CIRCULAR FOURIER-HOG FEATURES FOR ROTATION INVARIANT OBJECT DETECTION IN BIOMEDICAL IMAGES

Henrik Skibbe and Marco Reisert

Dept. of Diagnostic Radiology, Medical Physics, University Medical Center, Freiburg, Germany

henrik.skibbe@uniklinik-freiburg.de, marco.reisert@uniklinik-freiburg.de

ABSTRACT

In this paper we present a new system for generic rotation invariant 2D object detection based on circular Fourier HOG features. Our system combines the advantages of a dense voting scheme as it is used in the Holomorphic Filter framework with features based on local orientation statistics. Experiments on two different biological datasets show superior detection performance over four state-of-the-art reference approaches.

1. INTRODUCTION

In biomedical images the reliable detection of structures is quite challenging. In contrast to many detection tasks concerning our daily life, like pedestrian detection for surveillance purposes or object detection in car safety systems, scale invariance plays often a minor role in biomedical images. The exact size of objects is usually given by the image acquisition technique. In many biomedical images objects and organisms are sought to be located and analyzed in any number, at arbitrary positions, and, unlike in the pedestrian scenario, in every orientation.

In this paper we introduce a new way to compute and represent HOG (histograms of oriented gradients) features. We show how to use them to build up a trainable filter that can be used for rotation invariant object detection in 2D images.

HOG features [1] are widely used for e.g. object detection or for solving point matching problems [2] because they can be densely computed efficiently [3] and are highly discriminative. However, they neither show rotation invariance nor show a well defined rotation behavior. While being sensitive to rotations is often a wanted feature for many computer vision tasks it hinders a detection of arbitrary oriented organism in biological images.

We overcome the problem by representing circular HOG-features in Fourier domain. This offers a well defined rotation behavior (a rotation is just a multiplication with a complex number) while still allowing a dense computation and showing a very discriminative representation of local image features. Upon these features we build a new framework for 2D object detection in 2D images that comprises the Holomorphic Filter [4] framework. Since Fourier HOG are related to SHOG (spherical HOG) [5] which can be used for 3D object detection in volumetric images, we call this framework the **CHOG**-Filter, since it is based on **C**ircular **HOG**-features. In contrast to SHOG, CHOG is much less computational expensive and since it deals with 2D images it might be useful for a much broader community.



Fig. 1. a) Representation of circular HOG (here 32 bins). b) Fourier HOG with increasing number of coefficients: it describes the local gradient orientation statistic very precisely without discretization artifacts. When neglecting higher order coefficients we still have a valid, continuous representation of the histogram which is not the case for the standard representation a).

The basic idea of the resulting system is that densely computed local HOG features are steering a voting function pixel by pixel. This means, each pixel votes for an object position hypothesis by casting votes in a certain direction based on the orientation and appearance of the local HOG feature. The idea is highly related to the generalized Hough transform [6] and implicit shape models [7]. Such voting based systems are widely used because they can deal with partial occlusions and high intra class variations which is of particular interest when working with biological data.

The CHOG-Filter densely computes HOG features and densely casts votes to ensure that no objects have been left out. Moreover, the proposed approach learns a discriminative voting scheme based on the Holomorphic Filter [4] that actively suppresses responses to regions not supporting an object hypothesis. The CHOG-Filter can be trained in a discriminative manner to detect arbitrary shaped structures.

We exemplarily demonstrate the effectiveness of the approach in two detection tasks on biological images. In a di-

The authors are indebted to the Baden-Württemberg Stiftung for the financial support of this research project by the Eliteprogramme for Postdocs.

rect comparison to existing approaches (among them SIFT [8]), our new filter reveals superior performance.

2. FOURIER CIRCULAR HOG

Local descriptors based on orientation histograms, such as SIFT [8] and HOG [1], have revolutionized detection and matching in natural 2D images. Recently in particular HOG found its way in many applications because it can be computed efficiently and shows excellent performance.

What we propose here is a computation and representation of circular HOG in Fourier domain offering a well defined rotation behavior while still allowing a dense, efficient computation. Since Fourier HOG encodes the same information like ordinary HOG they both encode local image features in the same very discriminative way.

It is worth noting that the literature differs between R-HOG (rectangular spatial window) and C-HOG (circular, isotropic window) [1]. Since the rotation of objects plays an important role in our framework, we only consider the latter one. Given an image $f : \mathbb{R}^2 \to \mathbb{R}$. We denote a dense field of Fourier CHOG descriptors defined over the whole image domain as $\text{CHOG}\{f\} : \mathbb{R}^2 \times S_1 \to \mathbb{R}$, where S_1 denotes the unit-circle.

For capturing only the structure in a pixel's surrounding a window function $w : \mathbb{R}^2 \to \mathbb{R}$ is required. Such a window function is e.g. the 2D Gaussian function. We compute a local CHOG at position $\mathbf{x} \in \mathbb{R}^2$ by collecting all magnitudes of gradients within the window function w contributing to orientation $\mathbf{n} \in \mathbb{R}^2, \|n\| = 1$ according to the continuous distribution function

$$\text{CHOG}\{f\}_w(\mathbf{x}, \mathbf{n}) = \int_{\mathbf{r}\in\mathbb{R}^2} \|\mathbf{g}(\mathbf{r})\| \delta_{\mathbf{n}}(\hat{\mathbf{g}}(\mathbf{r})) w(\mathbf{x}-\mathbf{r}) d\mathbf{r} \ , \quad (1)$$

where $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{g} = \nabla f$ is the gradient field of the image f, $\hat{\mathbf{g}} := \mathbf{g}/\|\mathbf{g}\|$, $\hat{\mathbf{g}} : \mathbb{R}^2 \to S_1$ the gradient orientation field and $\mathbf{n} \in S_1$ is the current histogram entry (the direction) taken into account. $\delta_{\mathbf{n}} : S_1 \to \mathbb{R}$ denotes the Dirac delta function on the circle that selects those gradients out of $\mathbf{g}$ with orientation $\mathbf{n}$. In the following we consider unit-length vectors $\mathbf{n} = (x, y)^T \in \mathbb{R}^2$, $\|\mathbf{n}\| = 1$ w.l.o.g as points on the unit-circle $\mathbf{n} \in S_1$. We equivalently can represent $\mathbf{n}$ as angle ϕ, where $\phi = \text{atan2}(y, x)$.

In contrast to the standard representation of HOG we propose to keep the histogram continuous and realize the "binning" in the frequency domain by using the 1D periodic Fourier basis. We gain the following advantages: First, no interpolation is required because the descriptor is based on the true continuous distribution function. Furthermore, if the window function is isotropic, the descriptor rotates with respect to rotation of its underlying data without leading to any discrete binning artifacts in the histogram. This plays a very important role when aiming at detecting objects at any position and in any orientation using the Holomorphic Filter framework.

2.0.1. Computing Circular HOG in Fourier Domain

Since a CHOG is a function on the circle we can represent $\text{CHOG}_w(\mathbf{x})$ in terms of the orthogonal (periodic) circular Fourier basis functions $\mathbf{e}^\ell(\mathbf{n}) = \mathbf{e}^\ell(\phi) = e^{-\phi i \ell}$, namely

$$\text{CHOG}\{f\}_w(\mathbf{x}, \mathbf{n}) = \tfrac{1}{2\pi} \textstyle\sum_{\ell=-\infty}^{\infty} \mathbf{a}_w^\ell(\mathbf{x})\, \mathbf{e}^\ell(\mathbf{n}) \quad , \quad (2)$$

where $\mathbf{a}_w^\ell(\mathbf{x}) \in \mathbb{C}$ are the complex valued expansion coefficients completely representing the CHOG at image position $\mathbf{x}$ in the Fourier domain. When neglecting higher frequency components by limiting the band $\ell \leq L$ we obtain the best-approximation of CHOG_w in the finite subspace spanned by $\{\mathbf{e}^0 \dots \mathbf{e}^L\}$.

In figure 1 we exemplarily depict the band limited expansion of a gradient histogram for an increasing number of frequency components.

To compute the coefficients $\mathbf{a}_w^\ell(\mathbf{x})$ we plug the circular expansion of the Dirac delta function $\delta_{\mathbf{n}}(\mathbf{n}') = \sum \frac{\overline{\mathbf{e}}^\ell(\mathbf{n}')\mathbf{e}^\ell(\mathbf{n})}{2\pi}$ in eq. (1) and get

$$\begin{aligned} &\text{CHOG}\{f\}_w(\mathbf{x}, \mathbf{n}) = \textstyle\int_{\mathbf{r}\in\mathbb{R}^2} \|\mathbf{g}(\mathbf{r})\| \delta_{\mathbf{n}}(\hat{\mathbf{g}}(\mathbf{r})) w(\mathbf{x}-\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} \underbrace{\left(\left(\|\mathbf{g}\| \, \overline{\mathbf{e}^\ell}(\hat{\mathbf{g}}) * w \right)(\mathbf{x}) \right)}_{=\mathbf{a}_w^\ell(\mathbf{x}) \in \mathbb{C}} \mathbf{e}^\ell(\mathbf{n}) \quad . \end{aligned} \quad (3)$$

Hence we can densely compute the coefficients $\mathbf{a}_w^\ell : \mathbb{R}^2 \to \mathbb{C}$ representing the expansion coefficients of order ℓ for all image positions by convolving the 'higher order' gradient orientation states $\|\mathbf{g}(\mathbf{r})\|\mathbf{e}^\ell(\hat{\mathbf{g}}(\mathbf{r})) \in \mathbb{C}$ with the window function w component by component.

We can archieve robustness against nonlinear illumination and contrast changes by amplifying the influence of gradient orientations and suppressing the influence of the gradient magnitudes by introducing an initial gamma correction of the gradient:

$$\mathbf{g}_\gamma := \|\mathbf{g}\|^\gamma \hat{\mathbf{g}} \ , \text{ where } \gamma \in (0, 1]. \quad (4)$$

For densely computing higher order gradient orientation states $\mathbf{e}^\ell(\hat{\mathbf{g}})$ in an efficient manner we utilize complex multiplications pixel by pixel. Important in our case is that we can recursively derive Fourier basis functions of order $\ell + 1$ by multiplying two basis functions of order ℓ and 1 with $\mathbf{e}^{\ell+1} = \mathbf{e}^\ell \cdot \mathbf{e}^1$, for $\ell \geq 1$. Utilizing this property we gain a recursive rule with which we avoid an explicit, expensive computation of $\mathbf{e}^\ell(\hat{\mathbf{g}}(\mathbf{n}))$, namely $\mathbf{e}^{\ell+1}(\hat{\mathbf{g}}) = \mathbf{e}^\ell(\hat{\mathbf{g}}) \cdot \mathbf{e}^1(\hat{\mathbf{g}})$. The basis case is the gradient in complex notation: $\|\mathbf{g}\|\mathbf{e}^1(\hat{\mathbf{g}}) = \frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}$.

Once the gradient orientation states $\mathbf{e}^\ell(\hat{\mathbf{g}})$ are computed, the remaining computations are the convolutions with the window function w that can be realized efficiently by utilizing the Fast Fourier Transform.

2.1. The CHOG-Filter: Dense Voting with CHOG

The Holomorphic Filter [4] is a nonlinear polynomial filter that is designed to detect arbitrary structures in volumetric images. The most important characteristic of this filter is a trainable voting scheme. The scheme comprises local image features to train a voting function such that the filter responds only to certain structures of interest while responses to all remaining structures in the image are actively suppressed. This is achieved by learning a voting scheme in an initial training step by providing a reference image together with a binary-valued label image. The Holomorphic Filter uses Gaussian derivatives as local features (in complex notation). In the voting step these features are combined in a nonlinear way. The involved free coefficients are the filter parameters that are optimized during the training step. Since the angular expansion coefficients $\mathbf{a}_w^\ell$ of the CHOG obey the same rotation behavior like the Gaussian derivatives in the original filter, we replace the Gaussian derivatives by the CHOG features.

In addition to a non-linear combination of all expansion coefficients $\mathbf{a}_w^\ell$ (following [4]) we compute and combine coefficients derived from different window functions $w_{n=\{1,\cdots,M\}}$ (an angular cross-correlation of different local CHOG). According to section B in [4] the expansion coefficients of the voting functions are now

$$h^\ell[\mathbf{a}^0_{w_0},\cdots,\mathbf{a}^L_{w_M}] := \sum_{\substack{\ell_0+\ell_1=\ell, l\le L \\ n,m\in\{1,\cdots,M\}}} \alpha^{w_0,w_1,w_2}_{\ell_0,\ell_1,\ell} \overline{\mathbf{a}^{\ell_0}_{w_0}} \cdot \overline{\mathbf{a}^{\ell_1}_{w_1}} \cdot \mathbf{a}^\ell_{w_2} \quad (5)$$

where $\alpha^{w_0,w_1,w_2}_{\ell_0,\ell_1,\ell} \in \mathbb{R}$ are the coefficients (the filter parameters) that are learned in an initial training step. Thanks to the rotation preserving characteristic of the complex multiplication the expansion coefficients h^ℓ are "rotating" according to the rotation acting on the coefficients $\mathbf{a}_w^\ell$ and thus the filter's voting function is rotating, too. Consequently the filter response on an image f rotates smoothly with respect to the underlying image data.

It is worth noting that once the coefficients $\mathbf{a}_w^\ell$ have been computed, the application of the filter only requires one convolution per window function. Considering our experiments were we have an expansion up to order $L = 6$ and two different window functions we have $2(L + 1)$ convolutions for the features and 2 convolutions for the filter part. All remaining operations are local differentiations and point-wise complex multiplications. The whole system needs for a 166×158 image about 0.3 seconds using an unoptimized Matlab implementation [1].

3. EXPERIMENTS

For evaluating the performance of the CHOG-filter we aim at detecting landmarks and organisms in biological images in two experiments. For a comparison to existing methods we use the same datasets as in [4]:

[1] System: Intel i7 870, 2.93 GHz

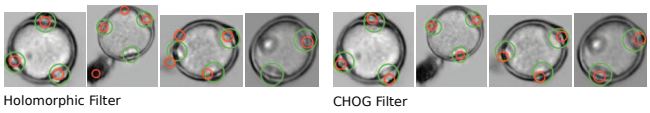


Fig. 2. Comparing the detection performance of Holomorphic Filters and our CHOG Filter. The green circles are indicating the ground-truth, the red circles the local maxima considered as detection in the experiments. The CHOG-Filter clearly suppresses responses to dust and other unwanted structures.

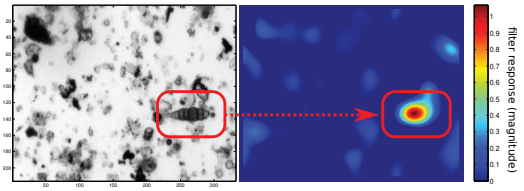


Fig. 3. We get a clear response despite the cluttered background (here a spore image left, response right)

Detection of Pollen-Porates: Palynology, the study and analysis of pollen, is an interesting topic with very diverse applications like in paleoclimatology or forensics. In the first experiment we aim at detecting porates in pollen grains [9] , small pores on the surface of the grain which are crucial for the determination of the species. We define a detection to be successful if the local maxima of the filter response is at most ten pixels apart from the labeled center (a porate has a length of about 40 pixels). All local maxima of the filter responses are collected as detection hypotheses. The filter strength at the putative detection sites are assigned to each hypothesis. This database consists of 150 segmented pollen grains with about 500 porates at all.

Detection of Fungal Spores: Asthma is one of the major respiratory diseases. There may be multiple factors, but it seems that sensitivity to Alternaria spores plays an important role with the onset asthma in certain areas [10]. This makes the counting, detection and forecasting of Alternaria an important task. In this experiment we used four spores in four images for training and a very challenging test set of 50 manually labeled spores in 20 images (details and images can be found in [4]).

We compared our method with the following four approaches (see [4] for implementation details and parameters)
SIFT & PCASIFT: A probabilistic voting procedure as done in the implicit shape model [7] based on SIFT features [8].
Invfeat: We extract a dense set of rotation invariant features based on the power spectrum of an expansion in terms of complex Gaussian derivatives. For each pixel we classify them whether they are an object center or not.
Holo: The original Holomorphic Filter [4]

The proposed CHOG-Filter: For the CHOG-Filter we must determine the following parameters: A filter degree $L \in$

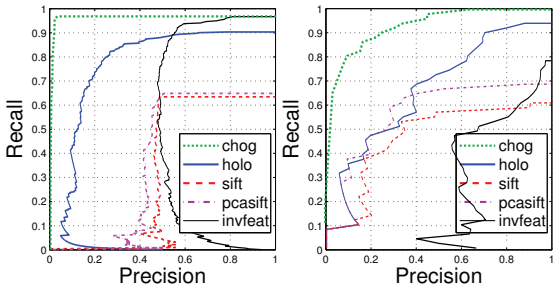


Fig. 4. Precision/Recall graphs for the pollen data set (left) and the spore data set (right).

$\mathbb{N}$ that limits the number of expansion coefficients. Furthermore, we need to set a parameter $\eta \in \mathbb{R}_{>0}$ that steers the scale of the voting function [4] or in other words, the size of a Gaussian window that restricts the CHOG features that can contribute to a local filter response. For the pollen and spore experiment we used $L = 6$ and $\eta = 12$ which we determined experimentally using the training set.

Another issue is the choice of the window function w. We first tested isotropic Gaussian windows of different scales showing promising results. However, finally, we found that the idea to use nested circles with different radii $d \in \mathbb{R}_{>0}$ worked best in our context. The radial profile of the circles is Gaussian smoothed ($e^{\frac{-(r-d)^2}{2\sigma^2}}, \sigma \in \mathbb{R}_{>0}$) ensuring that the corresponding CHOG descriptors are neither suffering from discretization effects nor from small deformations. For the pollen experiment we used two window functions with $w_1 := \{d{=}0, \sigma{=}2.5\}$ and $w_2 := \{d{=}2.5, \sigma{=}2.5\}$. For the spore experiment we used $w_1 := \{d{=}0, \sigma{=}3\}$ and $w_2 := \{d{=}4, \sigma{=}3\}$.

We observed that our filter performs best when using a gamma corrected gradient with $\gamma = 0.8$ (see eq. (4)).

Figure 4 shows the PR graphs for the pollen and spore experiment, respectively. We additionally show qualitative results of the CHOG-Filter and the Holomorphic Filter in figure 2. The CHOG-Filter clearly responds for the correct porate positions. All remaining regions of a pollen are successfully suppressed. This was also true for the spore experiments (see figure 3). Moreover, our Fourier CHOG-Filter significantly outperforms all reference approaches, manly due to two reasons: First, it benefits from the highly discriminative representation of local image patches in terms of gradient orientation histograms. Furthermore, it highly benefits from the dense computation of features and the dense voting scheme ensuring that no objects are left out during the evaluation step.

4. CONCLUSIONS

In this paper we presented an efficient way to compute dense circular Fourier-HOG (CHOG) descriptors. Upon theses descriptors we have built the CHOG-Filter, a dense voting scheme for generic 2D object detection. Our system combines the advantages of a dense voting scheme with features based on local orientation statistics.

We have shown the superior detection performance of our filter compared to state-of-the-art rotation invariant 2D object detection methods. Our approach outperforms other dense voting schemes based on wavelet-like features and shows better performance than sparse voting systems based on orientation statistics. *Source-code of the filter is publicly available*.

5. REFERENCES

[1] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. of the CVPR*, 2005, pp. 886–893.

[2] T. Brox and J. Malik, "Large displacement optical flow: descriptor matching in variational motion estimation," *IEEE Trans. on PAMI*, vol. 33, no. 3, pp. 500–513, Mar. 2011.

[3] Qiang Zhu, Qiang Zhu, Shai Avidan, Shai Avidan, Mei chen Yeh, Mei chen Yeh, Kwang ting Cheng, and Kwang ting Cheng, "Fast human detection using a cascade of histograms of oriented gradients," in *Proc. of the CVPR*, 2006, pp. 1491–1498.

[4] M. Reisert and H. Burkhardt, "Equivariant holomorphic filters for contour denoising and rapid object detection," *IEEE Trans. Image Processing*, vol. 17, no. 2, pp. 190–203, 2008.

[5] H. Skibbe, M. Reisert, and H. Burkhardt, "SHOG - Spherical HOG descriptors for rotation invariant 3D object detection," in *Proc. of the DAGM*, Frankfurt, Germany, 2011.

[6] Dana H. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[7] Bastian Leibe, Ales Leonardis, and Bernt Schiele, "Combined object categorization and segmentation with an implicit shape model," in *In ECCV workshop on statistical learning in computer vision*, 2004, pp. 17–32.

[8] David G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, November 2004.

[9] O. Ronneberger, Q. Wang, and H. Burkhardt, "3D invariants with high robustness to local deformations for automated pollen recognition," in *Proc. of the DAGM,*, Heidelberg, Germany, 2007, pp. 455–435, LNCS, Springer.

[10] R.K. Bush and J.J. Prochnau, "Alternaria-induced asthma.," *J Allergy Clin Immunol*, vol. 113, no. 2, pp. 227–34, 2004.