Optimal Edge Detectors for Ramp Edges

Maria Petrou and Josef Kittler

Abstract—**We argue that the best way to model an edge is by assuming an ideal mathematical function passed through a low pass filter and emersed in noise. Using techniques similar to those developed by Canny and Spacek, we derive optimal filters for ramp edges of various slopes. We also derive the optimal nonrecursive filter for ideal step edges as a limiting case of the filters for ramp edges. Because there are no step edges in images, we show that edge detection is improved when the ramp filter is used instead of the filters developed for step edges. For practical purposes we give in Table IV some convolution masks which can be used directly for edge detection without the need to go into the details of the subject.**

Index Terms— **Canny edge detector, edge detection, optimal filters, ramp edges, step edges.**

I. Introduction

The problem of edge detection is a fundamental one in image processing and has eluded scientists for many years. There is a plethora of papers in the subject on edge detection filters and algorithms which work with various degrees of success for different images. (A good selection of references is given in [7].) This is not a review paper on the subject, so we are not going to attempt to summarize all the work done in this area. Instead, we are concentrating on a particular approach to the problem, namely the approach established by Canny, which has rapidly gained widespread popularity and became a benchmark for evaluating the results of other approaches to the subject.

Canny was the first to set the foundations of the theory of an optimal edge detector: good signal to noise ratio, good locality, and maximum suppression of false responses. He derived quantitative measures for these three qualities and combined the first two to form a performance measure for an edge detector. He then maximized that measure under the additional constraint of reduction in the number of false responses. The equations he arrived at were long and complicated. Eventually, he proposed the derivative of a Gaussian as a good approximation to the best operator derived by the method described above. This operator is simple in form and its performance measure is 80% that of Canny's best operator [2], [3].

Spacek on the other hand formed a performance measure combining all three quantitative measures Canny had derived. In doing so, he simplified the differential equation, the solution of which gives the form of the optimal filter. The final form of the filter equation Spacek derived depended on six parameters, the numerical values of which had to be chosen so that the performance was optimal. In order to simplify the process, Spacek fixed two of those parameters and determined the remaining four from the boundary conditions [9]. Subsequently, he revised his work, but those results are less publicized [10]. Deriche used the same approach as Canny, but he allowed his filters to be of infinite extent. His main contribution lies in the idea that an edge filter of the kind can be implemented recursively and thus most efficiently. The filter he derived is optimal in its functional form, but it depends on two parameters which can be chosen freely by the user, so that the signal to noise ratio of the filter becomes as large as one wants at the expense of good locality and vice versa.

All the above work concerns ideal step edges emersed in white Gaussian noise. In reality, however, image edges are never ideal steps. Even if scene edges had idealized mathematical shapes to begin with, during the process of image capture and digitization the step edges will be converted into ramps. This is a direct consequence of the fact that any practical imaging system will have a finite bandwidth and therefore it will behave approximately as a low pass filter, blurring the edges [7], [8].

The purpose of this work is to extend Spacek's work by applying it to ramp edges. In Section II we model the ramp edges conveniently and derive the appropriate performance measure. In Section III we derive the form of the optimal filter for ramps and in Section IV we choose the values of its parameters for best performance. As an aside, in Section V we derive the optimal filter for ideal step edges. In Section VI we discuss some implementation details and finally, in Section VII we compare our filters with some other popular ones and present our conclusions.

II. Performance Measure for Detectors of Ramp Edges

We assume that the profile of the ramp edge we want to detect can successfully be modeled by the function

$$c(x) = \begin{cases} 1 - e^{-sx}/2 & \text{for } x \geq 0 \\ e^{sx}/2 & \text{for } x \leq 0 \end{cases} \tag{1}$$

where s is some positive constant.

This function has been chosen for convenience reasons (differential equation (9) can easily be solved) but at the same time, an ideal step edge passed through a low pass filter of the form $g(x) = se^{-s|x|}/2$ will take the above shape. This filter is the first approximation to any low pass filter involved in the imaging process. Figs. 1 and 2 show the profiles of various step edges randomly chosen from the images shown in Figs. 5 and 6. All the step edges in an image can be successfully modeled by (1) with the same value of s as can be seen from these figures. Further, we assume that the edges are emersed in white Gaussian noise. This assumption is based on the premise that the major source of image corrupting information is the thermal noise generated by the imaging system.

We want to derive a one-dimensional convolution filter $f(x)$ which will detect this edge when used along the direction perpendicular to the direction of the edge. This filter must have the following properties:

1) It must be antisymmetric, i.e., $f(x) = -f(-x)$ and thus $f(0) = 0$. This follows from the fact that we want to use it to detect an antisymmetric feature [3].
2) It must be of finite extent going smoothly to zero at its ends: $f(\pm w) = 0, f'(\pm w) = 0$ and $f(x) = 0$ for $|x| \geq w$, where w is the half width of the filter.
3) It must have a given maximum amplitude $|k| : f(x_m) = k$ where x_m is defined by $f'(x_m) = 0$ and x_m is in the interval $(-w, 0)$, say.

Following Canny, we can derive the signal to noise ratio for this filter as being proportional to

$$S = \frac{\int_{-w}^{0} f(x)[1 - e^{sx}]\,dx}{\sqrt{\int_{-w}^{0} |f(x)|^2\,dx}} \tag{2}$$

Also, if we consider the good locality measure to be inversely proportional to the standard deviation of the distribution of points

Manuscript received August 20, 1989; revised October 31, 1990. Recommended for acceptance by C. Brown.

The authors are with the Department of Electronic and Electrical Engineering, University of Surrey, Guildford GU2 5XH, England.

IEEE Log Number 9042521.

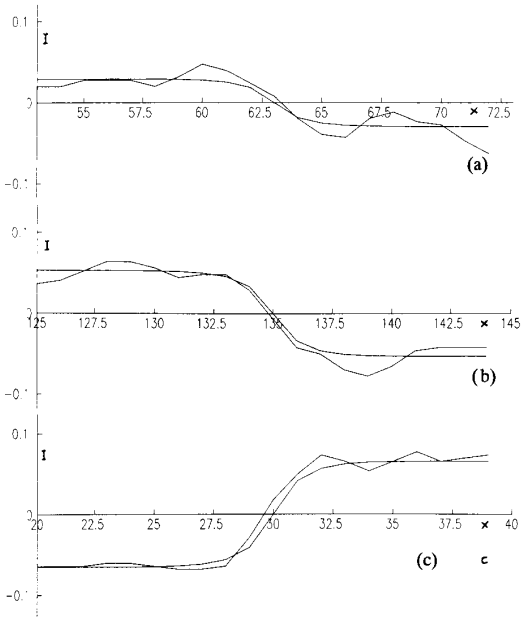

Fig. 1. Various "step" edges of Fig. 5(a) modeled by (1) with $s = 1$. (a) Face/eyebrow. (b) Face/hair. (c) Background/face.

where the edge is supposed to be, we can define it as

$$L = \frac{s^2 \int_{-w}^{0} f(x)e^{sx}\,dx}{\sqrt{\int_{-w}^{0} |f'(x)|^2\,dx}}. \tag{3}$$

It is worth clearing up at this point a controversy which seems to exist as to whether there is an uncertainty principle applicable to edge detection or not (see for example [1] and [8]). From the above two equations it is clear that one can choose a unique function $f(x)$ which will maximize both measures simultaneously [1], [3]. The uncertainty appears when one tries to choose the appropriate size of this function so that both measures attain maximal values. The larger the filter is, the better the signal to noise ratio and the smaller the filter is the better the localization of the edge will be [3].

Finally, we can define the measure for the suppression of false edges to be proportional to the mean distance between the neighboring maxima of the response of the filter to white Gaussian noise:

$$C = \frac{1}{w}\sqrt{\frac{\int_{-w}^{0} |f'(x)|^2\,dx}{\int_{-w}^{0} |f''(x)|^2\,dx}}. \tag{4}$$

Clearly, the optimal filter would be the one for which all these three performance measures have maximal values. It is impossible, however, to maximize all three simultaneously, so we follow Spacek [7] and define a combined performance measure of the filter given by

$$P = (SLC)^2 = \frac{s^4}{w^2}\frac{\left|\int_{-w}^{0} f(x)[1 - e^{sx}]\,dx \int_{-w}^{0} f(x)e^{sx}\,dx\right|^2}{\int_{-w}^{0} |f(x)|^2\,dx \int_{-w}^{0} |f''(x)|^2\,dx}. \tag{5}$$

In all the above definitions of the various performance measures we have omitted constant factors and factors which depend neither on the feature we want to detect nor on the filter we want to define.

Our purpose is to choose a function $f(x)$ which satisfies the boundary conditions (1)–(3) and at the same time maximizes the total performance measure P. This problem is addressed in the next section. However, the reader not interested in the mathematical details can find the derived filter in (15) and move on to Section VI which is concerned with its implementation.

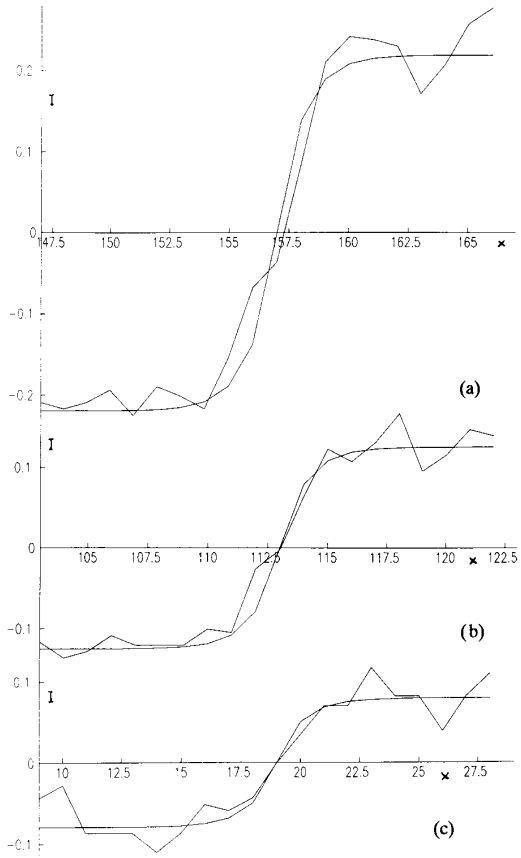

Fig. 2. Various "step" edges of Fig. 6(a) modeled by (1) with $s = 1$.

III. The Optimization Process

To choose a function which maximizes P is enough to extremize any one of the integrals appearing in its expression, assuming that the remaining integrals are constant [4]. We choose to minimize $\int_{-w}^{0} |f(x)|^2\,dx$ assuming that

$$\begin{aligned} \int_{-w}^{0} f(x)[1 - e^{sx}]\,dx &= c_1, \\ \int_{-w}^{0} f(x)e^{sx}\,dx &= c_2, \\ \int_{-w}^{0} |f''(x)|^2\,dx &= c_3, \end{aligned} \tag{6}$$

where c_1, c_2, and c_3 are some arbitrary constants.

Using the method of Lagrange multipliers we define the function $Z(f, f', f'')$ as

$$Z(f, f', f'') = f(x)^2 + \lambda_1 f(x)(1 - e^{sx}) + \lambda_2 f(x)e^{sx} + \lambda_3 f(x)''^2, \tag{7}$$

where λ_1, λ_2, and λ_3 are the Lagrange multipliers. This function must satisfy Euler's equation:

$$Z_f - \frac{d}{dx}Z_{f'} + \frac{d^2}{dx^2}Z_{f''} = 0. \tag{8}$$

By substitution, we derive the differential equation for the optimal filter:

$$f(x) + \mu_1 f(x)'''' = \mu_2 + \mu_3 e^{sx}, \tag{9}$$

where μ_1, μ_2, and μ_3 are some complex constants. The general solution of the above differential equation is then given by

$$\begin{aligned} f(x) = {} & A_1 e^{\beta x}[\cos(\alpha x) + i\sin(\alpha x)] \\ & + A_2 e^{-\alpha x}[\cos(\beta x) + i\sin(\beta x)] \\ & + A_3 e^{-\beta x}[\cos(\alpha x) - i\sin(\alpha x)] \\ & + A_4 e^{\alpha x}[\cos(\beta x) - i\ \sin(\beta x)] \\ & + A_5 + A_6 e^{sx}, \qquad \text{for } -w \le x \le 0. \end{aligned} \tag{10}$$

where α, β are real and $A_1 - A_6$ are complex.

This is a complex filter which depends on fourteen real parameters. We shall choose some of these parameters so that the imaginary part of the filter vanishes identically. If we use superscripts R and I to indicate the real and the imaginary parts of a quantity, we obtain

$$\begin{aligned} f^R(x) = {} & e^{\beta x}\left[A_1^R\cos(\alpha x) - A_1^I\sin(\alpha x)\right] \\ & + e^{-\alpha x}\left[A_2^R\cos(\beta x) - A_2^I\sin(\beta x)\right] \\ & + e^{-\beta x}\left[A_3^R\cos(\alpha x) + A_3^I\sin(\alpha x)\right] \\ & + e^{\alpha x}\left[A_4^R\cos(\beta x) + A_4^I\sin(\beta x)\right] + A_5^R + A_6^R e^{sx}, \end{aligned} \tag{11}$$

$$\begin{aligned} f^I(x) = {} & e^{\beta x}\left[A_1^I\cos(\alpha x) + A_1^R\sin(\alpha x)\right] \\ & + e^{-\alpha x}\left[A_2^I\cos(\beta x) + A_2^R\sin(\beta x)\right] \\ & + e^{-\beta x}\left[A_3^I\cos(\alpha x) - A_3^R\sin(\alpha x)\right] \\ & + e^{\alpha x}\left[A_4^I\cos(\beta x) - A_4^R\sin(\beta x)\right] + A_5^I + A_6^I e^{sx}. \end{aligned} \tag{12}$$

The trivial solution of the equation $f^I(x) = 0$ is the vanishing of all its coefficients. This, in combination with the antisymmetry condition (1) leads to the filter:

$$f(x) = \begin{cases} B(1 - e^{sx}) & \text{for } -w < x \le 0 \\ 0 & \text{for } -\infty < x \le -w, \end{cases} \tag{13}$$

where B can be chosen appropriately so that condition (3) is satisfied and the antisymmetry condition can be used to define the filter for positive x. This is a matched filter corresponding to the difference of boxes operator in the ideal step edge case. It attains its maximum value at $x_m = -w$, where, however, its first derivative is not zero. One can use this function naively in (5) to calculate the performance measure of this filter. The result, compared to the best results obtained for the filters we shall discuss in the next section, will appear spectacularly better. This is not the case, however, as the assumption of the vanishing first derivative of the filter at $\pm w$ has implicitly been built in the derivation of the performance measure P. So this solution of the optimization process has to be rejected on the grounds that it does not satisfy the boundary conditions.

The nontrivial way to satisfy equation $f^I(x) = 0$ is to require the separate vanishing of its x-dependent part and its x-independent part. The latter immediately leads to the choice $A_5^I = 0$. The vanishing of the x-dependent part for *every* x is not as easy. It can be achieved in the following two cases:

1) $\alpha = 0$ (or $\beta = 0$) which leads to

$$A_1^I = A_3^I = A_5^I = A_6^I = 0,$$
$$A_2^I + A_4^I = 0, \quad A_2^R - A_4^R = 0.$$

Using these in (10) we obtain the solution

$$\begin{aligned} f(x) = {} & K_1 e^{Ax} + K_2 e^{-Ax} + K_3\cos(Ax) \\ & + K_4\sin(Ax) + K_5 + K_6 e^{sx}. \end{aligned} \tag{14}$$

2) $\alpha = \beta$ (or $\alpha = -\beta$) which leads to

$$A_1^I = -A_4^I, \quad A_1^R = A_4^R, \quad A_2^I = -A_3^I,$$
$$A_2^I = A_3^I, \quad A_5^I = A_6^I = 0.$$

Using these conditions in (10) we obtain the solution

$$\begin{aligned} f(x) = {} & e^{Ax}[K_1\sin(Ax) + K_2\cos(Ax)] \\ & + e^{-Ax}[K_3\sin(Ax) + K_4\cos(Ax)] + K_5 + K_6 e^{sx}. \end{aligned} \tag{15}$$

The parameters K_1–K_6 and A which appear in the above two solutions are *real* quantities to be determined so that $f(x)$ satisfies conditions (1)–(3) and renders the value of P maximal.

IV. Choosing the Values of the Filter Parameters

The main difference between the case of ideal step edges and any other edge one might want to detect, is that ideal steps lack an intrinsic length-scale, i.e., the problem itself does not define a measure of distance. Filters for ideal step edges then can be developed in a scale-free manner and the user can scale them up or down accordingly, without missing anything in the performance of the filter. This, of course, is only in theory, because in reality one almost never has to detetct ideal step edges, unless it is the case of a simulated image.

In the problem we are considering here, the intrinsic length-scale is set by the parameter s which has dimensions [length]$^{-1}$. Thus, the size of the filter w is a parameter which has to be optimally chosen like the other parameters. By choosing s^{-1} as our length unit, we can measure all distances in terms of it. So, we shall set $s = 1$ from now on and determine the parameters of the filter that way. At the end, if one wants to use a filter with different s, one will have to divide the value of w with the desired value of s and to multiply the value of A with the desired value of s. The rest of the parameters of the filter remain unaffected as they are dimensionless.

We can reduce the number of free parameters which appear in the expression of $f(x)$ by making use of the three boundary conditions:

1) We can use $f(0) = 0$ to express the additive constant K_5 in terms of the other parameters.
2) By substituting the value of K_5 in the equations of condition (2), we obtain:

$$f(-w; K_1, K_2, K_3, K_4, K_6, A) = 0,$$
$$f'(-w; K_1, K_2, K_3, K_4, K_6, A) = 0.$$

This is a system of two linear equations for K_1, K_2, say, and it can be solved to express these two parameters in terms of the remaining five.

3) Substitution of the expressions for K_1, K_2, and K_5 in the equations for condition (3) yields the equations

$$f'(x_m; K_3, K_4, K_6, A) = 0,$$
$$f(x_m; K_3, K_4, K_6, A) = k.$$

TABLE I
PARAMETER VALUES FOR THE FILTER FOR RAMP EDGES

w	K_1	K_2	K_3	K_4	K_5	K_6	A	P
2	−0.028880	1.737065	−0.113414	−0.055020	−1.332045	−0.35	1.42	0.002283
3	0.556678	2.304465	−0.134374	−0.038797	−1.165668	−1.1	0.9	0.008851
4	0.924604	2.099287	−0.104453	−0.041305	−1.057982	−1.0	0.7	0.019114
5	1.400267	1.927899	−0.096130	−0.036598	−0.951301	−0.94	0.56	0.031085
6	1.561026	1.672333	−0.088292	−0.035464	−0.906869	−0.73	0.47	0.043196
7	1.582743	1.468015	−0.077481	−0.035918	−0.872097	−0.56	0.41	0.054604

We can choose k to be anything we like, because the performance measure P is independent of the amplitude of the filter. We choose $k = -1$ and solve the above system of equations for K_3, K_4 in terms of K_6, A, w and x_m.

Then, we substitute into the expression for $f(x)$ [equation (14) or (15)], and thus, we manage to reduce the number of free parameters on which $f(x)$ depends from eight, down to four. Finally, by substitution in the expression for P, we have a function $P(K_6, A, w, x_m)$ which has to be maximized with respect to its four variables.

The above calculation is far too tedious to be done manually, so algebraic computing was used and in particular the system REDUCE.

The surface $P(K_6, A, w, x_m)$ is not a smooth surface with one maximum. It is rather a corrugated one with lots of local maxima. So, the choice of the values of its arguments which render it maximum cannot be done automatically. Instead, the following procedure was used: The value of w was fixed and $P(w; K_6, A, x_m)$ was computed at the knots of a grid of points spanning the three-dimensional space (K_6, A, x_m). The grid became progressively denser as we were homing into the global maximum of P. At every stage, three profiles (one versus each of the three parameters) of the function $P(w; K_6, A, x_m)$ were plotted, so that the area where P attained its peak value was identified and the grid for the next stage was placed in that region. The process was repeated for various values of w.

The optimal filter is the one represented by (15) (for $-w \le x \le 0$) with parameter values given in Table I for various fixed filter sizes. The last column of the table gives the value of the performance measure for each filter. One might have expected that there would be an optimal filter size w for detecting a given ramp edge. This is not the case, however, as it can be seen from the values of P which increase for increasing w.

This is not difficult to understand: The larger the filter, the more insignificant the detailed structure of the edge at the center becomes and the more the edge "looks" to the filter like an ideal step edge. So, the better it will be detected by the filter as it appears more abrupt to it. This is further reflected in the values of the coefficient K_6 which multiplies the term of the filter that depends on the edge itself: the larger the filter, the smaller the value of K_6 becomes. In the limit when $w \to \infty$, one would expect to recover the optimal filter for ideal step edges.

V. The Optimal Filter For Ideal Step Edges

When w is very large, the value of $|x|$ can be arbitrarily high and it is clear from equation (15) that the e^{-Ax} term will explode unless $A \to 0$. To examine, therefore, the limiting case of $w \to \infty$, $A \to 0$, we define a new parameter for the problem $t \equiv Aw$, which remains finite as $w \to \infty$ and $A \to 0$, and a new independent variable $y \equiv Ax$. Equation (15) then becomes

$$f(y) = e^{y}(K_1 \sin y + K_2 \cos y) + e^{-y}(K_3 \sin y + K_4 \cos y) + K_5 + K_6 e^{y/A}, \quad \text{for } -t \le y \le 0. \tag{16}$$

From the above expression it is clear that as we approach the limit the term that depends on the feature we want to detect vanishes exponentially. So, in the discussion that follows we shall set $K_6 = 0$. This filter then has the same form as the one derived by Spacek for ideal step edges. For the filter to be acceptable, conditions (1)–(3) must also be satisfied. In the present case these conditions are modified as follows:

1) $f(0) = 0$ and $f(y) = -f(-y)$ for $y > 0$.
2) $f(-t) = 0, f'(-t) = 0$.
3) $f(y_m) = k, f'(y_m) = 0$ where $y_m = Ax_m$.

We can use these conditions as before to express the parameters K_1–K_5 in terms of t and y_m and thus arrive at an expression of the form $f(y; t, y_m)$.

Before we substitute into the expression for P we must see how the various integrals get modified in this limiting case. With a simple transformation of variables it is very easy to see that

$$\int_{-w}^{0} f(x)\,dx = \frac{1}{A}\int_{-t}^{0} f(y)\,dy \equiv \frac{1}{A} I_1$$

$$\int_{-w}^{0} f(x)e^{sx}\,dx = \frac{1}{A}\int_{-t}^{0} f(y)e^{y/A}\,dy \equiv \frac{1}{A}\tilde{I}_2$$

$$\int_{-w}^{0} |f(x)|^2\,dx = \frac{1}{A}\int_{-t}^{0} |f(y)|^2\,dy \equiv \frac{1}{A} I_3$$

$$\int_{-w}^{0} |f'(x)|^2\,dx = A\int_{-t}^{0} |f'(y)|^2\,dy \equiv A I_4$$

$$\int_{-t}^{0} |f''(x)|^2\,dx = A^3\int_{-t}^{0} |f''(y)|^2\,dy \equiv A^3 I_5,$$

where the above expressions are also used to define the quantities I_1, $\tilde{I}_2$, I_3, I_4 and I_5. The only one of these quantities which still depends on A is $\tilde{I}_2$ and we can deal with this as follows: if we perform the integration involved by parts and use the fact that $f(0) = f(-t) = 0$, we obtain

$$\frac{1}{A}\tilde{I}_2 = -\int_{-t}^{0} e^{y/A} f'(y)\,dy.$$

The factor $e^{y/A}$ is varying very fast over the range of integration, while $f'(y)$ varies much less. Since the main contribution to the integral comes from the point $y = 0$, we may assume that $f'(y) \sim f'(0)$ and take this factor out of the integral. Then we have

$$\frac{1}{A}\tilde{I}_2 = -f'(0)\int_{-t}^{0} e^{y/A}\,dy = -Af'(0) \equiv AI_2,$$

TABLE II
PARAMETER VALUES FOR THE SMOOTHING FILTER FOR RAMP EDGES

w	L_1	L_2	L_3	L_4	L_5	L_6	L_7	A
2	−0.6014735	−0.6218117	−0.02056117	−0.05930792	1.332045	0.35	1.497775	1.42
3	−1.589524	−0.9709929	−0.05309844	−0.09620634	1.165668	1.1	1.705717	0.9
4	−2.159923	−0.8390591	−0.04510553	−0.1041132	1.057982	1.0	2.259878	0.7
5	−2.971577	−0.4711000	−0.05315354	−0.1185067	0.9513013	0.94	2.533634	0.56
6	−3.439744	−0.1184117	−0.05619989	−0.1316546	0.9068693	0.73	2.974375	0.47
7	−3.720437	0.1399123	−0.05068549	−0.1382914	0.8720965	0.56	3.465760	0.41

where we have defined the quantity I_2.

All five quantities I_1–I_5 are now finite and well behaved and if we substitute the integrals in the expressions for the performance measures [equations (2)–(4)] we shall recover formulas similar to those used by Canny and Spacek for the ideal step edges. The overall performance measure P is given by

$$P = \frac{\left[\int_{-t}^{0} f(y)\,dy\right]^2 f'(0)^2}{t^2 \int_{-t}^{0} f(y)^2\,dy \int_{-t}^{0} f''(y)^2\,dy}.$$

If we substitute the equation for $f(y; t, y_m)$ and perform the integrations, we shall derive an expression for $P(t, y_m)$. We can then choose the values of the parameters t and y_m which will maximize P by numerically exploring the parameter space (t, y_m).

The $P(t, y_m)$ surface is a corrugated one, with a trough and a ridge. The trough is along the line $|y_m| = t/2$ which corresponds to filters which are symmetric about the point $y = y_m$ and have $f'(0) = 0$. If we use the definition of L as given by (3), these filters have $P = L = 0$. However, L can be expressed in terms of higher order derivatives of $f(y)$ if one starts again from the first principles and takes into consideration the vanishing of the first derivative of $f(y)$ at the point $y = 0$. Subsequently, the expression for P would change as well. We are not going to pursue this point any further as it is really of no importance.

The ridge of the surface $P(t, y_m)$ is approximately near the line $|y_m| = t/3$. The values of the parameters for which P is maximum are

$$K_1 = 1.316134, \quad K_2 = 0.8223482, \quad K_3 = -0.03800267$$
$$K_4 = -0.03552345, \quad K_5 = -0.78682475, \quad t = 3.16.$$

We have again chosen the value $k = -1$. If one wants to implement this filter, one should decide first upon the desired value of w and work out the value of A from the expression $t = Aw$. Then for negative x formula (15) should be used with $K_6 = 0$ and the values of the remaining parameters those given above. For x positive equation $f(x) = -f(-x)$ should be utilized.

VI. Implementation Considerations

All the above analysis concerns one-dimensional continuous filters. The extension to two dimensions is not trivial. For a start, there is no theoretically sound method by which one can derive an optimal filter in two dimensions given an optimal filter in one dimension.

At a much more elementary level, the filter we derived is antisymmetric and there is no way to circularize it as it stands. There is one way around this problem: we can integrate our filter to derive the optimal smoothing filter. Once the signal has been convolved with this smoothing filter, the edges can be detected by direct differentiation of the result without any loss of optimality [9]. Then one can replace the variable x by the variable r of the polar coordinates in the expression of the smoothing filter and have an axially symmetric two-dimensional filter. To locate the edges in an image which has been smoothed by this two-dimensional filter we must calculate first the derivative of the intensity in two orthogonal directions and thus find the magnitude and direction of the gradient at every position in the image. The edges will correspond to the maxima of the gradient values. In practice, of course, we have another problem too: the signal is not continuous, so a discretized version of the filter is required. Given that the filter is of finite impulse response, its bandwidth is infinite, so we are bound to lose some of its optimality by sampling it.

It is generally accepted, however, that if the one-dimensional filter is good and the sampling dense enough, the corresponding two-dimensional filter obtained by circularization will be good too [6].

The one-dimensional smoothing filter for ramps is given by

$$h(x) = e^{Ax}[L_1 \sin(Ax) + L_2 \cos(Ax)] + e^{-Ax}[L_3 \sin(Ax) + L_4 \cos(Ax)] + L_5 x + L_6 e^{sx} + L_7, \quad \text{for } -w \le x \le 0. \tag{17}$$

The filter is symmetric about the point $x = 0$. Table II gives the values of the parameters L_1–L_7 for $s = 1$ and various values of w. The integration constant L_7 has been chosen so that $h(\pm w) = 0$. If one is interested in the filter for a different value of s, one should scale everything as follows. The values of w, L_1, L_2, L_3, L_4, L_6, and L_7 are to be divided by the desired value of s while the value of A is to be multiplied by the desired value of s. L_5 remains the same.

We can calculate the Fourier spectrum of this filter is we consider it to be a periodic function with period $2w$. The filter will be given by:

$$h(x) = C_0 + \sum_{n=1}^{\infty} C_n \cos(n\pi x/w), \tag{18}$$

with the values of the Fourier coefficients up to the 10th harmonic given in Table III for $s = 1$ and various values of w. Although this filter is of infinite bandwidth, its Fourier coefficients from the 7th harmonic onwards are quite negligible as can be seen from Table III. We may assume, therefore, that our filter is band limited with cut-off frequency $\Omega_f = 6\pi/w$. According to the sampling theorem then, the filter will retain all its properties if it is sampled with frequency $2\pi/T \ge 2\Omega_f$ where T is the period of the sampling function, in our case the interpixel distance. The above inequality means that we must have $w/T \ge 6$, so that the filter must be at least 12 pixels wide in order to preserve its optimal properties. It is clear that such a filter will blur out details which have frequency higher than Ω_f or period less than $w/3$. One must, therefore, choose w so that the filter is adequately sampled ($w \ge 6$ pixels) and no desired detail is lost ($w \le 3\tau$ where τ is the period (in pixels) of the details one wants to detect). Of course, if one wants to detect very dense features ($\tau = 1$), like steps of a staircase, say, it is impossible to satisfy both conditions. This, however, is an extreme case.

Having chosen w, the value of s has to be decided. From the ramp model [equation (1)], it is clear that $s/2$ is the slope of the ramp at its middle point, divided by the total jump in the intensity across the

TABLE III
FOURIER COEFFICIENTS OF THE SMOOTHING FILTERS FOR RAMP EDGES

w	C_0	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
2	0.51872	0.59338	0.05927	−0.00791	0.00406	−0.00143	0.00080	−0.00040	0.00025	−0.00015	0.00010
3	0.76234	0.88215	0.09921	−0.01004	0.00592	−0.00189	0.00113	−0.00054	0.00035	−0.00020	0.00014
4	1.00928	1.17316	0.13868	−0.01204	0.00802	−0.00237	0.00150	−0.00068	0.00046	−0.00026	0.00019
5	1.24302	1.45656	0.18603	−0.01048	0.01009	−0.00263	0.00183	−0.00079	0.00056	−0.00030	0.00022
6	1.47795	1.74066	0.23334	−0.00903	0.01243	−0.00285	0.00220	−0.00088	0.00066	−0.00035	0.00036
7	1.71322	2.02511	0.28137	−0.00662	0.01505	−0.00299	0.00261	−0.00097	0.00077	−0.00039	0.00031

TABLE IV
WEIGHTS FOR SMOOTHING FILTERS FOR RAMP EDGES WITH WIDTH d INTERPIXEL DISTANCES (OR SLOPE AT THE CENTRAL POINT $s/2$ WHEN THE JUMP IN GRAY LEVEL VALUES ACROSS THE EDGE IS NORMALIZED TO 1). THE FILTERS ARE $(2w-1)\times(2w-1)$ IN SIZE. THE NUMBERS GIVEN ARE 100 TIMES THE ACTUAL WEIGHTS TO INCLUDE AS MANY SIGNIFICANT FIGURES AS POSSIBLE. A COMPLETE MASK CAN BE CONSTRUCTED BY REFLECTION ABOUT THE DIAGONAL TO CREATE A QUARTER OF THE FILTER AND REFLECTION AGAIN ABOUT THE TOP ROW AND THE LEFT COLUMN SUCCESSIVELY TO CREATE THE WHOLE MASK.

4.104123	3.63454	2.56577	1.41549	0.53485	0.08423
	3.23228	2.28535	1.24773	0.45502	0.06300
		1.60099	0.83549	0.26622	0.02135
$s = 2$			0.38409	0.08423	0.00049
$w = 6$				0.00394	0.00000
$d = 1$					
3.02828	2.76298	2.12973	1.38738	0.72760	0.26500	0.04048
	2.52959	1.95586	1.27080	0.65846	0.23217	0.03175
		1.51317	0.96905	0.48031	0.15077	0.01330
$s = 2$			0.59440	0.26500	0.06254	0.00116
$w = 7$				0.09161	0.00932	0.00000
$d = 1$						
3.95123	3.54210	2.55610	1.43591	0.54806	0.08650
	3.17855	2.28788	1.26853	0.46657	0.06469
		1.61991	0.85359	0.27332	0.02191
$s = 1$			0.39405	0.08650	0.00050
$w = 6$				0.00404	0.00000
$d = 2$					
2.93018	2.69807	2.11530	1.39782	0.73897	0.26950	0.04099
	2.48724	1.95006	1.28262	0.66910	0.23607	0.03213
		1.52143	0.98190	0.48851	0.15317	0.01343
$s = 1$			0.60424	0.26950	0.06339	0.00116
$w = 7$				0.09295	0.00941	0.00000
$d = 2$						
3.83311	3.46058	2.53746	1.45179	0.56374	0.09028
	3.12403	2.28095	1.28632	0.48086	0.06760
		1.63271	0.87229	0.28319	0.02298
$s = 0.667$			0.40687	0.09028	0.00053
$w = 6$				0.00425	0.00000
$d = 3$					
2.85322	2.64109	2.09522	1.40300	0.75058	0.27631	0.04230
	2.44556	1.93761	1.29004	0.68049	0.24222	0.03318
		1.52367	0.99292	0.49855	0.15748	0.01389
$s = 0.667$			0.61528	0.27631	0.06536	0.00121
$w = 7$				0.09574	0.00973	0.00000
$d = 3$						
3.79245	3.43309	2.53211	1.45757	0.56859	0.09129
	3.10617	2.27956	1.29256	0.48521	0.06837
		1.63761	0.87841	0.28605	0.02324
$s = 0.5$			0.41071	0.09129	0.00053
$w = 6$				0.00430	0.00000
$d = 4$					
2.79463	2.59565	2.07650	1.40541	0.76035	0.28292	0.04373
	2.41079	1.92491	1.29461	0.69029	0.24826	0.03433
		1.52338	1.00142	0.50766	0.16188	0.01440
$s = 0.5$			0.62496	0.28292	0.06747	0.00125
$w = 7$				0.09867	0.01009	0.00000
$d = 4$						

edge. Since s is a characteristic of the imaging device it does not need to be determined for each image separately. It will be the same for all edges which were step edges in the scene, in all images obtained by the same device. Once s has been obtained, the product ws will give the value of w in units of s^{-1} needed. The filter with the nearest value of w can be chosen from Table II and appropriately scaled for the desired value of s as described earlier. Then the convolution filter $2w \times 2w$ can be calculated using (17) with the variable x replaced by the polar radius r. At the end, the values of the discrete mask can be divided by an appropriately chosen constant, so that their sum is 1 and the smoothing filter does not distort the image in the absence of edges and noise. For practical purposes we give a few masks for four values of s and two values of w in Table IV.

VII. DISCUSSIONS AND CONCLUSIONS

We argue here that the best approach in developing an optimal filter is to model an edge as an ideal mathematical function plus a

TABLE V
THE PERFORMANCE MEASURE OF THE TRUNCATED GAUSSIAN AS A PERCENTAGE OF THE PERFORMANCE MEASURE OF THE CORRESPONDING OPTIMAL FILTER FOR THE VARIOUS CASES OF TABLE IV. THE DISCONTINUITY OF THE GAUSSIAN AT THE POINT OF TRUNCATION IS THE SAME AS THE DISCONTINUITY OF THE CORRESPONDING OPTIMAL FILTER DUE TO SUBSAMPLING

s	d	$w = 6$	$w = 7$
2	1	68%	72%
1	2	58%	64%
0.667	3	51%	57%
0.5	4	46%	52%

low pass filter plus noise. We consider unrealistic the omission of the low pass filter in other similar approaches to the problem, namely Canny's and Spacek's. Inevitably, in order to develop specific filters one has to adopt a specific function for each one of the above three items. We developed the theory around the specific case of an ideal step edge passed through an exponential low pass filter, which is the first approximation to any low pass filter, and immersed in white Gaussian noise. We derived the optimal filter for ideal step edges as a limiting case of the optimal filter for ramps. Our approach is based on Spacek's work designed to overcome the shortcomings of Canny's method. The performance measure of our filter for ideal step edges is $P = 0.1849$, which is higher than the performance measures of Canny's best filter ($P = 0.183$). This is because we maximized a combination of all three performance measures S, L and C. What Canny did was to maximize SL and choose the parameter values so that the error of the operator due to false maxima was minimal when expressed as a percentage of the error due to thresholding. This is not equivalent to maximizing SL and C as can be seen from the last line entry of the table in Canny's paper [3]. Spacek's original filter had even lower value of the performance measure $P = 0.1824$ [9]. His revised filter, however, had performance measure comparable to ours, although that result was never widely demonstrated [10].

In practice, Canny never implemented his optimal filter. He used instead the derivative of a Gaussian which, when assumed to be of infinite extent, has a performance measure SL equal to 0.92. We do not consider, however, this to be a meaningful comparison. When the Gaussian filter is implemented, it is always truncated at some finite half-width w. The standard deviation of the Gaussian can always be chosen so that the discontinuity at w can be almost anything. To calculate the value of P for such a filter is not strictly speaking correct, because, as was mentioned earlier, the smooth vanishing of $f(x)$ at $\pm w$ is inbuilt in the derivation of P. Even so, the derivative of a Gaussian is a very widely used filter and that is what is usually meant by the term "the Canny filter." To calculate its performance measure when used to detect real edges, we did the following: For each of the filters of Table IV we calculated the discontinuity at $\pm w$ due to subsampling and chose the standard deviation of the Gaussian so that it has the same discontinuity. Table V gives the performance measure of the Gaussian filter as a percentage of the performance measure of the corresponding optimal filter. It is clear that as the width of the ramp increases, the performance of the Gaussian filter becomes worse when compared to the performance of the optimal filter.

Another filter we can compare our work with is the cubic spline which was suggested by Spacek. If one considers only the three boundary conditions for $f(x)$ and assumes a polynomial expression for it, one can derive a filter of the form:

$$cs(x) = \frac{27k}{16}(x^3 + 2x^2 + x).$$

TABLE VI
THE PERFORMANCE MEASURE OF THE CUBIC SPLINE FILTER AS A PERCENTAGE OF THE PERFORMANCE MEASURE OF THE CORRESPONDING OPTIMAL FILTER FOR THE VARIOUS CASES OF TABLE IV.

s	d	$w = 6$	$w = 7$
2	1	84%	86%
1	2	73%	76%
0.667	3	66%	68%
0.5	4	61%	63%

It is very easy to calculate the performance measures for this filter and these are given in Table VI. From the entries of Tables V and VI it is obvious that the cubic spline filter is expected to perform better than the Gaussian filter although this filter is not determined by any optimization process but it only reflects the boundary conditions we imposed on our filter.

It would be interesting, however, to compare the various filters in practice. So we designed the simulated image of a stamp which contains ramp edges of characteristic slope $s = 0.43$, emersed in 100% Gaussian noise (this means that the standard deviation of the noise is the same as the jump in the intensity across every edge in the image). Fig. 3 shows the original image, the edges we try to recover and the result of running an edge detection program using a Gaussian filter (as suggested by Canny), the cubic spline filter (as defined above), the optimal filter for step edges and the corresponding optimal filter for ramps. All filters were implemented in the same program and the same threshold was used for all outputs. All masks were 13×13 pixels large. For comparison, the perforations in the image of the stamp are 8×8 pixels. These results clearly show how poorly the Gaussian filter performs when used to detect edges in blurred images. The optimal ramp filter is clearly the best of all four filters. It must be stressed here that neither hysteresis thresholding nor salt and pepper noise removal were used. We are only comparing the linear parts of the edge detection problem. It is our opinion that the nonlinear characteristics are those which make the Canny algorithm for detecting edges (as opposed to "Canny filter" which invariably is taken to mean a Gaussian filter) so superior to other edge detectors. It is worth noting that the cubic spline filter, which is almost as good as the optimal one for step edges, is not the result of any optimization process. It is determined only by the boundary conditions we impose on the filter. It is also worth noting that the $P(t, y_m)$ surface of the filter for step edges has its maxima in the (t, y_m) space almost along the line $|y_m| = t/3$ which is an equation exactly satisfied by the cubic spline filter. These two observations lead to the conclusion that this type of approach is heavily dominated by the boundary conditions.

Deriche's filter is infinite in extent too and should be implemented recursively. If, however, is truncated at a low enough value, it can be implemented within our scheme. We adopted the parameter values Deriche considers best ($\alpha = 1$, $\omega = 0.0001$) and chose the size of this filter to be 17×17 so that its discontinuity at the truncation was of the order 0.001. Fig. 4 shows the result of using it on the stamp image. The only postprocessing used was the same as that for the other filters, i.e., thresholding at a specific value. Clearly the filter has poor signal to noise ratio in comparison to the other filters. This is to be expected, as its shape is almost like a box operator at the center. One could improve its performance by experimenting with the values of the two filter parameters, and one could do the same by experimenting with the standard deviation of the Gaussian filter too. The whole point is that there is no guidance for these filters as to which are the best values of the parameters involved.

These results are confirmed by applying the various filters on real images. The images were specifically chosen to be bad because any

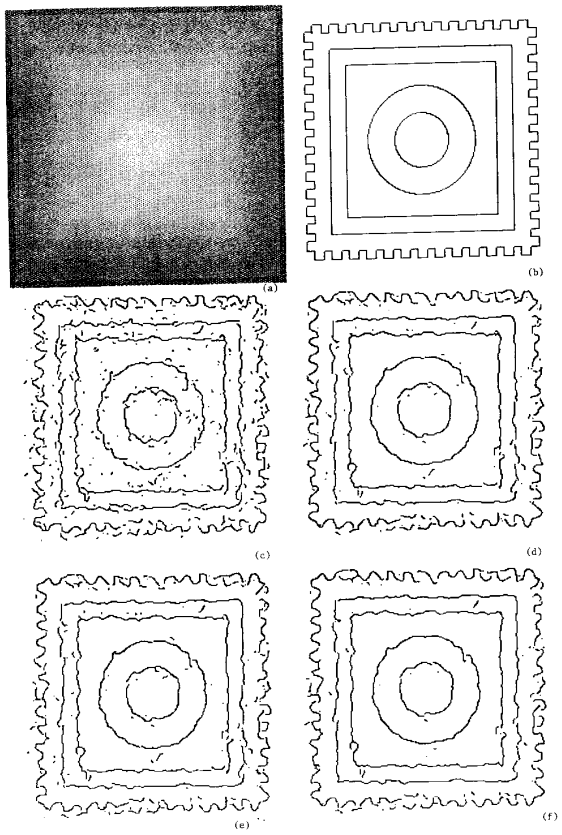

Fig. 3. (a) The simulated stamp image with 100% Gaussian noise. (b) The edges of the stamp. (c), (d), (e), (f) The edges detected by a 13 × 13 Gaussian filter, the optimal filter for step edges, the cubic spline filter and the appropriate optimal filter for ramp edges, respectively.

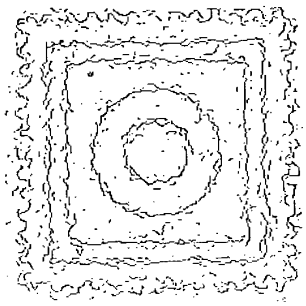

Fig. 4. The result of applying the Deriche filter to Fig. 3(a).

simple off the shelf operator would perform well on a good image. Fig. 5 shows the results of using a Gaussian filter, the wrong ramp filter and the appropriate ramp filter on the real image of a face. Fig. 6 shows similar results of applying the Gaussian filter, the filter for ideal step edges and the proper ramp filter to the real image of an industrial part. For both images the appropriate ramp filter was chosen by modeling a few step edges of each image as shown in Figs. 1 and 2. In both cases the Gaussian filter is clearly the worse

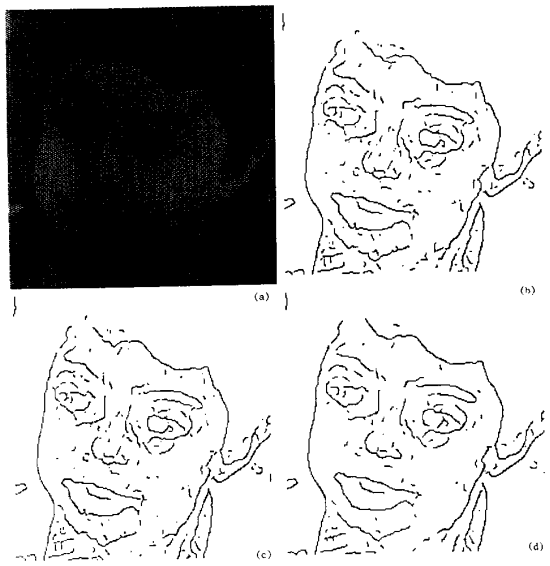

Fig. 5. (a) Original image. (b) Gaussian filter. (c) Optimal filter for ramp edges with $s = 0.5$. (d) Optimal filter for ramps edges with $s = 1$.

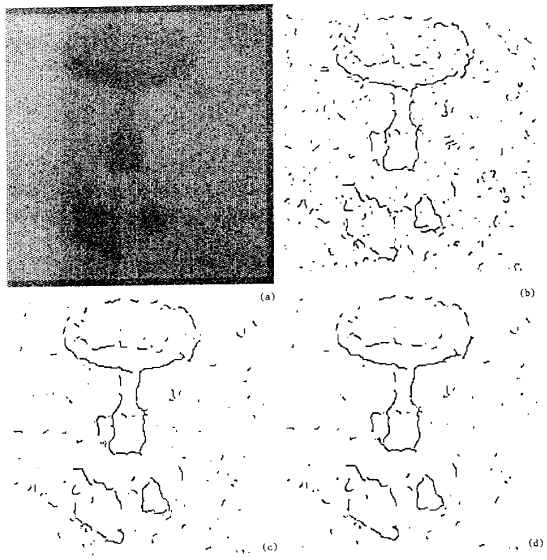

Fig. 6. (a) Original image. (b) Gaussian filter. (c) Optimal filter for ideal step edges. (d) Optimal filter for ramps.

while the proper ramp filters produce the best results.

Finally, we summarize the main points of this paper:

1) We showed that real image edges can be well modeled by function (1) and that they are never ideal steps.
2) The performance of the Gaussian filter varies significantly according to the value of the standard deviation chosen and is worse than the performance of the optimal filters for ramp edges.
3) The performance of the Gaussian filter improves as the edges become more like step edges, but it is unrealistic to expect image edges to be step-like.

4) We presented optimal filters for various types of ramp edges and showed them to be better for noisy and blurred images than the optimal filter for ideal step edges and much better than the commonly used Gaussian filter.

REFERENCES

[1] R. A. Boie, I. J. Cox and P. Rehak, "On optimum edge recognition using matched filters," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 1986, p. 100.

[2] J. Canny, "Finding edges and lines in images," MIT AI Lab., Tec. Rep. 720, 1983.

[3] ——, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-8, p. 679, Nov. 1986.

[4] R. Courant and D. Hilbert, *Methods of Mathematical Physics,* vol. 1. New York: Wiley Interscience, 1953.

[5] R. Deriche, "Using Canny's criteria to derive a recursively implemented optimal edge detector," *Int. J. Comput. Vision,* vol. 1, no. 2, May 1987.

[6] J. G. Fiasconaro, "Two-dimensional nonrecursive filters," in *Topics in Applied Physics, Vol. 6: Picture Processing and Digital Filtering.,* T. S. Huang, Ed. Springer-Verlag, 1979, p. 69.

[7] D. Lee, "Edge detection, classification and measurement," in *Proc. IEEE Comput. Soc. Conf. Computer Vision and Pattern Recognition,* June 1989, p. 2.

[8] V. S. Nalwa and T. O. Binford, "On detecting edges," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-8, p. 699, Nov. 1986.

[9] L. A. Spacek, "Edge detection and motion detection," *Image Vision Comput.,* vol. 4, p. 43, 1986.

[10] ——, "The detection of contours and their visual motion," Ph.D. dissertation, Univ. Essex, 1986.

[11] R. Wilson and M. Spann, *Image Segmentation and Uncertainty.* New York: Wiley (Research Studies Press), 1987.

Positioning Quadric Surfaces in an Active Stereo Imaging System

Joseph H. Nurre and Ernest L. Hall

Abstract—**The location of surfaces using stereo imaging techniques is an important area of research for robot guidance and machine inspection applications. This correspondence will investigate the underlying geometry of finite focal length stereo pinhole cameras. This is the model used in both active and passive stereo imaging systems. It is shown that the points of intersecting views from the pinhole models result in conic sections. This information is used to locate quadric surfaces in the inspection space. This work provides a basis for understanding the underlying geometry of stereo vision and uses that understanding for solving practical positioning problems.**

Index Terms—**Computer vision, conic sections, pinhole model, projection moiré, Ronchi grating, stereo imaging, triangulation.**

I. INTRODUCTION

The active stereo imaging technique is a universally recognized practical method for determining deformations of curved surfaces. This technique, also referred to as projection moiré, is commonly used in the fields of vibration analysis and stress analysis, where an object is acted upon by some force [1]. It has been demonstrated, however, that the method can also be used for inspection purposes [2]. The advantages of using projection moiré as an inspection technique include the fact that it is a noncontact and nondestructive method for making three dimensional measurements on smooth surfaces.

One commonly discussed method of inspection is to project a master grating of constant pitch onto a model surface and record the image from a different perspective. This generates an encoded grating. Objects to be inspected are then positioned in the same location with an identical master grating projected. A comparison of the encoded grating, with the grating generated by the new object at the same perspective, gives the location of any surface flaws.

A modification has been suggested to the method described above [3]. In order to simplify the comparison of gratings, the encoded grating is projected from the location of its conception, resulting in a duplication of the master grating at the location of its origin. In other words the encoded grating is projected onto the object to be inspected, and the resulting image of a nondefective part should be that of the master grating with its straight parallel lines. Several authors have suggested that this model grating might be generated by a computer model of the object [4]–[6].

When using the technique described above, it is necessary to maintain the relative location of the light source, image apparatus, and object, in order to accurately locate flaws on the surface. Although not as sensitive to small misalignments as holography, a moiré analysis must maintain fixture tolerances based upon the pitch of the master grating. This may prove to be difficult when the model grating is generated by a computer rather than physical components. Furthermore, fixturing does not lend itself to rapid inspection of mass produced, randomly oriented parts. Therefore, there is a need to be able to interpret moiré fringes as a specific misalignment of an object. The analysis of the fringe pattern should allow us to reorient the part for proper inspection to begin.

A limited amount of research has been conducted to address the orientation problem. When the projection system uses a parallel focal axis model and the surface being inspected is flat, the orientation problem is greatly simplified [7]. This special case is addressed in Section II of this correspondence. For more general shapes, Michalski *et al.* [3] demonstrates that the fringes recorded from a misaligned part can be interpreted as directional derivatives; however, it is not clear how this information simplifies the positioning task.

This correspondence shall describe a method of fringe interpretation, which uses the underlying geometry of a finite focal length stereo imaging system, to determine the position of the object. Although simple shapes are used, the method could be used to position subsets of more complex objects. In Section II, we discuss the intersection space of moiré projections which indicate the best surfaces to be inspected. Section III describes a method of using the intersection space with a physically realizable system, while Section IV will present experimental results from an implemented system. Finally, Section V will present a discussion of the results.

II. RAY TRACING FOR A MOIRÉ PROJECTION SYSTEM

In this correspondence, we will use incoherent light and assume a grid pitch is such that a purely geometrical analysis can be performed using ray tracings. The camera and projector are both modeled as pinholes, where light converges to or from a focal point. The models are further defined by a focal axis which is a vector from the focal

Manuscript received May 26, 1989; revised October 31, 1990. Recommended for acceptance by J. L. Mundy. This work was supported in part by the National Science Foundation under Grants ECS 8609524 and DMC 8806313.

J. H. Nurre is with the Department of Electrical and Computer Engineering, Ohio University, Athens, OH 45701.

E. L. Hall is with the Center for Robotics Research, University of Cincinnati, Cincinnati, OH 45221.

IEEE Log Number 9042520.