Here we will create regional descriptions by spherical tensor operation. The basic approach to describe a region around one point, is to compute features on multiple concentric shells centred at the selected point. Applying a shell-wise orthogonal tensorial expansion, we can get a large group of expansion coefficients $\hat{\mathbf{a}}_k^j(r,\ell) \in \mathscr{T}_{j+k}$ $(k = -j \sim j)$ for each tensor field $\tilde{\mathbf{d}}^\ell$ on the shell of radius r. For any coefficients of the same rank $(j+k = j'+k')$, they will transform with the same Wigner-D matrix under rotations. Because the Wigner-D matrices are unitary, their effect will be compensated in the complex inner product. So a general formula of our rotation-invariant features is

$$||\sqrt{<\hat{\mathbf{a}}_k^j(r,\ell), \hat{\mathbf{a}}_{k'}^{j'}(r',\ell')>}|| = ||\sqrt{<\mathbf{D}_\mathfrak{g}^{j+k}\hat{\mathbf{a}}_k^j(r,\ell), \mathbf{D}_\mathfrak{g}^{j'+k'}\hat{\mathbf{a}}_{k'}^{j'}(r',\ell')>}|| \qquad (2)$$

when $j+k = j'+k'$. The shell-wise expansion is not efficient for dense volume description. Within the same conception, the Spherical Gaussian Derivative (*SGD*) filtering [3] is a more efficient tool for getting voxel-wise description. After computing SGDs on the tensor fields $\hat{\mathbf{d}}^\ell(\mathbf{r})$, we again collect features by coupling the output of the same rank with inner product, like in Equation 2.

Experiment We demonstrate the effectivity of our description on Princeton Shape Benchmark and SHREC 2009 Generic Shape Benchmark. It achieves 72% nearest-neighbour accuracy on PSB (the accuracy of pure *SH* description and pose-normalized HOG description are 56% and 58%). We also show a biological image analysis application - the structural segmentation in Arabidopsis roots by voxel-wise classification based on the rotation-invariant description (Fig.3).

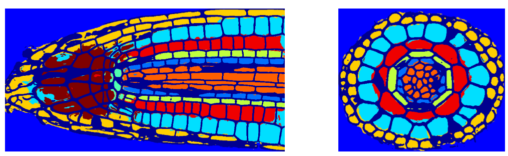


Figure 3: Voxel-wise classification based on rotation-invariant description

[1] D.M. Brink and G.R. Satchler. *Angular momentum*. Clarendon Press, 1962.

[2] G. Flitton, T.P. Breckon, N. Megherbi, and UK Cranfield. Object Recognition using 3D SIFT in Complex CT Volumes. In *BMVC*, 2010.

[3] M. Reisert and H. Burkhardt. Spherical tensor calculus for local adaptive filtering. In *Tensors in Image Processing and Computer Vision*, Advances in Pattern Recognition. Springer, 2009.

[4] Q. Wang, O. Ronneberger, and H. Burkhardt. Rotational invariance based on fourier analysis in polar and spherical coordinates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31: 1715–1722, 2009.