IMAGE RECONSTRUCTION FROM ZERO-CROSSINGS

Su-shing Chen

Department of Computer Science, University of North Carolina
Charlotte, NC 28223, U.S.A.

ABSTRACT
In the computational theory of human vision, developed by D. Marr and his school, zero-crossings detection is proposed as one of the first stages of human visual information processing. The justification of this theory reguires that the zero crossings are rich in information about the original image. In a paper, T. Poggio, H.K. Nishihara and K.R.K. Nielsen stated that a successful extension of B. Logan's analysis to two-dimensional images represents one of the critical unsolved steps of this computational theory. Moreover, sucn an extension is crucial to the question of spatiotemporal interpolation in vision. In this paper, we shall provide a solution to this problem using a theory of spherical perspective and spherical harmonics.

1. INTRODUCTION

The goal of low level vision is to detect changes in the reflectance of object surfaces and in surface orientation, geometry and depth value. In computer vision or the computational human vision, changes in image intensity of different scales of resolution is the optimal indicator of such changes. It has been proposed that different scales are set by filtering the image with 2-dimensional Gaussian filters of different sizes [2]. Intensity changes of object surfaces are spatially localized at different scales.

In the spherical model, we could extend the Gaussian filter on the plane to a filter (probability distribution) of Gaussian type, such as Fisher and von Mises filters on the sphere. It is quite natural to consider the exponential family of distributions on the sphere, because of its relation to spherical harmonics. We may consider some characterizations of the Gaussian to set up analogous filters in the spherical model. Alternatively, we may use the correspondence between the image plane and the image sphere (see next section) to transform filtered planar images to scaled images on the sphere. That is, we may do the filtering in the plane first and then transform the filtered image to an scaled image G*I on the sphere.

In both cases, zero-crossings are computed by finding the zero set of the spherical Laplace operator A on G*I. Detection of edges in an image is reduced to solving for zero-crossings. A key question is whether detected edges contain full information about the image. If we can reconstruct the image from detected edges, then the question is affirmatively answered. From the view point of visual information processing, we do not need to reconstruct the original image, but only to show that sufficiently enough information is provided.

An extension of B. Logan's theorem to the image sphere will be presented. Thus, the theory of Marr and Hildreth [2] will be supported by this paper. The sphere is parametrized by the elevation and azimuth angles θ and ϕ. Any square integrable function f(6,<|0) may be decomposed into a sum of spherical harmonics $\{Y_{km}(\theta,\phi)\}$, which are products of exponentials and Legendre functions (special functions), functions of two variables θ and ϕ. As we have stated earlier, this framework will provide the appropriate extension of Logan's theorem to two variables. It will be shown that such a function f can be reconstructed from the zero-crossings data, provided that enough of them are available.

This extension of Logan's theorem is the theoretical basis of zero-crossings and spatiotcmporal interpolation. Spatiotemporal interpolation is proposed in [1] and other references quoted there. The positional accuracy in human vision, corresponding to a fraction of the spacing between adjacent photoreceptors in the fovea [1], implies the spatial interpolation. The perception of continuous motion requires spatiotemporal interpolation [1], Our extension of Logan's theorem addresses both the problem of spatial interpolation and an extension to θ,ϕ,t-variables for spatiotemporal interpolation.

2. THE SPHERICAL MODEL

The two imaging models are not very different. The image sphere (or hemisphere) is replacing the image plane tangent to it at the north pole. The perspective projection is extended beyond the image plane to the spherical perspective projection onto the image sphere. In [5], a nonlinear transformation sets up the correspondence between points on the image plane and points on the image sphere. Image intensity is defined by this correspondence, under the Lambertian model. If we move the origin (viewer) of the spherical model to infinity, the limit is the orthographic model. There are other nice features of the spherical model. For instance, the vanishing points are finite points on the image sphere.

Like the Fourier descriptor method of expanding functions on $[0,1]$ in exponentials, we expand an image intensity function $I(\theta,\phi)$ in spherical harmonics $\{Y_{km}(\theta,\phi),\ -k\leq m\leq k,\ k=0,1,2,\ldots\}$ which form a basis of all square integrable functions on the unit sphere. There are other orthonormal expansions of image intensity functions, such as SVD and Karhunen-Loeve expansions; however, all the nice properties of Fourier descriptors are lost. This spherical harmonics expansion carries the group invariance over and forms a natural setting in the group representation theory of modern physics [5].

Each spherical harmonic $Y_{km}(\theta,\phi)$ is the product $\exp(im\phi)P_{km}(\theta),\ -k\leq m\leq k,$ for each k. In fact, there is a vector space spanned by Y_{km}, for each $k=0,1,2,\ldots.$ The Legendre function P_{km} is presented in [5]. For any square integrable $I(\theta,\phi)$, we have

$$I(\theta,\phi) = \Sigma_{k=1}^{\infty}\Sigma_{m=-k}^{k}\ c_{km}\ Y_{km}(\theta,\phi),$$

$$c_{km} = \int_0^{2\pi}\int_0^{\pi} I(\theta,\phi)\ Y_{km}(\theta,\phi)\ \sin\theta d\theta d\phi.$$

These coefficients can be calculated numerically and lookup tables are constructed to speed up numerical work involving with Legendre functions. In [5], a distance $d(I_1,I_2)$ is defined for equivalent shape classes I_1 and I_2. Thus, object recognition is performed in a standard fashion. If I_1 and I_2 have respectively Fourier coefficients $\{c^{(1)}_{km}\}$ and $\{c^{(2)}_{km}\}$, then the minimum of the vector $(c^{(1)}_{km}-RsSc^{(2)}_{km})$, over all rotations R, symmetries s and scalings S, is the distance between I_1 and I_2.

3. ZERO-CROSSINGS

The first order differential operators on the sphere are given by three basis operators

$$D_1=\exp(i\phi)(\partial/\partial\theta+i\cot\theta\partial/\partial\phi),$$
$$D_2= \exp(-i\phi)(-\partial/\partial\theta+i\cot\theta\partial/\partial\phi),$$
and $D_3= -\partial/\partial\phi.$

The effect of these three differential operators on spherical harmonics $\{Y_{km}(\theta,\phi)\}$ is given by

$$D_1Y_{km}=\surd(k-m)(k+m+1)\ \ Y_{k(m+1)},$$
$$D_2Y_{km}= \surd(k+m)(k-m+1)\ \ Y_{k(m-1)},$$
$$D_3Y_{km}= kY_{km}.$$

The spherical Laplace operator Δ is computed to act on Y_{km} by

$$\Delta Y_{km}= k(k+1)Y_{km},$$

Thus, the Laplace operator Δ is a scalar matrix when restricted to the subspace $V_k=\{Y_{km}|\ -k\leq m\leq k\}$ for each $k=0,1,2,\ldots.$This is a very nice feature of the spherical model.

The correspondence between $Q(X,Y)$ on the image plane and $S(\theta,\phi)$ on the image sphere is given by the nonlinear transformation between the image plane at the north pole and the sphere. The image intensity function $I(\theta,\phi)$ is conveniently defined as the image intensity function $I(X,Y)$ under this correspondence. This is a reasonable assumption near the fovea in human vision or near the optical axis in computer vision because the plane is tangent to the sphere. For peripheral views or wide viewing angles, problems arise and we need to rotate our eyeballs or cameras. This amounts to a rotation of the sphere and its tangent plane at the north pole.

Since the Gaussian filter (distribution function) in [2] is to set up different scales of resolution, it is quite reasonable to project the 2-dimensional Gaussian filter on the plane $G(X,Y,\sigma)=(1/\sqrt{(2\pi)}\ \sigma)\ \exp(-(X^2+Y^2)/2\sigma^2)$ to a filter on the sphere

$$G(\theta,\phi,\sigma)=(1/\sqrt{(2\pi)}\ \sigma)\ \exp(-\cot^2\theta/2\sigma^2).$$

Now, we consider $\Delta(G*I) = 0$. The zero set is the zero-crossings of I in the spherical model. The blurred image $G*I$ may be expanded in terms of spherical harmonics:

$$G*I(\theta,\phi)=\Sigma_{k=0}{}^{\infty}\Sigma_{m=-k}{}^{k}c_{km}Y_{km}(\theta,\phi).$$

The Laplace operator Δ acts on $G*I$ by

$$\Delta(G*I)(\theta,\phi)=\Sigma_{k=0}{}^{\infty}\Sigma_{m=-k}{}^{k}c_{km}\Delta Y_{km}(\theta,\phi)$$
$$=\Sigma_{k=0}{}^{\infty}\Sigma_{m=-k}{}^{k}c_{km}k(k+1)Y_{km}(\theta,\phi).$$

Thus, zero-crossings are given by solutions of the following nonlinear equation

$$\Sigma_{k=1}{}^{\infty}\Sigma_{m=-k}{}^{k}c_{km}k(k+1)Y_{km}(\theta,\phi)=0.$$

For each k and m, $Y_{km}(\theta,\phi)= \exp(im\phi)P_{km}(\theta)$, $-k\le m\le k$. The Legendre functions $P_{km}(\theta)$ could be calculated by lookup tables over a sufficiently fine subdivision of $(0,\pi)$. Let $d_{km}(\theta_0)$ be the value of P_{km} at a particular point θ_0. In the cross section at θ_0, we have

$$\Sigma_{k=1}{}^{\infty}\Sigma_{m=-k}{}^{k}c_{km}k(k+1)d_{km}(\theta_0)\exp(im\phi)=0.$$

For bandlimited functions, we have $\Sigma_{k=1}{}^{N}\Sigma_{m=-k}{}^{k}c_{km}k(k+1)d_{km}(\theta_0)\exp(im\phi)=0.$

By interchanging summation signs over k and m, we factor out

$$e_m(\theta_0)=\Sigma_{k=1}{}^{N}c_{km}k(k+1)d_{km}(\theta_0)$$

and the equation of zero-crossings in the cross section at θ_0 becomes

$$\Sigma_{m=-N}{}^{N}e_m(\theta_0)\exp(im\phi)=0.$$

Thus, we have successfully reduced the equation of zero-crossings at each cross section to the equation of B. Logan [1], [3]. Like [1], [3], we may assume that the function is bandpass with one octave bandwidth. The complex coefficients $e_m(\theta_0)$'s may be determined by the 2N solutions of the equation as the solutions of

$$\Sigma_{m=-N}{}^{N}e_m(\theta_0)\exp(im\phi_1)=0$$
$$\Sigma_{m=-N}{}^{N}e_m(\theta_0)\exp(im\phi_2)=0$$
$$\cdots$$
$$\Sigma_{m=-N}{}^{N}e_m(\theta_0)\exp(im\phi_{2N})=0$$

After the set $\{e_m(\theta_0)\}$ is determined, we solve the set of equations

$$\Sigma_{k=1}{}^{N}c_{km}k(k+1)d_{km}(\theta_0)=e_m(\theta_0), -N\le m\le N,$$

for the unknowns $\{c_{km}\}$. For each m, there are N unknowns $\{c_{km}\}$, k=1,2,...,N. We need to set up a system of N equations, for each m, and N cross-sections $\{\theta_1,...,\theta_N\}$:

$$\Sigma_{k=1}{}^{N}c_{km}k(k+1)d_{km}(\theta_1)=e_m(\theta_1)$$
$$\Sigma_{k=1}{}^{N}c_{km}k(k+1)d_{km}(\theta_2)=e_m(\theta_2)$$
$$\cdots$$
$$\Sigma_{k=1}{}^{N}c_{km}k(k+1)d_{km}(\theta_N)=e_m(\theta_N)$$

The samples $\{\theta_1, \theta_2,...,\theta_N\}$ may be taken from a suitable subdivision of $(0,\pi)$ such that the lookup tables of $d_{km}(\theta_i)$, together with $k(k+1)$, form determinants of maximum rank. Consequently, the image intensity function $I(\theta,\phi)$ is recovered from the spherical harmonics expansion whose coefficients $\{c_{km}\}$ are determined by the above $2N+1$ systems of N equations.

THEOREM. Under suitable conditions, image intensity functions in the spherical model may be recovered up to multiplicative constants from their zero-crossings data.

REFERENCES

1. T. Poggio, H.K. Nishihara and K.R.K. Nielsen, Zero-crossings and spatiotemporal interpolation in vision: aliasing and electrical coupling between sensors, MIT AI Lab, AI Memo No. 675, May, 1982.
2. D. Marr and E. Hildreth, Theory of edge detection, Proc. Royal Soc. London, B. 207 (1980), 187-217.
3. B. F. Logan, Information in the zero-crossings of band pass signals, Bell System Tech. Journal, 56 (1977), 510.
4. S. Chen, A new vision system and the Fourier descriptor method by group representation theory, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1985.
5. S. Chen, An intelligent computer vision system, International Journal of Intelligent Systems, 1986.