An Isotropic 3 3 Image Gradient Operator

1 author:

Irwin Sobel
Consultant (HP Labs Retired - 8Mar13)

History and Definition of the so-called "Sobel Operator",
more appropriately named the
Sobel-Feldman Operator
by Irwin Sobel
February 2, 2014
Updated June 14 2015

The work was never "published" by me... however it was first described and credited in a footnote "suggested by I. Sobel" in the book:

- Duda,R. and Hart,P., Pattern Classification and Scene Analysis, John Wiley and Sons,'73, pp271-2

  The "suggestion" was made in a talk:
  Sobel, I., Feldman, G., "A 3x3 Isotropic Gradient Operator for Image Processing", presented at the Stanford Artificial Intelligence Project (SAIL) in 1968.

There also a detailed historical account that I gave to

Prof. Per-Erik Danielsson
Institutionen for Systemteknik
Department of Electrical Engineering
Universitetet i Linkoping
S-581 83 Linkoping, Sweden
email: ped@isy.liu.se

which he kindly published as an appendix to a paper of his:

Danielsson, P.E., Seger, O., "Generalized and Separable Sobel Operators", in "Machine vision for three-dimensional scenes", Herbert Freeman (ed), Academic Press (1990).

****
****************************** correspondence follows **********************
****

[corrected by IES 15Dec94 - G' scale factor is 4 not 2]

February 6, 1989

Prof. Per-Erik Danielsson
Institutionen for Systemteknik
Department of Electrical Engineering
Universitetet i Linkoping
S-581 83 Linkoping
Sweden

Dear Prof. Danielsson,
...

The history of the "Sobel Operator" according to my best recollection is as follows:

In 1968, while a PhD candidate at the Stanford Artificial Intelligence Project I gave a talk, together with Gary Feldman (another graduate student and good friend of mine) on a relatively isotropic 3x3 gradient operator. This talk was presented at a time when the major piece of published work on computer vision was Larry Roberts' PhD Thesis from MIT wherein he defined a 2x2 gradient estimator then referred to as the "Roberts Cross" operator. I had previously thought up the operator via a line of reasoning presented in the accompanied short document and discussed it with Gary who enthusiastically proceeded to help me program it and test it. After doing so and satisfying ourselves that it gave at least visually desireable results, we presented these in a seminar at the Stanford Artificial Intelligence Project where we were employed as Research Assistants.

As this was an event that occurred more than 20 years ago my memory of it is somewhat foggy. Faculty that I most clearly remember in attendance at the seminar were Raj Reddy, John McCarthy and Arthur Samuels. I'm pretty sure that Peter Hart and/or Dick Duda and possibly Nils Nilsson from SRI were also there. Lester Earnest, project executive officer was most likely there. I'm pretty sure Karl Pingle who was employed as a project programmer and later wrote an edge follower incorporating the operator was there. Manfred Hueckel, a graduate student who later wrote a paper on a more robust and computationally expensive "edge detector", was I think also there. Lynn Quam, Jay "Marty" Tenenbaum, Gunnar Grape, Gil Falk, Dave Poole, and Phil Petit were other graduate students with the project and either were at the seminar or were working in such close proximity that they knew of the results.

My synopsis of what ensued was that Raj Reddy, who was then teaching one of the first courses on Computer Vision, coined the term "Sobel Operator" in contrast to the "Roberts Cross" and used it in his course. Subsequently Pingle published a paper (1969) describing it as part of his edge follower, and Duda and Hart mentioned it in their book.

In response to your gentle prod I polished up a note I had started to write several years ago for publication (I knew not where), giving my rationale for the derivation. I enclose it herewith with the request that you submit it as a technical/historical note with your forthcoming paper.

Sincerely,

Irwin Sobel
HPLABS, Measurement and Manufacturing Research Center

**************************** difop note follows ****************************

An Isotropic 3x3 Image Gradient Operator

by Irwin Sobel

We would like to document the derivation of a simple, computationally efficient, gradient operator which we developed in 1968. This operator has been frequently used and referenced since that time. The earliest description of this operator in the computer vision literature is [1], although it has been more widely popularized by its appearance in [2]. Horn [3] defines this operator and references 4 numerical analysis texts [4-7] with the statement:

> "Numerical analysis [4-7] teaches us that for certain classes of surfaces an even better estimate is obtained using a weighted average of three such central differences ... These expressions produce excellent estimates for the components of the gradient of the central point".

The motivation to develop it was to get an efficiently computable gradient estimate which would be more isotropic than the then popular "Roberts Cross" operator [8]. The principle involved is that of estimating the gradient of a digitized picture at a point by the vector summation of the 4 possible simple central gradient estimates obtainable in a 3x3 neighborhood. The vector summation operation provides an averaging over directions-of-measurement of the gradient. If the density function was truly planar over the neighborhood all 4 gradients would have the same value. Any differences are deviations from local planarity of the function over the neighborhood. The intent here was to extract the direction of the "best" plane although no attempt was made to make this rigorous.

To be more specific, we will refer here to the image function as a "density" function. (It could just as well be an "intensity" function - the difference depends on the physical nature of the image source.) For a 3x3 neighborhood each simple central gradient estimate is a vector sum of a pair of orthogonal vectors. Each orthogonal vector is a directional derivative estimate multiplied by a unit vector specifying the derivative's direction. The vector sum of these 4 simple gradient estimates amounts to a vector sum of the 8 directional derivative vectors.

Thus for a point on a Cartesian grid and its eight neighbors having density values as shown

```
 ___________
| a | b | c |
|___|___|___|
| d | e | f |
|___|___|___|
| g | h | i |
|___|___|___|
```

we define the magnitude of the directional derivative estimate vector 'g' for a given neighbor as

$$|g| = \langle\text{density difference}\rangle / \langle\text{distance to neighbor}\rangle$$

The direction of 'g' will be given by the unit vector to the appropriate neighbor. Notice that the neighbors group into antipodal pairs: (a,i) (b,h) (c,g) (f,d). Vector summing derivative estimates within each pair causes all the "e" values to cancel leaving the following vector sum for our gradient estimate:

$$\begin{aligned} G = & (c-g)/4 * [\,1, 1] \\ & +(a-i)/4 * [-1, 1] \\ & +(b-h)/2 * [\,0, 1] \\ & +(f-d)/2 * [\,1, 0] \end{aligned}$$

the resultant vector being

$$G = [(c-g-a+i)/4 + (f-d)/2,\ (c-g+a-i)/4 + (b-h)/2]$$

Notice that the square root fortuituously drops out of the formula. If this were to be metrically correct we should divide result by 4 to get the average gradient. However, since these operations are typically done in fixed point on small integers and division loses low order significant bits, it is convenient rather to scale the vector by 4, thereby replacing the "divide by 4" (doubleshift right) with a "multiply by 4" (doubleshift left) which will preserve low order bits. This leaves us with an estimate which is 16 times as large as the average gradient. The resultant formula is:

$$G' = 4*G = [c-g-a+i + 2*(f-d),\ c-g+a-i + 2*(b-h)]$$

It is useful to express this as weighted density summations using the following weighting functions for x and y components:

-1	0	1
-2	0	2
-1	0	1

x-component

1	2	1
0	0	0
-1	-2	-1

y-component

This algorithm was used as an edgepoint detector in the 1968 vision system [2] at the Stanford Artificial Intelligence Laboratory wherein a point was considered an edgepoint if and only if

$$|G'|**2 > T$$

where T was a previously chosen threshold. For this purpose it proved an economical alternative to the more robust, but computationally expensive "Hueckel operator" [9].

References:

[1] Pingle, K.K., "Visual Perception by a Computer", in Automatic Interpretation and Classification of Images, A. Grasselli (Ed.), Academic Press, New York, 1969, pp. 277-284.

[2] Duda, R.O. and Hart, P.E., pp. 271-272 in Pattern Classification and Scene Analysis, John Wiley & Sons, New York, 1973

[3] Horn,B.K.P., "Hill Shading and the Reflectance Map", Originally in Proc. DARPA Workshop in Image Understanding, Apr 24-25,1979, p. 85 Science Applications Inc. Report SAI-80-895-WA; Later in Geo- Processing 2(1982) p. 74, Elsevier Scientific Publishing, Amsterdam.

[4] Conte,D. and de Boor, C., Elementary Numerical Analysis, 1972, New York: McGraw Hill.

[5] Hamming, R.W., Numerical Methods for Scientists and Engineers, 1962, New York: McGraw Hill.

[6] Richtmeyer, R.D. and Morton, K.W., Difference Methods for Initial-Value Problems, New York: John Wiley pp.136-143.

[7] Hildebrand, F.B., Introduction to Numerical Analysis, 1956, 1974 New York: McGraw Hill.

[8] Roberts, L. G., "Machine Perception of Three-Dimensional Solids," in Optical and Electro-Optical Information Processing, pp. 159-197, J. T. Tippett, et al., (Ed.'s), MIT Press, Cambridge, Mass., 1965.

[9] Hueckel, M.H., "An Operator which Locates Edges in Digitized Pictures" in Journal of the Association of Computing Machinery, Vol.18,No. 1, January 1971, pp. 113-125.

****
****************************** end of correspondence **********************
****