2010 International Conference on Pattern Recognition

Local Rotation Invariant Patch Descriptors for 3D Vector Fields

Janis Fehr
HCI, Heidelberg University, Germany
janis.fehr@iwr.uni-heidelberg.de

***Abstract*—In this paper, we present two novel methods for the fast computation of local rotation invariant patch descriptors for 3D vectorial data. Patch based algorithms have recently become very popular approach for a wide range of 2D computer vision problems. Our local rotation invariant patch descriptors allow an extension of these methods to 3D vector fields. Our approaches are based on a harmonic representation for local spherical 3D vector field patches, which enables us to derive fast algorithms for the computation of rotation invariant *power spectrum* and *bispectrum* feature descriptors of such patches.**

***Keywords*-local feature; 3D vector field; invariance;**

I. Introduction

3D vector fields are a common data representation in computer vision tasks, e.g. carrying 3D flows or 3D gradient information. In this paper, we present two local rotation invariant features, which provide a discriminative representation of local patches for such 3D vector fields.

Parameterization of the Local 3D Vector Field Patches. Given a 3D vector field $\mathbf{X} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, we write $\mathbf{X}[z](\mathbf{x})$ to address the z component of the $(x, y, z)^T$-vector at position $\mathbf{x}$. We parameterize a local spherical patch at position $\mathbf{x} \in \mathbb{R}^3$ as a set of n concentric spherical surfaces $\mathcal{S}[r_i](\mathbf{x})$ with radii r_i and $i \in \{1, \ldots, n\}$ which are centered in $\mathbf{x}$:

$$\mathcal{S}[r_i](\mathbf{x}) := \{\mathbf{x}' \in \mathbb{R}^3 | \|\mathbf{x} - \mathbf{x}'\|_2 = r_i\}. \tag{1}$$

Related Work. To our knowledge, there have not been many notable publications on local rotation invariant features for 3D vector fields. There has been an approach to use Clifford Algebras for a template matching of 3D vector fields [1]. Also, the vector orientation histogram approach as described in the 3D SIFT descriptor [2] could be used. Our approaches are motivated by local rotation invariant features methods which have been introduced in the scalar domain of Spherical Harmonics (see section II). The *power spectrum* of this harmonic representation has been used as a rotation invariant descriptor, e.g. for 3D shape retrieval [3]. A *bispectrum* feature for the Spherical Harmonic domain has been introduced in [4].

II. Mathematical Foundations

We derive an orthonormal base for vectorial signals on the 2-sphere that provides a sound mathematical representation of the parameterized patches in the harmonic domain.

Spherical Harmonics. ($\mathcal{SH}$) [5] form an orthonormal base on the 2-sphere. Analogical to the Fourier Transform, any given real valued signal f on a sphere with its parameterization over the angles Θ, Φ (latitude and longitude of the sphere) can be represented by an expansion in its harmonic coefficients:

$$f(\Phi, \Theta) = \sum_{l=0}^{\infty} \sum_{m=-l}^{m=l} \widehat{f}_m^l Y_m^l(\Phi, \Theta), \tag{2}$$

where l denotes the band of expansion, m the order for the l-th band, $\widehat{f}_m^l$ the harmonic coefficients, $Y_m^l(\Phi, \Theta)$ the harmonic base functions and $\overline{Y_m^l}$ their complex conjugate.

Rotations in $\mathcal{SH}$. Rotations $\mathcal{R}$ in the Euclidean space find their equivalent representation in the Spherical Harmonic domain in terms of the so called Wigner D-Matrices, which form an irreducible representation of the rotation group $\mathcal{SO}(3)$ [5]. For each band l, $D^l(\mathcal{R})$ defines a band-wise rotation in the $\mathcal{SH}$ coefficients:

$$\mathcal{R}f = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \sum_{n=-l}^{l} D_{mn}^l(\mathcal{R}) \widehat{f}_n^l \overline{Y_m^l}. \tag{3}$$

Clebsch-Gordan Coefficients (CG). CG coefficients of the form $\langle lm | l_1 m_1, l_2 m_2 \rangle$ are commonly used for the representation of direct sum decompositions of $\mathcal{SO}(3)$ tensor couplings [5]. The CG define the selection criteria for such couplings, are only unequal to zero if the constraints $m = m_1 + m_2$ and $|l_1 - l_2| \leq l \leq l_1 + l_2$ hold and fulfill several useful orthogonality constraints (see [5] for more details).

Vectorial Harmonics. We derive our methods from a very general theory of Tensorial Harmonics [6], which provides expansions for arbitrary real valued tensor functions $\mathbf{f}$ of order d on the 2-sphere:

$$\mathbf{f}[r, d](\Theta, \Phi) := \sum_{l=0}^{\infty} \sum_{k=-d}^{k=d} \sum_{m=-(l+k)}^{m=(l+k)} \widehat{f}_{km}^l[r] \overline{\mathbf{Z}_{km}^l}(\Theta, \Phi), \tag{4}$$

where $\widehat{f}_{km}^l(r)$ is the expansion coefficient of the l-th band of tensor order d and harmonic order m at radius r. The

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.341

orthonormal Tensorial Harmonic base functions $\mathbf{Z}^l_{km}$ are given as:

$$\mathbf{Z}^l_{km}(\Theta,\Phi) := \mathbf{e}^{(l+k)}_m \circ_l Y^l_m[r](\Theta,\Phi), \tag{5}$$

where $\mathbf{e}^l_m$ are elements of the standard Euclidean base of $\mathbb{C}^{2d+1}$. $\circ_l$ denotes a bilinear form connecting tensors of different ranks:

$$\circ_d : V_{l_1} \times V_{l_2} \to \mathbb{C}^{2d+1}. \tag{6}$$

$l_1, l_2 \in \mathbb{N}$ have to hold $|l_1 - l_2| \leq l \leq l_1 + l_2$ and $\circ_l$ is computed as follows:

$$(\mathbf{e}^l_m)^T(\mathbf{v} \circ_l \mathbf{u}) := \sum_{m=m_1+m_2} \langle lm|l_1m_1, l_2m_2\rangle v_{m_1}u_{m_2}. \tag{7}$$

See [6] for details and proofs. If we limit the general form to tensors of order one ($d := 1$), then we obtain our Vectorial Harmonic expansions

$$\mathbf{f}[r,1](\Theta,\Phi) := \sum_{l=0}^{\infty}\sum_{k=-1}^{k=1}\sum_{m=-(l+k)}^{m=(l+k)} \widehat{f}^l_{km}[r]\overline{\mathbf{Z}^l_{km}}(\Theta,\Phi) \tag{8}$$

with the orthonormal base functions:

$$\mathbf{Z}^l_{km} = \begin{pmatrix} \langle 1 \;\; 1|(l+k) \;\; m,l \;\; (1-m)\rangle Y^l_{1-m} \\ \langle 1 \;\; 0|(l+k) \;\; m,l \;\; -m\rangle Y^l_{-m} \\ \langle 1 \;\; -1|(l+k) \;\; m,l \;\; (-1-m)\rangle Y^l_{-1-m} \end{pmatrix}^T \tag{9}$$

For fields of real valued vectors $(x, y, z) \in \mathbb{R}^3$, we need to transform the vector coordinates to $\mathbb{C}^3$ following the Spherical Harmonic relation:

$$\mathbf{u} \in \mathbb{C}^3 : \mathbf{u} := \begin{pmatrix} \frac{x-iy}{\sqrt{2}} \\ z \\ \frac{-x-iy}{\sqrt{2}} \end{pmatrix}. \tag{10}$$

Rotations in $\mathcal{VH}$. Complex 3D vector valued signals $\mathbf{f}$ with Vectorial Harmonic coefficients $\widehat{\mathbf{f}}$ are rotated [6] by:

$$\mathcal{R}\mathbf{f} = \sum_{l=0}^{\infty}\sum_{k=-1}^{k=1}\sum_{m=-(l+k)}^{l+k}\sum_{n=-(l+k)}^{l+k} D^{l+k}_{mn}(\mathcal{R})\widehat{f}^l_{km}\overline{\mathbf{Z}^l_{kn}}, \tag{11}$$

which is a straight forward extension of (3).

Discrete Vectorial Harmonic Transformation. For practical applications, we need a discrete version of the Vectorial Harmonic transform, i.e. we need to obtain the frequency decomposition of 3D vectorial signals at discrete positions on the discrete spherical surfaces $\mathcal{S}_i[r_i](\mathbf{x}) = \mathbf{X}|_{\mathcal{S}[r](\mathbf{x})}$ given by our parameterization (1) of the local spherical patches in $\mathbf{X} : \mathbb{Z}^3 \to \mathbb{R}^3$. To obtain the discrete Vectorial Harmonic transformation ($\mathcal{VH}$), we pre-compute discrete approximations $\tilde{\mathbf{Z}}^l_{k,m}[r]$ of the orthonormal harmonic base functions $\mathbf{Z}^l_{k,m}[r]$ (9). $\mathcal{VH}$ is then computed as:

$$\left(\mathcal{VH}\left(\mathbf{X}|_{\mathcal{S}[r](\mathbf{x})}\right)\right)^l_{k,m} := \sum_{\mathbf{x}'\in\mathcal{S}[r](\mathbf{x})} \langle \mathbf{X}(\mathbf{x}')\tilde{\mathbf{Z}}^l_{k,m}[r](\mathbf{x}')\rangle. \tag{12}$$

In order to compute the harmonic transformation of the neighborhoods around each voxel $\mathbf{x}$, we perform a fast convolution ($*$) of the pre-computed based functions with the discrete input data:

$$\begin{aligned}\left(\mathcal{VH}[r]\left(\mathbf{X}\right)\right)^l_{k,m} &= \mathbf{X}[x] * \tilde{\mathbf{Z}}^l_{k,m}[r,x] + \mathbf{X}[y] * \tilde{\mathbf{Z}}^l_{k,m}[r,y] \\ &+ \mathbf{X}[z] * \tilde{\mathbf{Z}}^l_{k,m}[r,z].\end{aligned} \tag{13}$$

III. Spectral Feature Descriptors

So far, we showed how to obtain a harmonic representation of spherical 3D vector field patches. It should be noted, that this representation is simply a result of changing base functions, we still have to construct our rotation invariant feature descriptors. Therefore, we take advantage of this frequency representation and adapt two well known signal processing methods, the *power spectrum* and the *bispectrum*, to obtain invariant features.

The Harmonic *Power Spectrum* ($\mathcal{VH}_{pow}$). The computation of the band-wise *power spectrum* feature in the Vectorial Harmonic representation is rather straight forward:

$$\mathcal{VH}_{pow}\left(\widehat{\mathbf{f}}^l_k\right) := \left(\widehat{\mathbf{f}}^l_k\right)^\dagger \widehat{\mathbf{f}}^l_k, \tag{14}$$

where $\widehat{\mathbf{f}}^l_k$ denotes a vector holding the $2(l + k) + 1$ components of the (l, k)th band representing a single frequency.

Rotation Invariance. We following (11), and using the fact that the D^{l+k} are orthonormal, it is easy to show the rotation invariance:

$$\begin{aligned}\mathcal{VH}_{pow}\left(D^{l+k}(\mathcal{R})\widehat{\mathbf{f}}^l_k\right) &:= \left(D^{l+k}(\mathcal{R})\widehat{\mathbf{f}}^l_k\right)^\dagger D^{l+k}(\mathcal{R})\widehat{\mathbf{f}}^l_k \\ &= \left(\widehat{\mathbf{f}}^l_k\right)^\dagger \left(D^{l+k}(\mathcal{R})\right)^\dagger D^{l+k}(\mathcal{R})\widehat{\mathbf{f}}^l_k \\ &= \left(\widehat{\mathbf{f}}^l_k\right)^\dagger \widehat{\mathbf{f}}^l_k.\end{aligned}$$

It should be noted, that *power spectrum* features are ambiguous, i.e. two different signals might have the same *power spectrum*. To overcome this problem, we introduce the complete *bispectrum* features.

The Harmonic *Bispectrum* ($\mathcal{VH}_{bi}$). Extending the complete *bispectrum* feature to the Vectorial Harmonic domain requires a coupling of two harmonic bands $\mathbf{f}^{l_1}_k$ and $\mathbf{f}^{l_2}_k$. Our approach is based on a fundamental ansatz provided by [7], who introduced a generalized *bispectrum* for all compact groups which have a invertible Fourier representation $\widehat{f}(\omega)$:

$$q(f,\omega_1,\omega_2) = C^\dagger\left(\widehat{f}(\omega_1)\otimes\widehat{f}(\omega_2)\right)^\dagger C\bigoplus_\omega \widehat{f}(\omega), \tag{15}$$

where C is a unitary matrix containing the coupling rules for the irreducible representations of the specific group.

While this approach is highly non-trivial for general groups, we find ourself in the convenient situation that the irreducible representation of $\mathcal{SO}(3)$ and its coupling rules are well known in form of Wigner D-Matrices and the CG-Coefficients. Hence, using (11) we rewrite (15) for our harmonic representation as *bispectrum* over $\mathcal{SO}(3)$:

$$\left(\mathcal{VH}_{bi}\left(\widehat{\mathbf{f}}_k^l\right)\right)_k^{l_1,l_2,l} := (C^{l_1,l_2,k})^\dagger \left(\widehat{\mathbf{f}}_k^{l_1} \otimes \widehat{\mathbf{f}}_k^{l_2}\right)^\dagger C^{l_1,l_2,k} \bigoplus_{(l,k)} \widehat{\mathbf{f}}_k^l, \tag{16}$$

where $C^{l_1,l_2,k}$ is a matrix containing the CG-Coefficients of the form

$$\left[C^{l_1,l_2,k}\right]_{(l+k,m)(m_1,m_2)} := \left\langle (l+k)m \middle| (l_1+k)m_1, (l_2+k)m_2 \right\rangle.$$

Since the CG-Coefficients are zero unless $m = m_1 + m_2$, we only need to consider the coefficients where $m_2 = m - m_1$. Exploiting orthonormal properties and the well known facts [5] that tensor products $\widehat{\mathbf{f}}_k^{l_1} \otimes \widehat{\mathbf{f}}_k^{l_2}$ rotate as:

$$\begin{aligned}\left(\widehat{\mathbf{f}}_k^{l_1} \otimes \widehat{\mathbf{f}}_k^{l_2}\right) &\rightarrow \left(D^{l_1+k}(\mathcal{R})\widehat{\mathbf{f}}_k^{l_1}\right) \otimes \left(D^{l_2+k}(\mathcal{R})\widehat{\mathbf{f}}_k^{l_2}\right) \\ &= \left(D^{l_1+k}(\mathcal{R}) \otimes D^{l_2+k}(\mathcal{R})\right)\left(\widehat{\mathbf{f}}_k^{l_1} \otimes \widehat{\mathbf{f}}_k^{l_2}\right) \\ &= (C^{l_1,l_2,k})^\dagger \left[\bigoplus_{l=|(l_1+k)-(l_2+k)|}^{l=(l_1+k)+(l_2+k)} D^l(\mathcal{R})\right] C^{l_1,l_2,k},\end{aligned}$$

we can simplify (16) to

$$\left(\mathcal{VH}_{bi}\left(\mathbf{f}_k^l\right)\right)_k^{l_1,l_2,l} = \left(C^{l_1,l_2,k}\left(\widehat{\mathbf{f}}_k^{l_1} \otimes \widehat{\mathbf{f}}_k^{l_2}\right)\right)^\dagger \widehat{\mathbf{f}}_k^l. \tag{17}$$

Introducing the abbreviation $\widehat{\mathbf{g}}_k^{l,l_1,l_2} := C^{l_1,l_2,k}\left(\widehat{\mathbf{f}}_k^{l_1} \otimes \widehat{\mathbf{f}}_k^{l_2}\right)$, which rotates according to (17):

$$\widehat{\mathbf{g}}_k^{l,l_1,l_2} \rightarrow D^{l+k}(\mathcal{R})\widehat{\mathbf{g}}_k^{l,l_1,l_2}, \tag{18}$$

we follow (17) and implement the elements of $\widehat{\mathbf{g}}_k^{l,l_1,l_2}$ as

$$\begin{aligned}\widehat{g}_{km}^{l,l_1,l_2} = &\sum_{m_1=-(l_1+k)}^{l_1+k} \overline{\widehat{\mathbf{f}}_{km_1}^{l_1}\widehat{\mathbf{f}}_{k(m-m_1)}^{l_2}} \\ &\left\langle (l+k)m \middle| (l_1+k)m_1, (l_2+k)(m-m_1) \right\rangle.\end{aligned} \tag{19}$$

This leads to the final implementation of the Vectorial Harmonic *bispectrum*:

$$\begin{aligned}&\left(\mathcal{VH}_{bi}\left(\mathbf{f}\right)\right)_k^{l_1,l_2,l} := \left(\widehat{g}_k^{l,l_1,l_2}\right)^\dagger \widehat{\mathbf{f}}_k^l \\ &= \sum_{m=-(l+k)}^{l+k} \sum_{m_1=-(l_1+k)}^{l_1+k} \overline{\widehat{\mathbf{f}}_{km_1}^{l_1}\widehat{\mathbf{f}}_{k(m-m_1)}^{l_2}}\widehat{\mathbf{f}}_{km}^l \\ &\left\langle (l+k)m \middle| (l_1+k)m_1, (l_2+k)(m-m_1) \right\rangle,\end{aligned} \tag{20}$$

Rotation Invariance. Given the rotational properties (18) of $\widehat{\mathbf{g}}_k^{l,l_1,l_2}$, the rotation invariance follows directly as in (15):

$$\begin{aligned}\mathcal{VH}_{bi}\left(D^{l+k}(\mathcal{R})\widehat{\mathbf{f}}_k^l\right) &= \left(D^{l+k}(\mathcal{R})\widehat{\mathbf{g}}_k^{l,l_1,l2}\right)^\dagger D^{l+k}(\mathcal{R})\widehat{\mathbf{f}}_k^l \\ &= \left(\widehat{\mathbf{g}}_k^{l,l_1,l_2}\right)^\dagger \left(D^{l+k}(\mathcal{R})\right)^\dagger D^{l+k}(\mathcal{R})\widehat{\mathbf{f}}_k^l \\ &= \left(\widehat{\mathbf{g}}_k^{l,l_1,l_2}\right)^\dagger \widehat{\mathbf{f}}_k^l.\end{aligned}$$

Patch Descriptors. Given our our parameterization (1), we describe local 3D vector field patches by 1D feature vectors that depend on several parameters: the radii r_i and the maximum expansion band b_{max}. Hence, $\mathcal{VH}_{pow}$ feature vectors have $b_{max}+1$ entries per radius while the number of $\mathcal{VH}_{bi}$ feature vectors entries is growing quadratically with b_{max}.

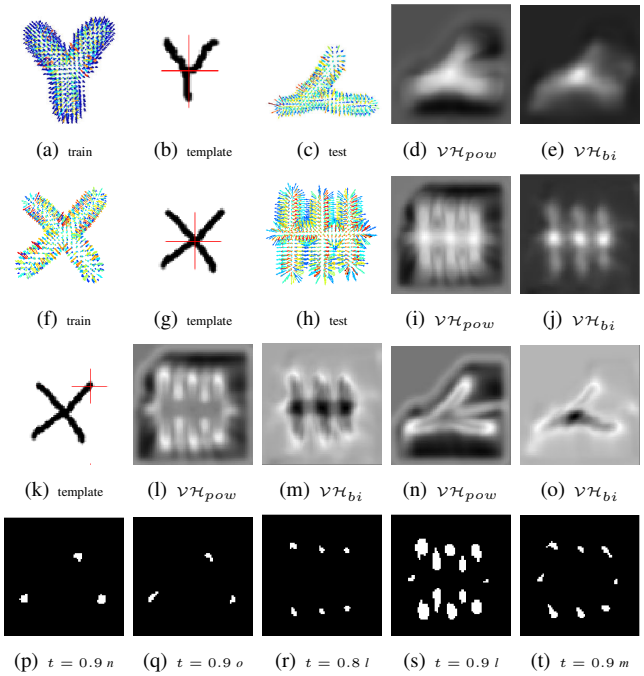
IV. Experiments

We evaluate the proposed features in a series of template matching experiments with artificial "toy data". These experiments are solely designed to provide a proof of concept and to discuss the general properties of the features. Due to space limitations, we have to focus on the theoretical aspects of our features and are unfortunately not able to present further experiments with real world data in this paper.

Toy Data. We obtained a set of discrete 3D vector fields which were obtained by applying the *gradient vector flow* algorithm [8] on manually generated scalar 3D volume data. In all experiments, we only used the directional information, hence we normalized the magnitude to 1. Figure 1 shows the experimental setup and results of the "toy data" experiments.

Experimental Setup. We divided the generated data into test and training samples, where we manually selected the positions of template patches in the training data (see figure 1 (b), (g) and (k)). We extracted the $\mathcal{VH}_{pow}$ and $\mathcal{VH}_{bispec}$ feature vectors at all positions in the test data and used the $L1$-norm to compute the similarity to the template feature vector. The patches were parameterized (1) in four radii and expanded to $b_{max} = 3$.

Results. The results given in figure 1 clearly support our our initial assumption that both features provide discriminative rotation invariant descriptors of local 3D vector field patches. It also showed that the phase information, which is preserved by the complete $\mathcal{VH}_{bispec}$ feature is indeed valuable. Figures 1 (r)-(t) illustrate an example where the $\mathcal{VH}_{bispec}$ feature is sufficient to detect all template like structures (t) with a simple threshold, while this is not possible using the $\mathcal{VH}_{pow}$ feature.
Figure 2 shows the impact of the maximum expansion band the matching results. Our empirical results indicate that it



(a) train (b) template (c) test (d) $\nu\mathcal{H}_{pow}$ (e) $\nu\mathcal{H}_{bi}$

(f) train (g) template (h) test (i) $\nu\mathcal{H}_{pow}$ (j) $\nu\mathcal{H}_{bi}$

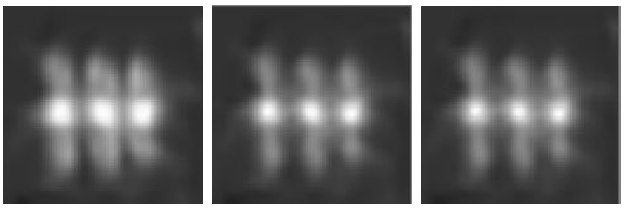
(k) template (l) $\nu\mathcal{H}_{pow}$ (m) $\nu\mathcal{H}_{bi}$ (n) $\nu\mathcal{H}_{pow}$ (o) $\nu\mathcal{H}_{bi}$

(p) $t = 0.9$ *n* (q) $t = 0.9$ *o* (r) $t = 0.8$ *l* (s) $t = 0.9$ *l* (t) $t = 0.9$ *m*

Figure 1. **Experiment I**: The red cross in (b) marks the position of the "Y-bifurcation" template patch in a xy-slice of the training data (a) which we try to locate in the test data (c). A xy-slice of template matching results for the $\mathcal{VH}_{pow}$ feature are shown in (d), and a xy-slice of template matching results for the $\mathcal{VH}_{bispec}$ feature in are shown (e). **Experiment II**: the detection of a "X-bifurcation" template is shown in images (f)-(j). **Experiment III**: the detection of a "terminal" template is shown in images (k)-(o). Images (p)-(t) show the maximum intensity projections of the results of the thresholding t on (l)-(o).

is sufficient to use low expansions (up to the third band) for local patches with small radii.

Conclusions. We presented two novel local rotation invariant features for 3D vector fields. Our experiments indicate that both features are potent descriptors. The complete $\mathcal{VH}_{bispec}$ feature outperforms the $\mathcal{VH}_{pow}$ feature which allows ambiguities. However, the $\mathcal{VH}_{pow}$ feature has only linear complexity in b_{max}, while the complexity of $\mathcal{VH}_{bispec}$ is quadratic.



(a) $b_{max} = 1$ (b) $b_{max} = 3$ (c) $b_{max} = 5$

Figure 2. Impact of the maximum expansion band b_{max} on the accuracy of the feature response. Results for **experiment II** with varying b_{max}.

References

[1] J. Ebling and G. Scheuermann, "Clifford convolution and pattern matching on vector fields," in *Proc. IEEE Conf. on Visualization*, 2003, pp. 193–200.

[2] P. Scovanner, S. Ali, and M. Shah, "A 3-dimensional sift descriptor and its application to action recognition," in *MULTIMEDIA '07: Proceedings of the 15th international conference on Multimedia*. New York, NY, USA: ACM, 2007, pp. 357–360.

[3] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The princeton shape benchmark," in *Shape Modeling International, Genova, Italy*, 2004.

[4] R. Kondor, "A complete set of rotationally and translationally invariant features for images," *CoRR*, vol. abs/cs/0701127, 2007.

[5] D. Brink and G. Satchler, *Angular Momentum, Second Edition*. Clarendon Press, Oxford, 1968.

[6] M. Reisert and H. Burkhardt, "Efficient tensor voting with 3d tensorial harmonics," in *CVPR Workshop on Tensors, 2008, Anchorage, Alaska*, 2008.

[7] R. Kakaraka, "Triple correlation on groups," PhD Thesis, UC Ervine., 1992.

[8] C. Xu and J. Prince, "Gradient vector flow: A new external force for snakes," in *Proc. IEEE Conf. on Comp. Vis. Patt. Recog. (CVPR), Los Alamitos*, 1997, pp. 66–71.