Image Processing in Catadioptric Planes: Spatiotemporal Derivatives and Optical Flow Computation

Kostas Daniilidis and Ameesh Makadia
GRASP Laboratory
and Department of Computer and Information Science
University of Pennsylvania

Thomas Bülow
Department of Electrical Engineering
and Computer Sciences
University of California, Berkeley

Abstract

Images produced by catadioptric sensors contain a significant amount of radial distortion and variation in inherent scale. Blind application of conventional shift-invariant operators or optical flow estimators yields erroneous results. One could argue that given a calibration of such a sensor we would always be able to remove distortions and apply any operator in a local perspective plane. In addition to the inefficiency of such an approach, interpolation effects during warping have undesired results in filtering. In this paper, we propose to use the sphere as the underlying domain of image processing in central catadioptric systems. This does not mean that we will warp the catadioptric image into a spherical image. Instead, we will formulate all the operations on the sphere but use the samples from the original catadioptric plane. As an example, we study convolution with the Gaussian and its derivatives as well as the computation of optical flow in image sequences acquired with a parabolic catadioptric sensor.

1 Introduction

Catadioptric sensors are systems consisting of a convex mirror and lens. They can be attached to conventional cameras (CCD or CMOS or film) producing usually an image with a hemi-spherical field of view. While catadioptric devices have long been used in telescopes, only recently they gained in popularity together with other omnidirectional vision systems based on fish-eye lenses or clusters of outwards-looking cameras. For an overview of the subject the reader is referred to a recent textbook [2], a review article [14], and the site for omnidirectional vision[1].

Images obtained by omnidirectional sensors suffer under significant deformations. An exception is the sensor constructed by Hicks [8] which however does not capture an entire hemi-sphere. If the equator of a hemi-spherical field of view is captured then necessarily the image has to be deformed. It is worth noticing that deformations appear also in cylindrical panoramas like concentric mosaics [7].

The deformation has a continuous component - that means it would appear also on a continuous imaging surface - and a discrete component which is due to the decimation induced by sampling. It helps to keep these two aspects separate. This makes deformations in omnidirectional sensors different than deformations in space-variant sensors like log-polar retinas where the deformation is based on the non-uniform sampling.

With a few exceptions of appearance-based methods [9], most approaches analyzing omnidirectional imagery apply image operators "blindly", which means they apply the same operators they would apply on non-deformed cartesian images. A simple example is template matching or matched filtering where we usually search for a match of a translated/rotated/scaled or in general an affinely deformed pattern in a conventional image. Obviously, these are not the local deformations that will appear in an omnidirectional image. The same applies for a filter kernel with a particular orientation or scale. Neighborhood operators like the local structure tensor $\Sigma \nabla I \nabla I^T$ [11, 3] where I is the image do not indicate the local orientation distribution anymore. As far as computation of optical flow is concerned, assumptions about constancy in a local support are not valid in the catadioptric plane.

The argument against "blind" application of classical image operators is that the underlying assumptions for the application of these operators were done on the perspective image plane. One might argue, that with proper calibration we can always remove the distortion and work on local virtual projective planes. However, this is a computationally very expensive and inaccurate approach since the virtual prospective planes contain artificially interpolated intensities.

We argue that signal processing has to take place in

[1] `http://www.cis.upenn.edu/~kostas/omni.html`

the originally sampled values. We need to find another non-deformed space where the old assumptions hold, then write the image operations in that domain, and finally with variable substitution perform the operation in the catadioptric plane. We could choose a frontoparallel image plane as the non-deformed plane and map it to the catadioptric plane. Unfortunately, this mapping does not provide any pre-image for the horizon mapped in most catadioptric images. Even if the horizon could be excluded, the region next to it is mapped to such distant positions in the frontoparallel plane that it suffers from extreme perspective projection.

A natural choice of a non-deformed domain for a hemispherical field of view is the hemi-sphere. Such a choice is also mathematically grounded in the case of catadioptric systems with a unique effective viewpoint [12] and is described in section 2. Given a map from catadioptric image coordinates (u, v) to spherical coordinates (θ, ϕ) we can transform the catadioptric image plane (Fig. 1, left) back to the sphere (Fig. 1, right) if we know the principal point of the catadioptric plane and the sphere radius corresponding to the coupled scaling factor of the orthographic lens and the focal length of the mirror.

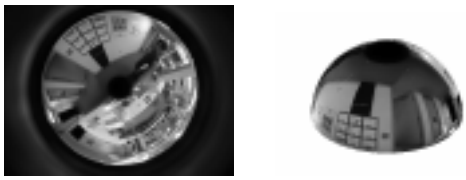

Figure 1. On the left a parabolic catadioptric image; on the right its 1:1 mapping to a hemisphere.

We choose the sphere as the domain where we will be making assumptions about the parameters of our filter and as the domain where operators will be shift-invariant. Convolution with a kernel $G_s(\theta, \phi)$ will be defined on the spherical image $I_S(\theta, \phi)$ giving a response $R_S(\theta, \phi)$ (see Fig. 2). We will introduce the Gaussian function on the sphere and compute the derivatives on the sphere by convolving with its spatiotemporal derivatives. We will keep the derivatives on the sphere for the computation of the optical flow, however, sampled on the catadioptric plane.

We will proceed with the computation of optical flow on the sphere. The computation of optical flow presumes a preservation principle, namely that either the gray-value or the response to a gray-value function is preserved. If we assume gray-value preservation we can easily derive the well-known Brightness Change Constraint Equation (BCCE)

$$I_u \dot{u} + I_v \dot{v} + I_t = 0$$

where (I_u, I_v) and I_t are respectively the gradient and the

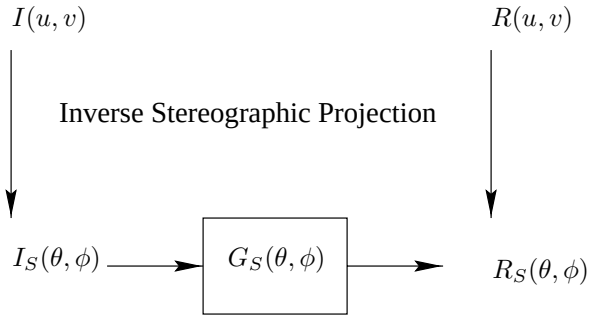

Figure 2. The original image I on the catadioptric plane is mapped through inverse stereographic projection to I_S on the sphere and subsequently convolved on the sphere with a kernel G_S yielding a response R_S which is stereographically projected to R on the plane. The stereographic mappings do never take place as warpings, instead the convolution is performed on the sphere sampling mapped from a uniform sampling on the catadioptric plane. In optical flow computation, we actually use R_S instead of R.

temporal derivative of the intensity function $I(u, v)$. The BCCE equation reveals one main indeterminacy in computing optical flow: it is impossible to estimate the flow component perpendicular to the gradient of the image. To overcome this aperture problem we have to make an assumption, for example, that locally the flow variation is negligible. The aperture problem is overcome as soon as more than one orientations are included in the support window - observe for example the circular aperture on the left of Fig. 3.

Figure 3. The images of a checker board part on a conventional image plane (left) and as obtained by a catadioptric system (right) . Although it is not obvious the circular aperture of the right image contains a curvilinear edge.

Let us turn our attention to the image obtained by a parabolic catadioptric system - see an excerpt in Fig. 3 (right). Although it is only slightly observable in this Figure we know from the geometric properties that the line is now a circular arc. A solution based on sampling the BCCE in a neighborhood [10] will lead to a system of full rank. The

polar deformation from the mirror appears to be eliminating the aperture problem. Of course this is not true. What really happens in a catadioptric image sequence is that an assumption about the local constancy of the flow in the new coordinate system for the same 3D-motion and geometry does not hold.

To proceed with this central idea of analyzing a signal mapped to another domain, we have been inspired by the exponential chirp transform introduced in [4]. However, in that paper the mapping is from cartesian to the logarithmic polar plane. Here, the mapping is from the sphere to the catadioptric plane. While signal processing in the cartesian plane is textbook material, signal processing on the sphere is an ongoing research problem which we study in section 3 after justifying the introduction of the sphere in section 2. We define the Gaussian on the sphere and describe how we perform convolution on the sphere. In section 4 we introduce a new technique for the computation of the optical flow in catadioptric planes. We conclude with experimental verification in real omnidirectional video in section 5.

2 The projection model

Catadioptric systems with a unique effective viewpoint have been proven to be convex reflective surfaces of revolution with a parabolic or hyperbolic profile. Geyer and Daniilidis [6] showed that such projections are equivalent with a projection on the sphere followed by a projection from a point on the sphere axis to the plane. Here, we will elaborate on the parabolic case, in which we will also present our new signal processing approach. In Fig. 4, left, we show a parabolic mirror and an image plane through the mirror focus. The effective viewpoint is the mirror focus and the second center of projection is at infinity yielding a parallel projection onto the image plane. The position of the image plane is thus irrelevant, it only has to be perpendicular to the mirror axis. In [6] it is proven that exactly the same image can be obtained by first projecting on a sphere and then stereographically projecting from the north-pole to the equator plane - Fig. 4, right.

If the equation of the paraboloidal mirror is

$$z = -f + \frac{x^2 + y^2}{4f}$$

then the projection of a point (x, y, z) in the scene to a point (u, v) in the catadioptric plane reads

$$\begin{aligned} u &= \frac{2fx}{-z + \sqrt{x^2 + y^2 + z^2}} \\ v &= \frac{2fy}{-z + \sqrt{x^2 + y^2 + z^2}}. \end{aligned}$$

The radius of the sphere in Fig. 4 is equal to the latus rectum of the parabola and thus equal to $2f$. The stereographic projection from the sphere to the catadioptric plane reads

$$\begin{aligned} u &= \cot\frac{\theta}{2}\cos\phi \\ v &= \cot\frac{\theta}{2}\sin\phi. \end{aligned}$$

By assuming the catadioptric plane to be the plane containing the equator in Fig. 4, the catadioptric image is mapped onto the lower hemisphere in the same figure.

3 Filtering in catadioptric planes

Given a catadioptric image $I(u, v)$ we define its inverse stereographic mapping onto the sphere as

$$I_S(\theta, \phi) \stackrel{def}{=} I(\cot\frac{\theta}{2}\cos\phi, \cot\frac{\theta}{2}\sin\phi).$$

Given a kernel (impulse response of a filter) $G_S(\theta, \phi)$ defined on the sphere, we call $R_S(\theta, \phi)$ the result of the continuous convolution

$$R_S(\theta, \phi) = (I_S * G_S)(\theta, \phi)$$

on the sphere. We call $R(u, v)$ the response on the catadioptric plane which is the stereographic mapping of the response R_S on the sphere.

To realize the convolution on the sphere we have to propose a sampling scheme for the sphere. Since one of the main premises in this paper is to avoid any resampling operation, we will keep the sampling of the catadioptric plane. We assume that a discrete image on the catadioptric plane has been obtained by uniform sampling (at least in the cases using a conventional CCD camera):

$$I[m, n] = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} I(u, v)\delta(u - m\delta u, v - n\delta v).$$

The discrete response reads then

$$\begin{aligned} R[m, n] &= \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} \delta(u - m\delta u, v - n\delta v) \\ &\quad \iint (u', v')K(u, v, u', v')du'\, dv'. \end{aligned}$$

where $K(u, v, u', v')$ is the kernel on the sphere expressed in coordinates of the catadioptric plane. We will see that it can not be written as a shift and that it will have a varying form dependent on the corresponding position on the sphere. Hence, filtering on the catadioptric plane is not a shift-invariant operation and thus a computationally expensive operation. It is also worth to note that the inverse stereographic projection of the uniform samples in the plane yields a non-uniform distribution on the sphere becoming denser towards the equator.

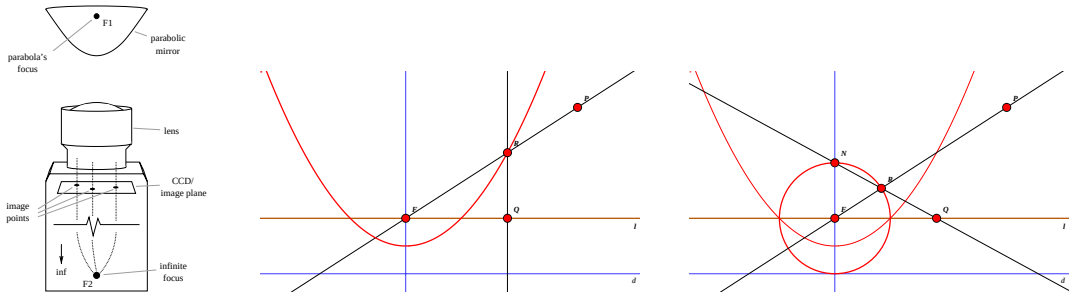

Figure 4. On the left, a sketch of a parabolic catadioptric camera with the two centers of projection F_1 and F_2 [6]. In the middle, the projection of a point P on the parabolic profile with subsequent parallel projection to Q in the catadioptric plane. On the right, the projection of a point P on the sphere with subsequent stereographic projection to Q' on the catadioptric plane, for a proof of $Q = Q'$ see [6]

3.1 Definition of a Gaussian on the sphere

Definition of filter kernels on the sphere is not trivial. Imagine defining a Gaussian function with (θ, ϕ) as an argument. In addition to the singularity at the north-pole, the function would not be periodic, which is a natural property of functions defined on the sphere. In this paper, we use as a filtering example the computation of spatial derivatives on the sphere. We will not delve into the discussion of the derivative definition on the sphere, which involves the spherical harmonics of the sphere, at this point. Let us assume that a "good" derivative estimator on the sphere is the response to the first Gaussian derivative on the sphere.

We introduce the Gaussian function on the sphere as follows [1]: We start with the definition of a Gaussian on a plane tangent at the south-pole of the sphere. We map it by inverse stereographic projection from the north-pole to the south hemisphere. The line at infinity is thus mapped to the north-pole at which we define the value of the inverse mapping to be zero. The Gaussian with offset at the south-pole reads then

$$G_s(\theta, \phi) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\cot^2\frac{\theta}{2}}. \qquad (1)$$

It is just a coincidence that we use stereographic projection in our catadioptric model as well as in the definition of the Gaussian. The reader should be aware that to perform the convolution we rotate the Gaussian with offset at the south-pole to the specified (θ, ϕ). Because it is a rotation, the operation is shift-invariant and thus a convolution on the sphere. This is different than first shifting on the cartesian plane and then projecting on the sphere which would yield a shift-variant operation on the sphere. In [5], it is shown how to define the Gauss function as the solution of the diffusion PDE on the sphere.

The computation of the spatial derivatives in θ and ϕ is also obtained from the inverse stereographic projection and rotation of a Gaussian first derivative with respect to u. Recall that the all derivatives are smoothed in the orthogonal directions including the temporal derivatives which are 0th order Gaussians on the sphere.

3.2 Convolution on the sphere

To perform convolution on the sphere, we define a function at either the north-pole or the south-pole and we rotate the pole to the new position. The reader might have already realized (1) that this cannot be written with a shift in the θ and ϕ angles.

We embed the sphere in R^3 and write an element $\eta \in S^2$ as η:

$$\eta = (\cos(\phi)\sin(\theta),\ \sin(\phi)\sin(\theta),\ \cos(\theta)).$$

We will write a function $I_S(\theta, \phi)$ as $i_S(\eta)$.

Rotations in R^3 will be parameterized by Euler angles such that any $g \in SO(3)$ will be written as

$$g(\gamma, \beta, \alpha) = R_z(\gamma)R_y(\beta)R_z(\alpha), \qquad (2)$$

where R_y and R_z denote rotation about the y-, and z-axis, respectively. A rotation of a function to a new position can be formulated as the action of the operator $\Lambda(g)$ which is defined by

$$\Lambda(g)f(\eta) := f(g^{-1}\eta). \qquad (3)$$

In our case, we constrain the rotations to the subgroup $g(\gamma, \beta, \alpha) = R_z(\gamma)R_y(\beta)$. The convolution will be defined as

$$(I_S * R_S)(\beta, \gamma) = \int_{\eta \in S^2} i_S(\eta) r_S(g^{-1}\eta) d\eta \qquad (4)$$

where $g = R_z(\gamma)R_y(\beta)$ and $d\eta = \sin\theta d\theta d\phi$.

4 Optical Flow Computation

The central assumption in optical flow computation is that the projection of a scene point will generate the same grey-value in each image of the motion sequence. If we assume this grey-value preservation, we can easily derive the well known Brightness Change Constraint Equation (BCCE)

$$I_x\dot{u} + I_v\dot{v} + I_t = 0$$

where (I_u, I_v) and I_t are respectively the gradient and the temporal derivative of the intensity function $I(u, v)$. The optical flow vector is denoted by $(\dot{u}, \dot{v})$. To overcome this aperture problem we have to make an assumption, for example, that locally the flow variation is negligible. This allows us to use the flow constraints from a neighborhood of pixels to estimate the flow at a single location. We can then estimate the flow using a weighted least squares estimation technique. To find the flow vector at any point $p_0 = (u_0, v_0)$, we minimize the function

$$\sum_{(p)\in N(p_0)} W(p)(I_u(p)\dot{u_0} + I_v(p)\dot{v_0} + I_t(p))^2$$

where $W(p)$ is a normalized neighborhood weighting function and $N(p_0)$ is a neighborhood region surrounding p_0. This minimization problem can be solved using singular value decomposition. This method for estimating the optical flow was first described by Lucas and Kanade [10].

Theoretically, the orientation variation can be studied with a differentiation of the BCCE above [11, 13] yielding, including the BCCE itself, three equations

$$\begin{aligned} I_u\dot{u} + I_v\dot{v} + I_t &= 0 \\ I_{uu}\dot{u} + I_{uv}\dot{v} + I_{ut} &= 0 \\ I_{uv}\dot{u} + I_{vv}\dot{v} + I_{vt} &= 0 \end{aligned}$$

Assuming that the gradient does not vanish, the indeterminacy of this system clearly depends on the rank of the Hessian of the intensity function $I(u, v)$ which depends on whether there is more than one local orientation in the intensity function.

If we examine the Hessian of the intensity function on the catadioptric plane, we realize that it might be full rank even if the corresponding Hessian on the sphere or a Hessian in a perspective image are rank deficient.

Let us now examine the optical flow equation on the sphere. If we define the signal on the sphere as $I_S(\theta, \phi)$, then the optical flow equation on the sphere is assumed to be:

$$\frac{1}{\sin\theta}\frac{\partial I_S}{\partial \phi}\dot{p}_\phi + \frac{\partial I_S}{\partial \theta}\dot{p}_\theta + I_t = 0 \qquad (5)$$

where $p = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is a sphere point and $(\dot{p}_\theta, \dot{p}_\phi, 0)$ is the flow vector in the tangential coordinate system. Given the projection model from before, we can define:

$$\begin{aligned} I_S(\theta, \phi) &= I(u(\theta, \phi), v(\theta, \phi)) \\ \frac{\partial I_S}{\partial \theta} &= \frac{\partial I}{\partial u}\frac{\partial u}{\partial \theta} + \frac{\partial I}{\partial v}\frac{\partial v}{\partial \theta} \\ \frac{\partial I_S}{\partial \phi} &= \frac{\partial I}{\partial u}\frac{\partial u}{\partial \phi} + \frac{\partial I}{\partial v}\frac{\partial v}{\partial \phi} \end{aligned}$$

We add the extra assumption that the optical flow field be locally constant. To find the flow vector at any point p_0 on the sphere, we minimize the function

$$\sum_{p\in N(p_0)} W(p)(\frac{1}{\sin\theta}\frac{\partial I_S}{\partial \phi}(p)\dot{p_0}_\phi + \frac{\partial I_S}{\partial \theta}(p)\dot{p_0}_\theta + \frac{\partial I_S}{\partial t}(p))^2 \qquad (6)$$

where $W(p)$ is a normalized neighborhood weighting function and $N(p_0)$ is a neighborhood region surrounding p_0. Note that during implementation, p is a stereographically projected point from an already discretized catadioptric plane. This minimization problem can again be solved using SVD. Observe that the final result of this minimization is the optical flow on the sphere. Depending on the task at hand, we can use it directly for egomotion computation or stereographically map it on the catadioptric plane.

We will test two methods based on this technique in the following section. One method will calculate image derivatives using a shift-invariant gaussian defined on the sphere (Sec. 3) and the other will use the same gaussian kernel scaled with a shift-variant σ.

5 Experiments

In this section, we compare the results of our new approach described in last section with a conventional approach [10] based on the same principle but applied on the catadioptric image as if it were a conventional cartesian image. We denote our approach by SI (for shift-invariant) and the conventional by LK (Lucas and Kanade [10]). Recall that the new approach assumes that the computation of the spatiotemporal derivatives on the sphere is shift-invariant, which means the filter support on the sphere is independent of the sphere position. However, this results to an effective support on the catadioptric plane increasing towards the periphery. As we will observe in the real sequences, in many occasions the periphery contains high-frequency image variation because of the projection of distant objects. To aleviate blurring this region, we introduced a modification of the new approach, denoted by SI_σ, where we assume that the support is constant on the fronto-parallel plane instead of the sphere. Let us define σ_x as the shift-invariant scale

factor on the fronto-parallel plane. We look for a σ on the sphere, denoted by σ_s, corresponding to the shift-invariant σ_x. If x is a point on the fronto-parallel line, and θ is the corresponding point on the unit circle, then we can define the relationship (in one dimension) to be:

$$tan(\theta + \sigma_s) = x + \sigma_x.$$

Solving for σ_s:

$$tan(\sigma_s) = \frac{\sigma_x}{(1 + x^2 + x\sigma_x)}.$$

As the corresponding point on the fronto-parallel moves further away from the origin, σ_s decreases. To successfully implement this technique, you must put a limit on the minimum value allowable for σ_s (here started equal to 3) determined by the minimum region of neighborhood support you would like. The temporal support of the derivatives is 7 frames ($\sigma = 1$). The spatial support for the derivatives in LK is $\sigma = 2$ (17x17). The same was the support for the Gaussian used in SI before its inverse stereographic projection to the sphere.

Each method was tested on two greyscale image sequences. All images were taken using a digital camera (Nikon 995), with a parabolic mirror attachment (Remote Reality OneShot360). The motion between captures was generated by moving the camera one-half inch along a slider apparatus, ensuring purely translational motion. Each sequence consists of eight images having a resolution of 700x700 pixels. In the first image sequence, most of the texture can be found in regions of large scene depth. The second sequence was taken from inside a small room, resulting in images containing only small scene depth (see Figure 5).

To evaluate the three approaches, we use two measures: The first is the location of the focus of expansion t (or translation direction) on the sphere obtained from the epipolar constraint $t^T(p \times \dot{p}) = 0$. The second measure is the distribution of the flow for a given flow density, which means using thresholds that yield the same percentage in all three approaches.

Examining Figure 6 we can see that all three methods pick up stronger flow at the periphery because that is where large scene depth, high-frequency content is located. In particular, SI_σ detects far more flow in the outer region than the middle region relative to the other methods. This is expected because σ is being scaled and so the derivative filter becomes more sensitive as θ increases. Also note that at the density of 75%, SI_σ picks up much noise, especially in the untextured, small scene depth region.

As we can see from Table 1, SI and SI_σ both outperform LK up to the 50% threshold. However, at the 75% threshold, SI_σ is the least accurate. This is due to the sensitivity of the method which is more easily affected by noise

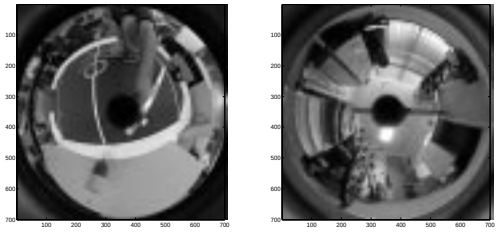

Figure 5. Images taken using a parabolic mirror attachment. On the left is an image from the first sequence (significant depth variation). On the right is am image from the second sequence (small room, less depth variation).

τ	LK	SI	SI_σ
10%	1°	1°	1°
50%	23°	19°	17°
75%	28°	24°	39°

Table 1. Azimuth errors for the estimated Focus of Expansion for the first image set. τ is the percentage of flow density preserved before errors were calculated.

at lower thresholds. The reason SI_σ outperforms the other two methods at higher thresholds can be explained by noting that the images contain high-frequency information at the outer region.

Examining Figure 7 we notice that the flow is detected more evenly throughout the image for all methods. This occurs because the second image sequence contains only small scene-depth information. There does not exist the high-frequency content at the periphery as there was in the first image set. Since SI_σ is sigma-scaled, and since the scene has little high frequency content, it picks up much more noise throughout the image than the other two methods.

τ	LK	SI	SI_σ
10%	1°	1°	2°
50%	10°	8°	15°
75%	15°	14°	22°

Table 2. Azimuth errors for the estimated Focus of Expansion for the second image set. τ is the percentage of flow density preserved before errors were calculated.

As we can see from Table 2, SI and SI_σ have relatively

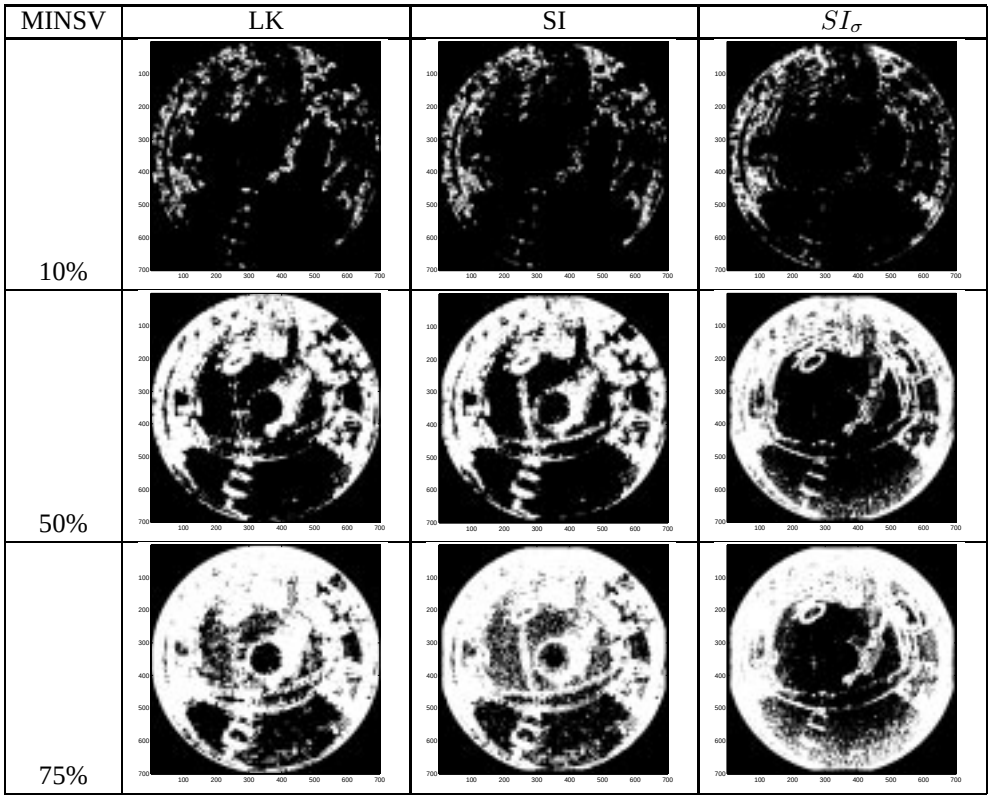

Figure 6. These are the results of each method applied to first image set. The first row displays the lowest singular values of the linear systems thresholded at the strongest 10%. The densities are 50% and 75% for the second and third row, respectively. The standard deviation used for LK and SI are the same. The starting standard deviation used in SI_σ is $\sigma = 3$;

similar accuracies at the three different thresholds, while SI_σ is less effective at all three thresholds. This is assumed to be the cause of the nature of the scene captured in the second image set. Comparing results generated for the two image sets we see that each method performs better when the scene has smaller depth variation.

6 Conclusion

We have introduced a framework for filtering and flow computation in catadioptric images. Convolutions take place on the sphere but their realization is a space-variant operation on the discrete samples of the original catadioptric image. We have described how a Gaussian function is defined on the sphere and how the image derivatives necessary for flow computation are estimated. We have reformulated the BCCE on the sphere and solved for the optical flow on the sphere by assuming that it is locally constant. Results on real sequences showed that the new approach improves on the computation of heading direction as well as in the quality of the flow field for the same field density. In our future work, we study filter design and optical flow through the frequency domain of the sphere consisting of the spherical harmonic coefficients. We also search for a fast computation of space-variant filtering in the catadioptric plane. The proposed methods can be generalized to mutiple other image modalities whose deformations are well understood.

Acknowledegments

The authors are grateful for support through the following grants: NSF-IIS-0083209, NSF-EIA-0218691, NSF-IIS-0121293, NSF-EIA-9703220, a DARPA/ITO/NGI subcontract to UNC, and a Penn Research Foundation grant.

References

[1] J.-P. Antoine and P. Vandergheynst. Wavelets on the 2-sphere: A group-theoretical approach. *Applied and Computational Harmonic Analysis*, 7:262–291, 1999.

[2] R. Benosman and S. Kang. *Panoramic Vision*. Springer-Verlag, 2000.

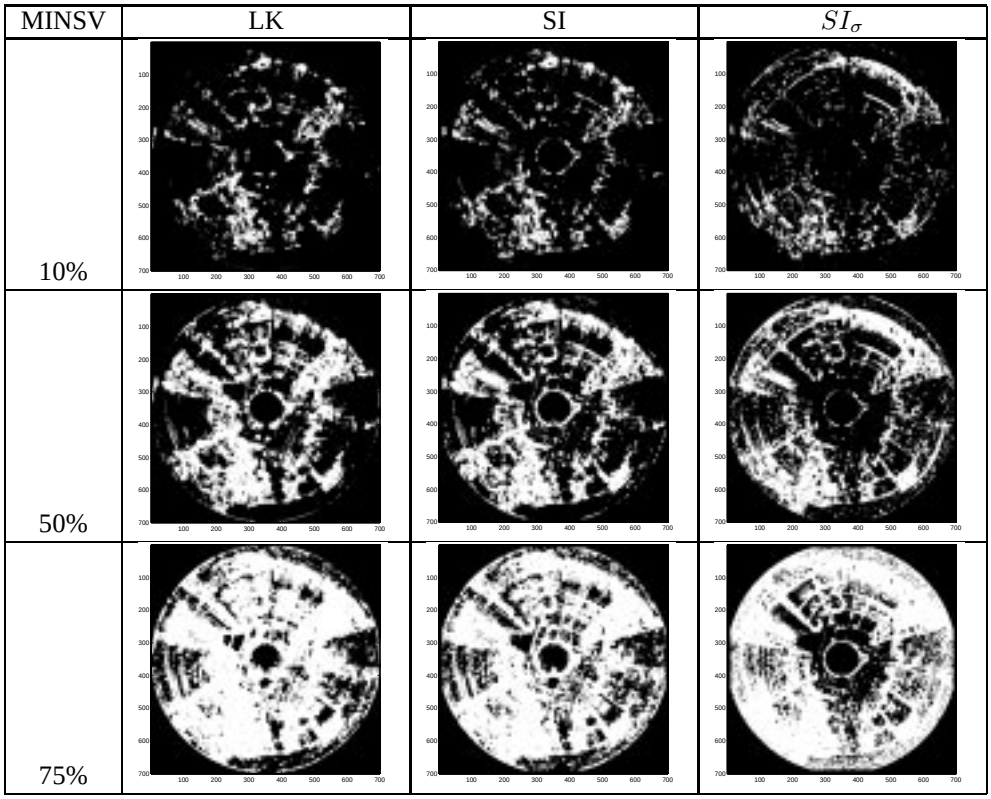

Figure 7. These are the results of each method applied to the second image set. The first row displays the lowest singular values of the linear systems thresholded at the strongest 10%. The densities are 50% and 75% for the second and third row, respectively. The standard deviation used for LK and SI are the same. The starting standard deviation used in SI_σ is $\sigma = 3$.

[3] J. Bigün, G. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13:775–790, 1991.

[4] G. Bonmassar and E. Schwartz. Space-variant fourier analysis: The exponential chirp transform. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19:1080–1089, 1997.

[5] T. Bülow. Spherical diffusion. Technical Report MS-CIS-01-38, Computer and Information Science Department, University of Pennsylvania, 2001.

[6] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *International Journal of Computer Vision*, 43:223–243, 2001.

[7] L.-W. H. H.-Y. Shum. Rendering with concentric mosaics. In *Proceedings of SIGGRAPH, in Computer Graphics Proceedings, Annual Conference Series*, pages 299–306, 1999.

[8] R. Hicks and R. Bajcsy. Catadioptric sensors that approximate wide-angle perspective projections. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 545–551, Hilton Head Island, SC, June 13-15, 2000.

[9] A. Leonardis and M. Jogan. Robust localization using eigenspace of spinning-images. In *IEEE Workshop on Omnidirectional Vision, Hilton Head, SC, June 12*, pages 37–46, 2000.

[10] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *DARPA Image Understandig Workshop*, pages 121–130, 1981.

[11] H.-H. Nagel. Displacement vectors derived from second-order intensity variations. *Computer Vision, Graphics, and Image Processing*, 21:85–117, 1983.

[12] S. Nayar. Catadioptric omnidirectional camera. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 482–488, Puerto Rico, June 17-19, 1997.

[13] O. Tretiak and L. Pastor. Velocity estimation from image sequences with second order differential operators. In *Proc. Int. Conf. on Pattern Recognition*, pages 16–19, 1984.

[14] Y. Yagi. Omnidirectional sensing and its application. *IEICE Trans. Inform. & Systems*, E82-D:568–579, 1999.