Correlation of gradients and the algorithm susan f^r motion estimation and tracking

R. Aburto G.[1,2], A.C. González G.[2,3], J.H. Sossa A.[3]

[1]Centro de Estudios Tecnológicos Industrial y de Servicios # 34, Cd. Lázaro Cárdenas, Michoacán México.
raburto@sct.gob.mx
[2]Instituto Tecnológico de Toluca, Depto. de Ingría. Eléctrica y Electrónica
Av. Instituto Tecnológico S/N Ex Rancho La Virgen, Metepec México, C.P. 52140
alaing@sagitario.cic.ipn.mx alaing@ittoluca.edu.mx
[3]Instituto Politécnico Nacional, Centro de Investigación en Computación
Av. Juan de Dios Batíz esq. Miguel Othón de Mendizábal, C.P. 07000 México, D.F.
hsossa@cic.ipn.mx

Abstract. The motion estimation problem is important when the basic objective is to determine the spatial localization of mobile objects inside images, the motion detection is basically based on brightness changes in an images sequences. In this work present a correlation of space-temporal gradient and the algorithm SUSAN for determine the motion estimation and the direction.

1. Introduction

The motion analysis in an image sequences is growing in many application. Some of them are: Mobile robotics, Satellite image application, Objects tracking, Autonomous navigation, Medical image processing, Virtual reality and Surveillance applications.

Tracking the objects in a images sequences is a common procedure, when we wants to know if an objects inside of the image sequences is moving and the direction they follow; if there are moving some objects, there will be changes in the image pixels brightness.

2. Motion estimation

There are many techniques available for motion estimation, the main groups are: Gradients, Correlation, and Frequency based methods.

2.1 Gradients

Motion estimation based on gradient is the technique most using actually, this is due to a good estimation of the motion field. Methods based on gradients usually presents some restriction, normally before start the processing they require filtering the images. One of the most gradients based method is Horn Shunk [1], this methods applied a global smoothness to the motion field, while the Lucas and Kanade [2] applied a local smoothness to the motion field, while the common disadvantage of both gradients methods is the change in brightness.

2.2 Correlation

One way to avoid limitations of the gradients based methods is to consider image regions instead of motion estimation pixel by pixel. In general, correlation based methods are less sensible to noise; due to they take more information from images in motion estimation process.

The techniques based on correlation minimize the difference between two block of pixels located in images I_t and I_{t+1} belong to an images sequence [3]. Basically the correlation based algorithm divide the image in series of block with same dimensions. For every blocks in the image I_t, applied a search in the image I_{t+1}, looking for a correlation that minimize a match function.

If block size is fixed constant, the motion estimation field is limited due to be not possible to search for blocks that contents various movements. Some researches still working for solving this problem in multiresolution block correlation [11]. Others are using the Hough Transformer for motion estimation [4][5].

One disadvantage that shares the correlations methods is that they depend to the translation model. The blocks must be small enough in order to consider the translational model valid.

2.3 Frecuency domain

One of the must used methods in the frequency domain are the methods based on the phase. This takes de advantage that a change in the spatial domain produces a change in the phase in the frequency domain [6].

Techniques based in the phases are in use in order to estimate the motion [7]; in this methods a gradients based technique is apply to the phase component of the filters output, sintonized to different velocities.

There is also a technique in the frequency domain based in the correlation [8].
The techniques based in the frequency domain have some advantages over methods in space domain, due to they are less sensible to the global changes in brightness and they have good response against the noise.

3. SUSAN

The algorithm SUSAN is a filter that can be used as corner and edge detector of the objects contain in a digital images, it was development by Smith in 1997 [12]. The acronym of SUSAN is "Smallest Univalue Segment Assimilating Nucleus". The algorithm SUSAN can be used as a filter to reduce the noise in the images, one important application of this algorithm is in searching of corners and edges of the objects contain in the images.

The SUSAN principle is based on the circular mask, having a center pixel, which shall be known as the nucleus. The brightness of each pixel within a mask is compared with the brightness of that mask's nucleus.

This area of the mask shall be known as the USAN (Univalue Segment Assimilating Nucleus), this concept of each image point having associated with it a local area of similar brightness is the basis for the SUSAN principle. The local area or USAN contains much information about the structure of the image. It is effectively region finding on a small scale. From the size, centroid and second moments of the USAN two dimensional features and edges can be detected.

The area of an USAN conveys the most important information about the structure of the image in the region around any point in question. The USAN area is at a maximum when the nucleus lies in a flat region of the image surface, its falls to half of this maximum very near a straight edge, and falls even further when inside a corner. It is this property of USAN's area, which is used as the main determinant of the presence of edges and corners in two-dimensional features.

This algorithm use two thresholds g and t, where g is call geometric threshold and t is the contrast threshold, the geometric threshold clearly affects the quality of the corners detected, and t affects the numbers of corners found.
Due to the fact that SUSAN does not use image derivatives in the search of corners and edges it has good results in presence of noise.

SUSAN analyses different region separately and made local measurement finding the place where the boundaries of two regions has the function in his minimum value, and found the corner between this two regions. There are not conditions about the image structure for using the algorithm SUSAN

4. Implementation

The system is basically formed for three processing image modules.

The first module is the algorithm in charge of calculating the special-temporal gradients of one image sequence [9], the output of this module is the objects edges that move from image I_t to image I_{t+1}, both belonging to an image sequence.

Second module is the algorithm SUSAN, applied as a corner and edges detector, it find the objects corners and edges contained in image I_t and image I_{t+1} and determine the point that change the position between both images.

The third module find the correlation between the space-temporal gradients and the corners and-edges found after applied the algorithm SUSAN.

After found the correlation of both algorithms, the global center of gravity is calculate for the space-temporal gradients and also for algorithm SUSAN and a comparison is made in order to determine the direction of the motion estimation, figure 1.

4.1 Detection of moving objects edges based on space-temporal gradients

One-way of detecting the moving objects edges in an images sequence or built the optical flow is trough the space-temporal gradients [3], equation 1.

$$E(x,y,t)=dF(x,y,t)/ds*dF(x,y,t)/dt \quad (1)$$

Where $dF(x,y,t)/ds$ is the spatial gradient and $dF(x,y,t)/dt$ is the temporal gradient and the symbol * is the AND operation between them.

For detecting the edges of the objects that move from image I_t to image I_{t+1}, the gradient based Sobel filter is used, equation 2.

$$\Delta f=[Gx\ Gy]^T \quad (2)$$

The gradient absolute value is calculated with equation 3.

$$|\Delta f|=|Gx|+|Gy| \quad (3)$$

The angle of the gradient vector is calculated with equation 4.

$$A(x,y)=\tan^{-1}(Gx/Gy) \quad (4)$$

The gradient in an image is determined calculating the partial derivative in the x (df/dx) and in the y (df/dy) direction in every pixel position. The form to implement the partials derivatives in digital image is using the mask of the Sobel filter in both directions, figure 2.

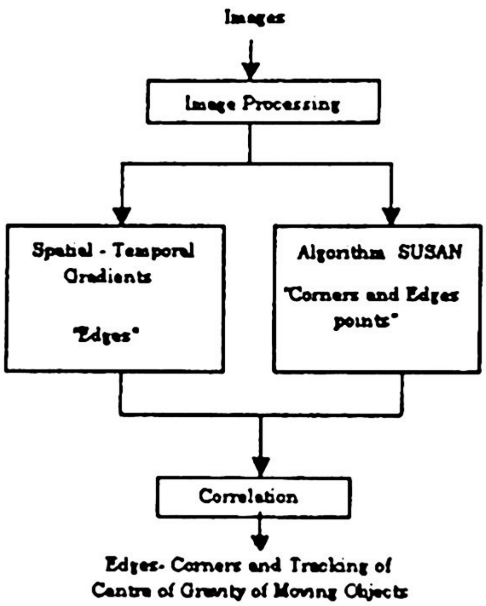

Fig. 1. Algorithm.

Direction x $\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$

Direction y $\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$

Fig. 2. Mask of Sobel filter.

After calculation of gradients magnitude and it's angle, the pixel in the center of the mask takes these values, this process is repeated for all the pixels of the images, and the gradients results are storage in a vector. This procedure is used in all image sequence.

For determining the objects inside the images, that move from image I_t to image I_{t+1} is necessary the following procedure:

- Calculate the difference between the spatial gradients of images I_t and image I_{t+1} in order to obtain the temporal gradient.
- The temporal gradient obtained in the prior step is multiply (AND operation) by the spatial gradient of image I_t
- The values obtained of gradients AND operation should be normalized from 0 to 255, after that a threshold has to be used in order to show only the objects edges that experiment a motion from the image I_t to image I_{t+1}.

4.2 Algorithm SUSAN

The algorithm SUSAN is implemented through the following procedure:

1. - Places a circular mask around the pixel in question.
2. - Using equation 4.5 calculate the number of pixels within the circular mask, which have similar brightness to the nucleus. These pixels define the area USAN.

$$\begin{aligned} C(r,r_o) &= 1 \text{ if } \quad |I(r) - I(r_0)| \le t \\ &= 0 \text{ if } \quad |I(r) - I(r_0)| > t \end{aligned} \tag{5}$$

Equation 5 gives good results, but a much more stable and sensible equation to use for C is equation 6.

$$C(r,r_o)=e^{-((I(r)-I(ro)/t)6} \tag{6}$$

Where:

$I(r)$ is the brightness of the nucleus

$I(r_o)$ is the brightness of any other pixel inside the mask

t determines the minimum contrast of features, which will be detected, and also the maximum amount of noise, which will be ignored.

3. - Obtain the difference between USAN and the geometric threshold.

$$R(r_o) = g - n(r_o) \text{ si } n(r_o) \qquad (7)$$
$$= 0 \text{ otherwise}$$

Where $n(r_o) = \Sigma C(r, r_o)$.
4. - Storage the corners found in a vector.
5. - Test for false corners detection. This can occur with real data where blurring of boundaries between regions occurs. Calculating the center of gravity of the USAN and the distance from the center of gravity to the nucleus can eliminate this problem. Clearly an USAN corresponding to a proper corner will have a center of gravity that is not near the nucleus while a thin line passing through the nucleus will have a short corresponding distance from center of gravity to the nucleus.

4.3. Correlation

If M represents a matching function which return a value proportional to the match of two given features, such as the absolute difference between the two pixels intensity values (E_1 and E_2), then the match strength M(x,y;u,w) for a point (x,y) and displacement (u,w) is calculated by taking the sum of the match values between each pixel in the displaced region in the image I_t and corresponding pixel in the in the actual region of the image I_{t+1}, the match function is calculating using equation 8.

$$M(x,y;u,w) = \sum \Phi(E_1(i,j) - E_2(i+u,j+w)) \qquad (8)$$

The actual motion of the pixel is taken to be that of the particular displacement with the maximum neighborhood match strength, that is equivalent to the minimum region difference.

In this particular work, a correlation was applied to the results of space-temporal gradients and the corners found by the algorithm SUSAN. A 5x5 mask center in the position of the corners is displaced for all corner's vector, searching in the same position of the space-temporal gradients image, giving as a results only the corners that match with the edges of the moving objects found by the space-temporal gradients.

5. Results

Figures 3 and 6 show two images belonging to a monocular image sequence, after applied the space-temporal gradients algorithm, the image in figure 4 was obtained. Figure 5 show the image of the corners and point of the edges found by the application of the algorithm SUSAN.

Figure 6 show the image that combine both algorithms and figure 7 and 8 shows the trajectory that follows the global center of gravity, calculated for around 20 images belonging to the sequences.

Fig. 3. 1st. Image

Fig. 4. Gradient

Fig. 5. Corners

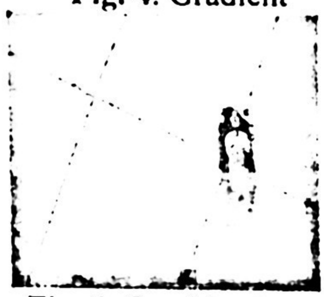

Fig. 6. Combination

Fig. 7. Trajectory

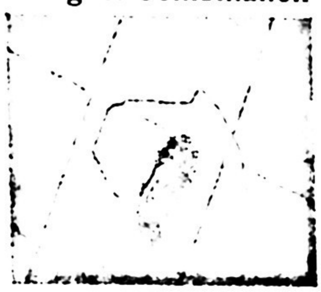

Fig. 8. Trajectory

Both centers of gravity were calculated after application of region correlation to the resulting of both algorithms, and projected over one of the images.
The application of correlation to both algorithms belonging to the same image sequence, reinforce the motion estimation, and the following of the moving objects trajectory is more accurate, during the time interval that the images were taken.}

The application of both algorithms gives good results and can be used in tasks like the surveillance.

References

[1] Horn B. K. P. & Schunck B. G. "Determining optical flow"; Artificial Intelligence, 17, 1981, pp. 185-203.

[2] Lucas B. & Kanade T. " An iterative image registration technique with an application to stereo vision", proceeding DARPA, Image Understanding Workshop, 1981, pp. 121-130.

[3] Jain J. R. & Jain A. K. "Displacement measurement and its application in interframe image coding", IEEE Transactions on Communications, vol. 29, No. 12, 1981, pp. 1799-1808.

[4] Adiv G., "Determining tree-directional motion and structure from optical flow generated by several moving objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 7, no. 4, 1985, pp. 384-401.
[5] Baber & Kitter J. , "Robust motion analysis" Proceeding CVPR'94, Seattle, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994, pp. 947-952.
[6] Popoulis A. Signal Analysis, Mc Graw Hill, 1984.
[7] Fleet D. J. & Jepson A. D. "Computation of components image velocity from local fase information", International Journal of Computer Vision, 5, 1990, pp. 77-104.
[8] Calway A. D., Knutsson H. Wilson R., "Multiresolution estimation of 2-D disparity using a frequency domain approach", Proceeding of British Machine Vision Conference, 1992.
[9] Jain, Ramesh, Rangachar Kasturi y Brian G. Schunck, Machine Vision, Singapore; Mc Graw-Hill Book Co., 1995.
[10] Kearney J. K. , Thompson W. B. & Boley D. L., " Optical flow estimation: an error analysis of gradient-based methods with local optimization", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 9, No. 2, 1987, pp. 229-244.
[11] Burt P. J. & Adelson E. H., "The Laplacian pyramid as a compact image code", IEEE Transaction on Communication, vol. 31 no. 4, 1983.
[12] Smith, Stephen M. "SUSAN", Low Level Image Processing, 1997.
[13] H. Liu, T. Hong, M. Herman, R. Chellapa, "A general Motion Model and Spatio-Temporal Filters for Computing Optical Flow". International Journal of Computer Vision, 1996.
[14] G. D. Hager and P. N. Belhumeur. Real-Time Tracking of Image Regions with Changes in Geometry and Illumination. IEEE Proc.. of CVPR, 1996.
[15] J. M. Letang, V. Rebuffel & P. Bouthemy, "Motion detection robust to perturbations" A statistical regularization and temporal integration framework". In Proc. ICCV, 21-23, 1993.
[16] L. Wixson, "Illumination assessment for vision-based traffic monitoring," In Proc. ICPR, 57-60, 1996.
[17] F. Bremond & M. Thonnat, "Tracking multiple non-rigid objects in a cluttered scene," In Proc. SCIA, 1997.
[18] S. P. Liou & R. C. Jain, "Motion detection in spatio-temporal space," CVGIP 45: 227-250, 1989.

Control y Tiempo Real