A Polar Representation of Motion and Implications for Optical Flow

Yair Adato[1] Todd Zickler[2] Ohad Ben-Shahar[1]

adato@cs.bgu.ac.il zickler@seas.harvard.edu ben-shahar@cs.bgu.ac.il

[1]Computer Science Department, Ben-Gurion University of The Negev, Beer Sheva, Israel
[2]Harvard School of Engineering and Applied Sciences, Cambridge, MA, USA

Abstract

We explore a polar representation of optical flow in which each element of the brightness motion field is represented by its magnitude and orientation instead of its Cartesian projections. This seemingly small change in representation provides more direct access to the intrinsic structure of a flow field, and when used with existing variational inference procedures it provides a framework in which regularizers can be intuitively tailored for very different classes of motion. Our evaluations reveal that a flow estimation algorithm that is based on a polar representation can perform as well or better than the state-of-the-art when applied to traditional optical flow problems concerning camera or rigid scene motion, and at the same time, it facilitates both qualitative and quantitative improvements for non-traditional cases such as fluid flows and specular flows, whose structure is very different.

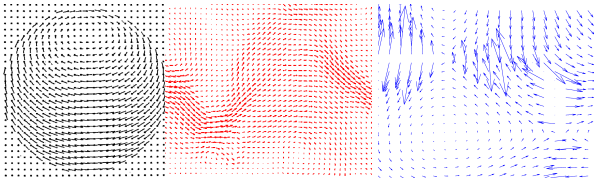
1. Introduction

Understanding motion is important to both human and computational vision systems. To understand scene motion, most computational systems estimate optical flow – the field of apparent velocities of the brightness pattern on the image plane. Since the physical causes of motion can be quite diverse (Fig. 1), inferring optical flow can be a difficult task.

In their seminal study, Horn and Schunck [16] represented the optical flow field using the velocity components in a Cartesian coordinate system:

$$\textbf{flow(x,y)} = (u(x,y), v(x,y)) \triangleq (\frac{dx}{dt}, \frac{dy}{dt}) \,. \qquad (1)$$

Naturally, the representation itself does not determine the flow, for which Horn and Schunck [16] applied their (now classical) *constant brightness assumption*. However, since this assumption alone cannot resolve the flow between consecutive frames due to the aperture problem, additional global constraints are needed to render the optical flow task well-posed. Hence, ever since Horn and Schunck [16], the study of optical flow estimation has focused almost exclusively on the design of proper regularized objective functions whose optimization provides a flow instance in Cartesian representation.



a. **b.** **c.**

Figure 1. Various motion fields. **a.** Piecewise smooth motion flow as would be measured from a dynamic scene involving typical rigid objects (Replicated from [16]). **b.** Fluid motion. This particular motion field represents air currents over North America. Observe the turbulence and occasionally pseudo chaotic behavior. **c.** Specular flow, i.e., the optical flow that is created due to reflection from specular objects. Observe the very large velocities and the unique singularities in certain regions.

Unfortunately, however, designing objective functions for optical flow is often considered an art and relies primarily on intuition[1]. Still, progressing over time, this type of exploration has indeed improved flow estimation performance on standard motion flows [3], though less successfully so on non-traditional ones (e.g., see [2, 12]).

This paper aspires to take optical flow estimation research to a different direction. Inspired by Marr's discussion on the importance of representations [17], we ask whether *qualitative* improvement in optical flow estimation and objective function design can be achieved by representing our objects of study differently. In particular, we propose representing image velocities in terms of their polar coordinates (magnitude and orientation) instead of their Cartesian ones. As advocated by Marr, we show that this not only "makes certain [flow] properties explicit", but it also "greatly affects how easy it is to do certain things with it" [17, page 21]. In particular, we show how different

[1]Notable exceptions include physics-based regularizers for fluids [14] and regularizers learned entirely from training data [19].

classes of motion flow fields exhibit qualitatively different statistical properties in this new representation (as opposed to the Cartesian one), and perhaps more importantly, how switching the representation provides access to a flexible and powerful family of regularizers that both simplify and generalize the design process of objective functions. By adjusting the magnitude and orientation terms differentially, we can create objective functions that lead to good performance not only on traditional (rigid motion) flows, but on more complicated and qualitatively-structured flows such as those induced by fluid motion (Fig. 1b) or specular objects (Fig. 1c).

Finally, we also note that an important property of our approach is the flexibility it offers. Adjustments to the regularizers for different motion classes can be affected by modifying only a few lines of computer code. We show the experimental advantages of this approach using three different datasets: the Middlebury benchmark [3], a specular flow benchmark [2], and synthetic images in the spirit of fluid motion benchmark [15]. To facilitate further scholarly exploration, all our source code is made available for public use via email to the authors.

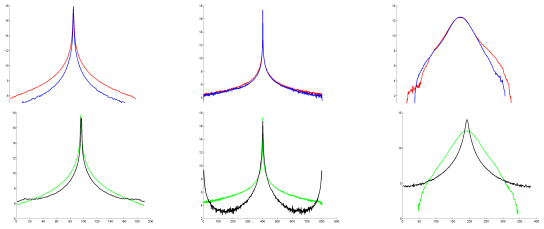
2. Related work

The optical flow literature is far too extensive and diverse to allow an exhaustive review here. As mentioned above, modern optical flow research began in the early 1980s when Horn and Schunck [16] combined the brightness constancy assumption with a regularizer that expresses a hypothetical *piecewise smoothness* behavior, i.e., the assumption that the flow's derivatives are zero almost everywhere. The two assumptions were combined to form a single energy functional

$$E(u,v) = \int\int_{\Omega} \begin{array}{l} (I(x+u,y+v,t+1) - I(x,y,t))^2 + \\ \alpha(|\nabla u|^2 + |\nabla v|^2) dx dy \end{array} \quad (2)$$

where $I(x,y,t)$ is the image intensity at pixel $(x,y) \in \Omega$ in time t and α controls the strength of regularization. With this functional defined, the goal of the computational algorithm is to find the vector field (u,v) that minimizes it, for example by applying the Euler-Lagrange equations or alternative methods (e.g. [7,24,26]). Over the years, optical flow algorithms have also incorporated various enhancements to improve accuracy such as coarse-to-fine strategies to deal with local minima (e.g., [8, 18]), median filtering to reduce the noise in the flow (e.g., [20,24]), robust penalty functions to handle outliers (e.g., [5]), and texture decomposition to minimize effects of lightning variations (e.g., [23, 24]).

Somewhat surprisingly, an algorithm based on the basic energy function in Eq. 2, combined with modern minimization approaches and other enhancements, can estimate the optical flow of traditional motion sequences (e.g., as in the Middlebury benchmark) fairly accurately [20]. However, the regularization term proposed by Horn and Schunck evaluates to zero only for flows with zero divergence, and for this reason, it is unsuitable for fluids and other non-traditional cases [2, 12]. The same is true for alternative regularizers based on differential properties of the flow [21]. Of course, special-purpose regularizations can be designed for these cases, and examples for fluids include approaches that minimize second-order divergence and curl [14, 25] or employ the Helmholtz decomposition [13]. However, these methods are often much more difficult to optimize and implement, they are harder to use, and they do not transfer easily to new types of flows.



'Traditional' optical flows [19] **Specular flows [2]** **Complex fluid motion [15]**

Figure 2. Spatial derivative histograms (log scale) of optical flow in different databases. In all figures red and blue lines denote the gradient magnitude of the horizontal ($|\nabla u|$) and vertical ($|\nabla v|$) components, respectively. The black and green lines represent the gradient magnitude of the flow orientation($|\nabla \theta|$) and flow magnitude ($|\nabla m|$), respectively. Note how the polar representation reveals hitherto neglected statistical properties of optical flows, which can be leveraged for its estimation from image sequences.

3. Motivation for representational shift

To the best of our knowledge, all optical flow algorithms represent the flow using Horn and Schunck's classic Cartesian representation (Eq. 1). However, there are no *physical* reasons to choose that particular representation of a vector field. In fact, vectors, and vector fields, are more naturally represented by their magnitude and orientation components which indeed carry distinct information and have different significance in the context of motion description. In this paper we therefore suggest to employ a polar representation for optical flows:

$$\textbf{flow(x,y)} = (m(x,y), \theta(x,y)) \; . \quad (3)$$

The rest of this section is dedicated to justifying this choice more rigorously, before moving on to formulate its application for flow estimation algorithms. We do note that preliminary conceptual advantages of representing optical flow and image gradients in polar coordinates have been briefly discussed by Schunk in the early days of optical flow research [20]. Unfortunately, however, his early remarks have never evolved into more rigorous computational framework, and in particular, to our best knowledge, no optical flow algorithm based on polar representation has ever been suggested.

3.1. Flow's statistics

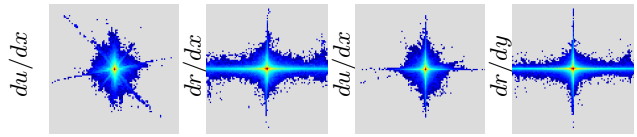
Recently, the spatial statistics of optical flow have been investigated and exploited for estimation algorithms by Roth and Black [19]. Here we repeated their study and compared the statistics of various optical flow classes in their Cartesian and polar representations. In particular, we used three different databases of qualitatively different motions. First, we used the Roth and Black [19] synthetic database (thereafter DB-A) that was created from a collection of depth maps of typical scenes. Second, we used the imaging model by Adato *et al*. [1] and created 200 flows (DB-B) that simulate the observed motions, or specular flows, in scenes containing highly specular (mirror-like) objects. Finally, we used 100 synthetic fluid flows (DB-C) from the FLUid project [15], most of which are characterized by high divergence and turbulence which are common in fluid dynamics. For all flows, we measured the distribution of their spatial derivatives, averaged over all samples within each class.

As is evident from the results in Fig. 2, the differences between classes are far more pronounced in the proposed polar representation. Whereas in the classical Cartesian representation the distributions of the spatial derivatives of the two flow components (u and v) share the same statistics in all databases (Fig. 2 - first row), significant differences are revealed when we switch to the polar representation. See, for example, how magnitude and orientation show qualitatively different distribution in DB-C, or how the orientation component in DB-B becomes bimodal[2] as opposed to the frequently encountered long tail behavior in the Cartesian representation.

The differences between classes in the polar representation immediately suggest the use of different regularizations for each class of flows. Since the polar representation reveals significant statistical differences between the two polar coordinates, we conclude that these two components should be treated separately and differently (an insight further strengthened by the very different scales of these coordinates).

Roth and Black [19] also investigated the statistical independence between the spatial derivatives of the horizontal and vertical components of the flow. They obtained empirical joint histograms of derivatives of the flow and the mutual information (MI) between the various derivatives. In the same spirit, here we calculated the MI and the joint histograms of derivatives of the flow represented in polar coordinate using database DB-A. Figure 3 shows that the spatial derivatives of a flow exhibit larger statistical independence when it represented in polar coordinates (compare to the corresponding results in [19]).

[2]Recall that orientation is a periodic function, and hence that the left and right ends of the graph connect.



	dv/dx	$d\theta/dx$	dv/dy	$d\theta/dy$
$MI =$	0.040	0.023	0.015	0.007

Figure 3. Joint log-histograms of derivatives of optical flow in polar coordinates and their mutual information (MI). Small MI values indicate approximate statistical independence.

In summary, we argue that while certain important aspects of optical flows are concealed in the classical Cartesian representation, they become explicit in the polar representation. This information can be used as prior knowledge for motion estimation by facilitating better energy models and regularizers fit better the class of motions at hand. In the bulk of this paper we study these advantages in greater depth.

3.2. Flexibility of the framework

In addition to providing natural coordinate system and revealing hidden structure, the polar representation has the significant advantage of easily incorporating physical properties of the flow at hand (should these become known in advance) within the common variational framework for flow estimation. For example, consider a case where the flow's orientation tends to change significantly faster than its magnitude (as is the case in complex fluid flows), or the case when flow magnitude grows rapidly while the orientation remains relatively stable (as is often the case in specular flow). Obviously, better flow estimation in these cases would require treating each of the magnitude and orientation differently. Doing this in the Cartesian representation comes with a price – it is much more complicated to isolate them using the Cartesian (u, v) coordinates. In the polar representation this would be done in a straight forward way, for example by smoothing the first derivative of the first component and the second derivative of the second (see Fig. 5 and 6 and Fig. 7). As we show later, this flexibility is a natural byproduct of the proposed representation.

3.3. Performance evaluation

As already discussed in the literature [2–4], the commonly used angular error (AE) and the end point (EP) measures are problematic in terms of evaluating the performance of flow estimation algorithms. For example, AE is biased toward small flows compared to large flows, i.e., the same orientation error in the estimated flow is played down in the AE measure as the magnitude of the flow increases [3, 4]. Moreover, both AE and EP could easily obscure essential aspects of the estimation process, and in particular, two estimated optical flows having similar AE or EP error levels might be qualitatively different [2].

Recently, an alternative approach has been proposed by Adato *et al*. [2], who argued that the performance of optical flow algorithms should be evaluated based on *separate* consideration of the orientation (O) and magnitude (M) components of the estimated flow, namely via an *Orientation Error* (OE) map and a *Magnitude Error* (ME) map. Each of these two error maps can be further averaged (spatially) to obtain a two dimensional error descriptor (AOE, AME).

Following these arguments, we further observe that performance analysis based on orientation and magnitude become even more meaningful once the estimation process itself is based on the same dimensions, possibly leading to better understanding of the source of errors, and as a result, to better estimation algorithms.

4. A variation framework for Optical flow estimation in polar representation

In this section we develop a general framework for optical flow estimation using the proposed polar representation. First, note that formulating the brightness constancy and a smoothness regularization into an energy functional based on polar representation is quite straight forward:

$$\begin{aligned} &E(\theta(x,y), m(x,y)) = \\ &\int_\Omega \psi(I(x + m\cos\theta, y + m\sin\theta, t+1) - I(x,y,t)) \\ &\quad +\alpha_\theta \psi_\theta(\rho_\theta(\theta))) + \alpha_m \psi_m(\rho_m(m)) d\mathbf{x} \end{aligned} \tag{4}$$

where $\theta(x,y)$ and $m(x,y)$ are the orientation and magnitude of the flow, respectively, ψ_θ and ψ_m are robust penalty functions (which, as suggested above, could be different based on prior statistics), and ρ_θ and ρ_m are differential operators which can be selected for the different terms and components according to the class of flow or application of interest. (For the sake of a simpler exposition, most of our derivations assume ρ to be a first order derivative, but arbitrary extensions are of course possible.) α_θ and α_m are weight parameters that control the desired regularization strength in the orientation or the magnitude components, a decision that could based on the properties of the estimated flow or the application of interest.

To use Eq. 4 under the polar representation, we realize that m must be non-negative. To avoid this difficulty we define the following equivalence relation $\sim$ over the values of m and θ

$$(m,\theta) \sim \begin{cases} (m,\theta) & m > 0 \\ (-m, \theta+\pi) & m < 0\,. \end{cases} \tag{5}$$

To avoid the problems due to the periodic nature of θ, we further re-formulate Eq. 4 by over parameterizing the orientation component. We define

$$\left.\begin{aligned} s(x,y) &= \sin\theta(x,y) \\ c(x,y) &= \cos\theta(x,y) \end{aligned}\right\} \text{ subject to } \; s^2 + c^2 = 1 \tag{6}$$

where the constraint $s^2 + c^2 = 1$ is referred to as the *coherence constraint*. Now we can re-write Eq. 4 as a constrained minimization problem

$$\begin{aligned} E(c,s,m) = \; & \int_\Omega \psi(I(x+mc, y+ms, t+1) - I(x,y,t)) \\ & +\alpha_\theta \psi(|\nabla c|^2, |\nabla s|^2) + \alpha_m \psi(|\nabla m|^2) d\mathbf{x} \\ \text{subject to} \quad & s^2 + c^2 = 1. \end{aligned} \tag{7}$$

We stress that the coherence constraint is a key component of the minimization problem since it ensures that variables s and c become genuine representation of orientation.

To minimize Eq. 7 one can use a Lagrange multiplier λ and solves a true four-parameter (c,s,m,λ) minimization problem. While this is possible, in this paper we take a simplified heuristic approach and let λ be a pre-determined parameter (which possibly varies spatially). Of course, doing so comes with a risk. λ values too small will cause a strong deviation from the constraint $s^2 + c^2 = 1$. On the other hand, values too high may cause the estimated optical flow to fall into a local minimum far from the desired results.

Naturally, one can reason about "good" values for λ. When $s^2 + c^2$ is close to 1 the value of λ should be as small as possible to allow the rest of the terms to express themselves more significantly. As the sum of $s^2 + c^2$ departs from 1, we would like λ to grow and enforce the coherence constrain more forcefully. Hence, we suggest that λ be a pixel-wise pre-determined parameter that is updated before each iteration of the minimization process to reflect the insights just discussed. We therefore set

$$\lambda = e^{(s^2+c^2-1)^2}, \tag{8}$$

evaluated at the previous iteration values of c and s. With this definition we now obtain the following energy functional:

$$\begin{aligned} E(c,s,m) = & \int_\Omega \lambda(s^2 + c^2 - 1)^2 \\ & +\psi(I(x+mc, y+ms, t+1) - I(x,y,t)) \\ & +\alpha_\theta \psi(|\nabla c|^2, |\nabla s|^2) + \alpha_m \psi(|\nabla m|^2)\; d\mathbf{x}\,. \end{aligned} \tag{9}$$

With Eq. 9 as our final flow model, our goal is finding a flow (c,s,m) that minimizes the incremental version (using coarse-to-fine approach) of Eq. 9. Conveniently, at this point, we can also take full advantage of the complete library of "enhancements" that the vision community has developed for variational Cartesian formulations.

5. Minimization in polar coordinates

The polar representation raises several issues that one must consider when formalizing the optimization process. As always, to estimate the optical flow one should find the minimizer (c,s,m) of the (incremental version of the) energy functional in Eq. 9. To represent the corresponding Euler-Lagrange equations more clearly we employ the following notations (similar to Brox *et al*. [7]):

$$\begin{gathered} I_* \triangleq \partial_* I(x+mc, y+ms, t+1), \quad * \in x, y \\ I_z \triangleq I(x+mc, y+ms, t+1) - I(x,y,t) \\ \psi'_d \triangleq \psi'(I_z^2) \quad \psi'_{cs} \triangleq \psi'(|\nabla c|^2, |\nabla s|^2) \quad \psi'_m \triangleq \psi'(|\nabla m|^2) \end{gathered}$$

where I_z is the derivative in the direction of the flow. With these notations, the necessary (Euler-Lagrange) conditions for the minimization of Eq. 9 become

$$\begin{aligned} \psi'_d m I_z I_x + \lambda c(c^2+s^2-1) &= \alpha_\theta div(\psi'_{cs} \cdot \nabla c) \\ \psi'_d m I_z I_y + \lambda s(c^2+s^2-1) &= \alpha_\theta div(\psi'_{cs} \cdot \nabla s) \\ \psi'_d I_z(cI_x + sI_y) &= \alpha_m div(\psi'_m \cdot \nabla m). \end{aligned} \quad (10)$$

Clearly, Eq. 10 is a highly non-linear system. To linearize it we employ a coarse-to-fine strategy by building an image pyramid where in each level the incremental dc, ds and dm are calculated. The Taylor expansions (where k is pyramid level) for this case become

$$\begin{aligned} I_z^{k+1} &= I_z^k + I_x^k d(mc)^k + I_y^k d(ms)^k \\ &= I_z^k + I_x^k(cdm + mdc)^k + I_y^k(sdm + mds)^k \ . \end{aligned}$$

The second order derivatives can be either ignored or resolved using fixed point iteration, and in our present implementation we ignore it. Representing the linear expressions in matrix form, we obtain the following sparse system:

$$A \begin{pmatrix} dc \\ ds \\ dm \end{pmatrix} = - \begin{pmatrix} \psi'_d I_x I_z + \lambda(c^3 + cs^2 - c) + O(d^2) \\ \psi'_d I_y I_z + \lambda(cs^2 + s^3 - s) + O(d^2) \\ \psi'_d I_z(I_x c + I_y s) + O(d^2) \end{pmatrix} \quad (11)$$

where $\mathbf{A}$ is

$$\mathbf{A} \triangleq \mathbf{D}^T\mathbf{D} + \lambda \begin{pmatrix} 3c^2+s^2 & 2cs & 0 \\ 2cs & 3s^2+c^2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{matrix} \alpha_\theta \\ -\alpha_\theta \\ \alpha_m \end{matrix} \begin{pmatrix} div\psi'_{cs}\nabla(s+ds) \\ div\psi'_{cs}\nabla(c+dc) \\ div\psi'_m\nabla(m+dm) \end{pmatrix},$$

$\mathbf{D} = \psi'_d(I_x m, I_y m, I_z + I_x c + I_y s)$, and $\lambda = e^{(s^2+c^2-1)^2}$. As before, α_θ and α_m are the smoothness parameters of the orientation and magnitude, respectively. Interestingly, the three parts of matrix $\mathbf{A}$ correspond to the data, coherence, and smoothness terms. The system itself can be solved using standard solvers such as successive over-relaxation (SOR).

Unlike previous optical flow algorithms, the optimization procedure cannot be initialized at zero. In polar coordinates, points specified by $c = s = 0$ are not well defined, they are unstable for numerical purposes, and they do not satisfy coherence constraint. Hence, we initialize by $c = s = \sqrt{2}/2$.

Importantly, since the orientation is periodic, the interpolation between pyramid levels should be implemented with care. For examples, cubic interpolation can not be used in a naive way because otherwise the absolute value of c and s may exceed one. Similarly, the median filtering (which is known as a crucial tool in many optical flow algorithm [20]) must be extended to circular functions, a topic we consider important future work. Finally, as mentioned above, the interpretation of magnitude subject to the equivalence relation in Eq. 5, which compensates for negative values, must be applied.

As we argue, although the numerical issues are non-trivial, the proposed framework and the polar representation hold much promise for optical flow estimation. As we show in the next section, even without rigorous handling of some of these issues, the results of our suggested approach are at least as good as the state-of-the-art on traditional optical flows, and they exceed it both qualitatively and quantitatively for non-traditional optical flows such as those induced by specular scenes or fluid motions.

6. Applications and results

In order to explore the potential of the suggested framework, an algorithm based on the new representation was implemented and compared to the state-of-the-art in optical flow algorithms [10, 20]. As mentioned above, we focused on three different types of flows, and used three flow databases of different characteristics:

1. DB-A: The Middlebury optical flow benchmark [3].
2. DB-B: The Ben-Gurion University specular flow database [2]
3. DB-C: our own set of synthetic images rendered from fluid flow vector field in the spirit of the FLUid project [15].

As we shall see, the suggested method based on the polar representation outperforms the state-of-the-art optical flow algorithms in all of the three benchmarks.

6.1. Performance on traditional optical flows

Among these three datasets, the Middlebury database is the most popular and focuses on relatively traditional scenes while avoiding specularities, transparencies, or non rigid moving objects. Hence, the observed motion fields follow the piecewise smoothness assumption and exhibit no outstanding velocities, structural singularities (except for line discontinuities due to occlusions), turbulence, or chaotic regions. Since our goal was to explore what are the possible benefits (if any) of the the polar representation, we compared our suggested method to the state-of-the-art implementation by Sun *et al.* [20], while using their very own optimization approach, median filtering heuristic, and texture decomposition for our method as well. This comparison is summarized in Table 1 in the *Classic NL* and *Polar Rep. NL* rows. For the sake of comparison we also bring the results from another public domain implementation of state-of-the-art optical flow algorithm due to the gpu4vision group [10], although we note that their objective function and optimization techniques are different from ours, making it difficult to isolate the effect of flow representation in the comparative performance measures.

The results of the seven Middlebury public ground truth sequences are displayed in Table 1. In each box the errors

	Dimetrodon	Hydrangea	RubberWhale	Grove2	Grove3	Urban2	Urban3
Classic NL	0.051, 0.216 0.244	**0.035**, 0.255 0.285	**0.054**, 0.093 **0.115**	**0.026**, 0.151 0.188	**0.096**, 0.826 0.970	0.114, 5.378 5.450	0.183, 3.546 3.699
Polar Rep. NL	**0.045**, **0.089** **0.138**	0.048, **0.221** **0.234**	0.105, **0.070** 0.131	0.044 , **0.097** **0.183**	0.124 , **0.425** **0.647**	**0.104**, **0.352** **0.502**	**0.072**, **0.444** **0.579**
TL-V1	0.053, 0.112 0.166	0.042, **0.132** 0.185	0.080, 0.072 0.120	0.056, 0.134 0.237	0.126, 0.457 0.680	**0.054**, **0.323** **0.377**	0.103, 0.417 0.624

Table 1. Results of our suggested method and two state-of-the-art optical flow algorithms [10, 20]. First line in each box shows the average orientation and magnitude error. Second line show the standard end point error. Numbers highlighted in green emphasize best results among the classic NL and Polar Rep. NL. Highlighted in blue are best results when obtained by the TL-V1 algorithm (though as mention in the text, this comparison is much less informative). Note that the polar based algorithm outperforms the equivalent algorithm.

were calculated in two ways – first by the average orientation and magnitude error pair (top of each entry), and then by the standard end point error [4]. As can be seen from the table, the suggested optical flow algorithm outperforms the equivalent algorithm in most of the sequences and since the only difference between these cases is the representation (and all that is entailed by it), we can attribute the better results to the polar representation. Note in particular that in most of the sequences the magnitude components are estimated significantly more accurately than in the algorithm based on the Cartesian presentation.

6.2. Performance on specular flows

Motion estimation from image sequences of scenes with highly specular (mirror like) objects provides optical flows with unique properties. The accurate estimation of this "specular flow" is essential for various computational vision tasks such as specular shape inference [1, 9] or 3D posing [11]. This inference task, however, is much more challenging than the estimation of traditional optical flow, since the flow's magnitude grows very large (and even unbounded) in certain regions, even if the physical motion in the scene is very small. Moreover, it incorporates several unique structural singularities at the projection of parabolic lines [2, 22].

To the best of our knowledge, the recent work by Adato *et al*. [2] is the first and only attempt to handle specular flows more robustly (under the Cartesian representation, of course). In that work it was argued that all optical flow algorithms fail to estimate this class of flows because the common regularization that all algorithms share is improper in the presence of parabolic singularities and large magnitude. Indeed, in the presence of such singularities, virtually all optical flow algorithms exhibit severe errors in the flow magnitude and significant distortions in the inferred orientation structure. Adato *et al*. [2] suggested to fix these problems with a new regularization term for the parabolic regions, assuming one can locate those areas beforehand (or estimate their location during the optimization process if enough frames are provided). However, as their results show, this solution is still short of estimating large magnitude flows and its results are still insufficient for subsequent visual tasks that require accurate specular flows.

The added complexity in specular flows makes this class of flows a natural candidate for testing the possible advantages of our new polar framework. Furthermore, a database of specular scenes sequences with their ground truth flows is publically available [2]. Hence, we repeated the type of test described for DB-A but with the specular flow data from DB-B. We do note that this type of flows requires additional care in the evaluation of performance. As discussed in Adato *et al*. [2], since specular flows can grow unbounded, most error metrics might report errors of infinite size, which make little sense by way of comparison or performance evaluation. instead, we follow their mapping of the magnitude error to a finite interval using the Geman-McClure function [6].

Using the very same functional (Eq. 9) and minimization process as used for traditional flows in DB-A, the results of our polar-based estimation on specular flows are depicted in Fig. 5 and Fig. 6 second row, and immediately show better performance. The suggested algorithm estimates the flow significantly more accurately than the state-of-the-art optical flow algorithm [20], especially in the magnitude components. Compare, for example, the ground truth in Fig. 4 to Fig. 5.

However, it is also evident that using the basic polar-based optimization (Fig. 5 and Fig. 6) does not provide accurate results, in particular because in specular flows the gradient of the magnitude component is *n*ot close to zero, in contrast to the requirement of the model (see Eq. 2 and Eq. 9). At this point comes the second advantage of our framework – the flexibility to apply more appropriate regularizations according to the physical properties of the flows of interest in the magnitude and orientation domains. Here, for example, we can smooth the flow's orientation using a first derivatives ∇s and ∇c (to obtain deep smoothing) while regularizing its magnitude using the second derivative Δm (which would permit more variations in this component). In other words, for specular flows we propose to minimize

$$\begin{aligned} E(c,s,m) &= \int_\Omega \lambda(s^2+c^2-1)^2 \\ &+\psi(I(x+mc,y+ms,t+1)-I(x,y,t)) \\ &+\alpha_\theta\psi(|\nabla c|^2,|\nabla s|^2)+\alpha_m\psi(|\Delta m|^2)\,d\mathbf{x}\,, \end{aligned} \tag{12}$$

which requires modifying a single line of code from the

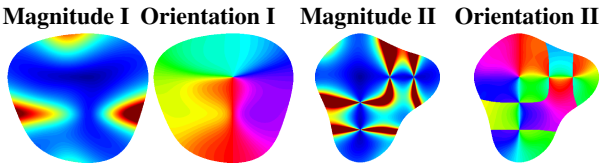


Figure 4. Two ground truth flows from [2]. Note the regions with very large magnitude (**Magnitude I**). Observe the regions with parabolic singularities, i.e., the fixed 180° orientation jump and unbounded magnitude (**Magnitude II and Orientation II**).

implementation of the previous functional.

The results of this minimization process are shown in the third rows of Figs. 5 and Fig. 6 and illustrate how the estimation of both the magnitude and orientation improved significantly. Importantly, we emphasize that there are no obvious equivalent regularizers in the Cartesian representation. Clearly, it seems meaningless to apply different regularizations to the different Cartesian components, and applying it to both coordinate functions has a different meaning altogether. Indeed, to achieve the same effect one must resort to more complex Cartesian expressions, which likely to complicate the minimization process as well.

6.3. Performance on fluid flows

When one examines optical flows due to complex fluid motion, image sequences along with their corresponding ground-truth motion are hard to obtain. Thus, to systematically study the utility of a polar representation in this domain, we synthetically rendered images, as done in the FLUid project [15]. However, while doing so we preferred motions with physical interpretation and therefore we used motion that represents air currents over North America (Fig. 1b).

Given the properties of the flow in question (i.e., relatively large and smooth variations in orientation with comparatively smaller variations in magnitude), for this task we optimized

$$E(c,s,m) = \int_\Omega \lambda(s^2+c^2-1)^2 \\ +\psi(I(x+mc,y+ms,t+1)-I(x,y,t)) \\ +\alpha_\theta\psi(|\Delta c|^2,|\Delta s|^2)+\alpha_m\psi(|\nabla m|^2)\,d\mathbf{x}\,. \quad (13)$$

Once again, we compared the results of the suggested algorithm to the one proposed by Sun *et al*. [20] and show the results in Fig. 7.

As the results show, the polar representation offers significantly better performance. Importantly, this was achieved *without* considering any of the physics behind fluid dynamics (e.g., Navier-Stokes equations) or the effect of fluid diffusion that frequently violates the brightness constancy assumption. Still, merely changing the representation and acting directly on the orientation and magnitude components improves performance qualitatively while using intuitive regularization terms with standard techniques from the traditional optical flow literature.

7. Conclusion

In this work we study optical flow estimation using polar representation of motion. We show that this representation reveals statistical structure that otherwise is concealed in the Cartesian representation. Furthermore, the flow's components seems to be more statistically independent when polar

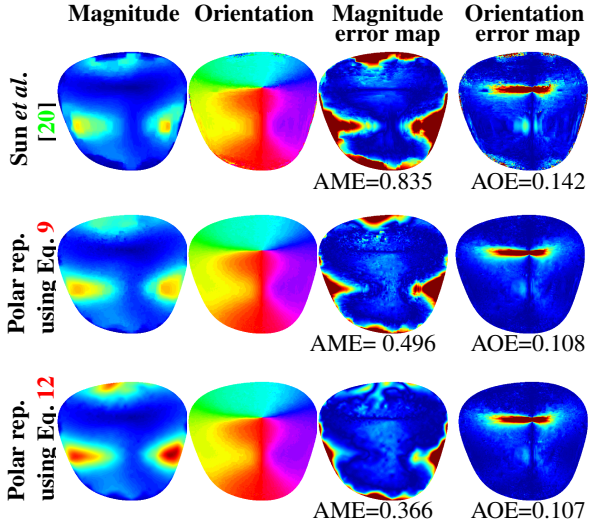


Figure 5. The performance of optical flow algorithms based on Cartesian and polar representations on specular flow sequences with no parabolic singularities. The polar representation based algorithm shows 20 percent less error in orientation and 65 percent error in magnitude than the state-of-the-art algorithms.

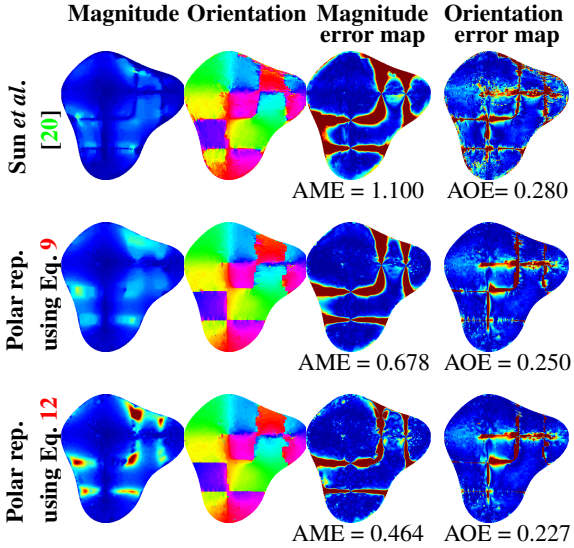


Figure 6. The performance of the algorithms in a scene that contains parabolic singularities. Despite of the difficult structure of the flow in question, the polar representation based algorithm shows significant less error than the state-of-the-art algorithms.

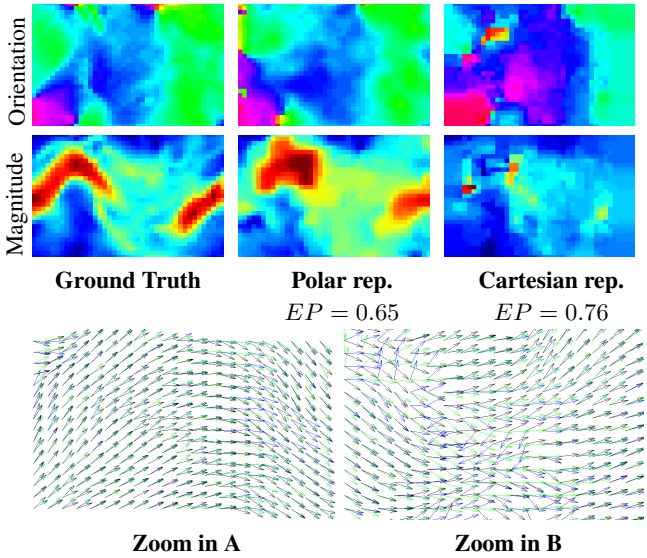


Figure 7. Qualitative performance comparison of the polar and Cartesian representation on image sequences of complex fluid motions. **Top.** The fluid motion ground truth and the two estimated flows coded by their orientation and magnitude. **Bottom.** Zoom in on two regions of interest. In this vector field representation black arrows depict the ground truth, green vectors show the result of our suggested algorithm, and blue show the result of Sun *et al*. [20]. Though neither algorithm exhibit perfect results, note how the suggested polar-based algorithm captures the variations in the flow much more accurately. Please zoom in using the electronic version.

representation is employed. We then demonstrate the flexibility that one can gain when designing a motion estimation algorithm using polar representation. Lastly, we show how performance compares well with state-of-the-art optical flow algorithm on traditional optical flows, and exceeds it on more complex flows such as specular flows of those induced by fluid motion. While much work remains to fully understand this new framework (e.g., how one handles certain filtering or interpolation operations in the cyclic orientation space), we believe that this research direction holds much promise for optical flow estimation in general.

Acknowledgments

This project is funded by the Israel Science Foundation under grant No. 1245/08 and the US National Science Foundation under grant IIS-0712956. O.B.S and Y.A. also thank the generous support of the Frankel Fund and the Paul Ivanier Robotics Center at Ben-Gurion University.

References

[1] Y. Adato, Y. Vasilyev, O. Ben-Shahar, and T. Zickler. Toward a theory of shape from specular flow. In *Proceedings of the IEEE International Conference on Computer Vision*, 2007. 1147, 1150

[2] Y. Adato, T. Zickler, and O. Ben-Shahar. Toward robust estimation of specular flow. In *British Machine Vision Conference*, 2010. 1145, 1146, 1147, 1148, 1149, 1150, 1151

[3] S. Baker, D. Scharstein, J. Lewis, S. Roth, M. Black, and R. Szeliski. A database and evaluation methodology for optical flow. In *Proceedings of the IEEE International Conference on Computer Vision*, 2007. 1145, 1146, 1147, 1149

[4] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994. 1147, 1150

[5] M. Black and P. Anandan. Robust dynamic motion estimation over time. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 296–302, 1991. 1146

[6] M. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. In *Computer Vision and Image Understanding*, 1999. 1150

[7] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *Proceedings of the European Conference on Computer Vision*, 2004. 1146, 1148

[8] A. Bruhn, J. Weickert, and C. Schnorr. Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231, 2005. 1146

[9] G. Canas, Y. Vasilyev, Y. Adato, T. Zickler, S. Gortler, and O. Ben-Shahar. A linear formulation of shape from specular flow. In *Proceedings of the IEEE International Conference on Computer Vision*, 2009. 1150

[10] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *preprint*. http://gpu4vision.icg.tugraz.at. 1149, 1150

[11] J. Y. Chang, R. Raskar, and A. Agrawal. 3d pose estimation and segmentation using specular cues. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2009. 1150

[12] T. Corpetti, E. Mémin, and P. Pérez. Estimating fluid optical flow. In *In Proc. International Conf. on Pattern Recognition*, 2000. 1145, 1146

[13] A. Cuzol and É. Mémin. Vortex and source particles for fluid motion estimation. In *In Proc. 5th Int. Conf. on Scale-Space and PDE methods in Computer Vision (Scale-Space05*, pages 254–266. Springer-Verlag, 2005. 1146

[14] A. Doshi and A. Bors. Robust processing of optical flow of fluids. *IEEE Trans. Image Processing*, 19(9):2332–2344, 2010. 1145, 1146

[15] Fluid image analysis and description. http://fluid.irisa.fr/index.html. 1146, 1147, 1149, 1151

[16] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981. 1145, 1146

[17] D. Marr. *Vision*. W.H.Freeman and Company, 1982. 1145

[18] H. Nagel and W. Enkelmann. An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:565–593, 1986. 1146

[19] S. Roth and M. J. Black. On the spatial statistics of optical flow. In *Proceedings of the IEEE International Conference on Computer Vision*, 2005. 1145, 1146, 1147

[20] D. Sun, S. Roth, and M. Black. Secrets of optical flow estimation and their principles. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2010. 1146, 1149, 1150, 1151, 1152

[21] D. Suter. Motion estimation and vector splines. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1994. 1146

[22] S. Waldon and C. Dyer. Dynamic shading, motion parallax and qualitative shape. In *Proc. IEEE Workshop on Qualitative Vision*, pages 61–70, New York City, NY, USA, 1993. 1150

[23] A. Wedel, T. Pock, J. Braun, U. Franke, and D. Cremers. Duality tv-l1 flow with fundamental matrix prior. In *IVCNZ*, 2008. 1146

[24] A. Wedel, T. Pock, C. Zach, D. Cremers, and H. Bischof. An improved algorithm for tv-l1 optical flow. In *Dagstuhl Motion Workshop*, 2008. 1146

[25] J. Yuan, C. Schnorr, and É. Mémin. Discrete orthogonal decomposition and variational fluid flow estimation. *Journ. of Mathematical Imaging and Vision*, pages 267–278, 2006. 1146

[26] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime tv-l1 optical flow. In *DAGM*, 2007. 1146