Optical Flow Estimation Using Line Image Sequences

Philippe Thévenaz & Heinz Hügli
Institut de microtechnique de l'Université de Neuchâtel
2 rue A.-L. Breguet, CH-2000 Neuchâtel (Switzerland)

Abstract

This paper presents and evaluates a method for measuring the speed of objects in line image sequences. In a line sequence, a line corresponds to a fixed line position in the real scene, and the objects move against it. The line image sequence is a space-time two dimensional image giving a good record of moving objects. The method uses two such line image sequences and estimate the object speed by optical flow computation. Unidirectional movement of the objects is assumed which simplifies the optical flow computation and makes it a simple method to implement. The usefullness and performance of the method is shown by an example comprising several vehicles of different speed. The performance evaluation shows good linearity and low error.

1. Introduction

1.1. PREVIOUS WORK ON OPTICAL FLOW

Optical flow estimation by intensity gradient based techniques has been investigated by many workers [5], [6], [7]. They mainly show how to compute a multidimensional flow field, that not only gives rise to the velocity magnitude, but also to its direction. Unfortunately, this cannot be done without assuming some a priori knowledge [10], [11], [13], [14], because of the aperture problem, that basically specifies that motion can locally only be measured along the intensity gradient.

1.2 LINE IMAGE SEQUENCE

The scope of this paper covers the simpliest of these assumptions, namely the assumption that the motion is along a straight line of known direction. According to that, we use a line image camera that produces a space-time two-dimensional image [1]. Such a camera shoots only one line image at a time, and always aims at the same place; a time sequence is collected and the line images are put aside. The two-dimensional image produced is spatially defined along one axis, and temporally along the other. Such an image is quiet different from an ordinary one. When nothing moves, the background only is shot, and no significant feature is present in the image. When an object crosses the line, then it leaves a trail on the record; if it crosses the line slowly, then it will have a longer trace than when it goes faster. Using such a camera, we will be able to survey every moving object crossing the line of acquisition on a path perpendicular to it. This sort of camera is also used for order decision, e. g. in horse races or bicycle races.

1.3. METHODS FOR MEASURING SPEED

We have seen that a moving object has a longer trace when it goes slow than when it goes fast.

Knowing a priori its real length, and measuring the length of its trace (that is, its duration), we could determine its speed; but the real length of an object is usually not available, so we have to resort to other schemes. One of them is obvious: it is to automatically identify the object and then fetch its length from a table. The other schemes take advantage of a second line image camera, aiming at a place just in front of the first one (figure 1). By knowing the space interval Δx between the lines, and by timing the appearance of locally identifiable features, it is possible to extract speed measurements valid for these features. This is termed a correspondence scheme [8]. The next method used for measuring velocity is the Fourier Method [2], [4], [12]. In this case, we assume that only one object is moving. That assumption lets us predict that one image is the shifted version of the other. The amount of shifting is proportional to the inverse of the speed and can be measured by calculating the phase difference between the Fourier Transforms of each image.

1.4. OPTICAL FLOW

The last scheme is also termed the method of differentials [3], [9]. Its interest resides in the fact that it can compute locally the optical flow, without any a priori knowledge. We will develop this method below, adapting it to a pair of line image sequences.

2. Optical flow of line image sequences

2.1. DEFINITION

The optical flow is a method for computing a velocity field from a set of two or more images. In our case, the field obtained is scalar. It represents the value $V_x(Y, T)$ of the speed measured on every point of the line image sequences $(i_1(Y, T), i_2(Y, T))$, where Y relates to the spatial axis and T to the time axis.

2.2 COMPUTATION

Let us consider a set of general time-varying images:

$$i = f(x, y, t) \tag{1}$$

Developing them in a Taylor series expansion we get:

$$f(x+\Delta x, y+\Delta y, t+\Delta t) = f(x, y, t) + \Delta x \cdot (\partial f/\partial x) + \Delta y \cdot (\partial f/\partial y) + \Delta t \cdot (\partial f/\partial t) + \varepsilon_1 \tag{2}$$

In that expression, ε_1 represents the higher order terms, that we will consider as negligible. If anything moves in the image, then it will be possible to track it and find many sets $(\Delta x, \Delta y, \Delta t)$ such that:

$$f(x+\Delta x, y+\Delta y, t+\Delta t) = f(x, y, t) \tag{3}$$

Whenever such a set exists, we can write:

$$\Delta x \cdot (\partial f/\partial x) + \Delta y \cdot (\partial f/\partial y) = - \Delta t \cdot (\partial f/\partial t) \tag{4}$$

In our case, we know that the motion is perpendicular to the line of acquisition:

$$\Delta y = 0 \tag{5}$$

If the set is not trivial, that is if $\Delta t \neq 0$, then we have:

$$(\Delta x/\Delta t) \cdot (\partial f/\partial x) = - (\partial f/\partial t) \tag{6}$$

From this equation, we can extract the velocity V_x, with no further reference to the particular set we used:

$$V_x = \Delta x/\Delta t = - (\partial f/\partial t)/(\partial f/\partial x) \tag{7}$$

To do that, the necessary conditions are i) the time and space derivatives should be defined, and ii) the space derivative in the motion direction has to be non-zero. This last condition means that the moving object shouldn't be of constant intensity.

Let us consider now a set of two discrete line image sequences taken at locations X_1 and X_2:

$$i_1 = f(X_1, Y, T) = f_1(Y, T) \tag{8}$$
$$i_2 = f(X_2, Y, T) = f_2(Y, T) \tag{9}$$

Choose some points of these images, remembering that the y-coordinate has no effect on the velocity:

$$A = f(X_1, Y, T_1) \tag{10}$$
$$B = f(X_1, Y, T_2) \tag{11}$$
$$C = f(X_2, Y, T_1) \tag{12}$$
$$D = f(X_2, Y, T_2) \tag{13}$$

It can easily be shown that the velocity (7) can be approximated by:

$$v_x(A, B, C, D) = - (\Delta X/\Delta T) \cdot ((B+ D - A - C)/(C + D - A - B)) \tag{14}$$

Where $\Delta X = X_2 - X_1$ is the distance betwen the two acquisition lines, and $\Delta T = T_2 - T_1$ is the time interval between two shots. Note that the condition i) has disappeared, whilst condition ii) is still valid.

2.3. CONFIDENCE

Let us now have a look at the sensitivity of that method. What we want to know is the consequence of an error on the gray levels specified in (10) through (13). For that purpose, we will use again a Taylor series expansion, where the higher than first order terms ε_2 are dropped:

$$v_x(A+\Delta A, B+\Delta B, C+\Delta C, D+\Delta D) = v_x(A, B, C, D) + \Delta A \cdot (\partial v_x/\partial a) + \Delta B \cdot (\partial v_x/\partial b) + \Delta C \cdot (\partial v_x/\partial c) + \Delta D \cdot (\partial v_x/\partial d) + \varepsilon_2 \tag{15}$$

We term:

$$\Delta v_X = v_X(A+\Delta A, B+\Delta B, C+\Delta C, D+\Delta D) - v_X(A, B, C, D) \quad (16)$$

That gives:

$$\Delta v_X = -2 \cdot (\Delta X/\Delta T) \cdot (((B - C) \cdot (\Delta A - \Delta D) + (D - A) \cdot (\Delta B - \Delta C))/(C + D - A - B)^2) \quad (17)$$

Considering that the errors are of same magnitude:

$$|\Delta A| = |\Delta B| = |\Delta C| = |\Delta D| = \Delta E \quad (18)$$

We have in the worst case approximation:

$$|\Delta v_X| = 4 \cdot |\Delta X/\Delta T| \cdot ((|B - C| + |D - A|)/|(B - C)^2 - (D - A)^2|) \cdot \Delta E \quad (19)$$

Finally, we define the sensitivity S, and the relative sensitivity S_r by:

$$S = |\Delta v_X|/\Delta E \quad (20)$$
$$S_r = S/|v_X| = 4 \cdot ((|B - C| + |D - A|)/|(B - C)^2 - (D - A)^2|) \quad (21)$$

2.4. THRESHOLDING

It is clear that when S_r is high, a low confidence has to be set on $V_X(Y, T)$, since any small error (for example: quantization) has a great effect on the measured velocity. Indeed, we could introduce a measure for that confidence, that would be like $1/S_r$; but this step is not necessary. Rather, we choose a threshold S_{r0}, above which we no longer accept $V_X(Y, T)$ as reliable. Finally, we have the optical flow field:

$$V_X(Y, T) = \begin{cases} v_X(A, B, C, D) & \text{if } S_r(A, B, C, D) < S_{r0} \\ \text{undefined} & \text{else} \end{cases} \quad (22)$$

3. Experiments and results

3.1. EXPERIMENT

To verifiy the usefullness and performance of the method, we constructed a bench on which some train models could be run. A single motor was used, but three different pulleys allowed three different speeds at a time. The transmission was designed to allow two pulling directions for each wire. We used a CCD video camera to look at the scene. From each frame, we retained two adjacent columns of pixels to simulate the two line image cameras. The acquisition rate was 12.5 [Hz].

3.2. RESULT: EXAMPLE

Figure 2 shows one of the two line image sequences obtained (the other is visually indistinguishable from the first one). Three models have a negative speed (first floor), and three others a positive one (second floor). The time axis is horizontal; the past-future direction is from left to right. Its range is about 41 [s]. Notice a

property of line image sequences: although some models were moving to the left and some to the right in the real scene, the line image sequence shows only left-going models. The reason is simply that they move head first and tail last. Superimposed over the image are frames enclosing each model. These frames are the result of some segmentation process we don't describe here.

Figure 3 shows the full optical flow field associated with figure 2. One can see that much noise is present, but at least the shape and direction (light: positive, dark: negative) of each model can be recognized. Nevertheless, the high level of noise forbids us from simply thresholding $V_x(Y, T)$ in order to segment the image.

Figure 4 shows the relative sensitivity field; the less sensitive measurements are coded more darkly. In this image, it is not difficult to find a threshold S_{r0} that conserves only points where confidence is sufficient. We see that most such points are located where the intensity gradients are large, that is, where simultaneously the denominator and numerator of the last term of expression (7) are large.

Figure 5 shows the result of relative sensitivity thresholding. White encodes accurate measurements. Grey encodes points above the S_{r0} threshold, for which the confidence is insufficient. Black stands for undefined velocity as well as for undefined relative sensitivity. The threshold used was S_{r0} = 100%; it means that the correct velocity could have had twice the measured value, or as well have been zero.

Figure 6 shows the figure 3, masked with figure 5. A lot of noise has disappeared, and velocities are now contrasted enough to recognize the velocity gradations between models.

The next step is to investigate the correctness of these measurements. For this purpose, we will compute the mean and standard deviation of measured velocities for each model. Figure 7 shows the computations performed within each model frame, before and after thresholding. Note that the frame shape is well suited to all models but the motorcyclist, for which the zero-velocity points have a priori some influence. The last column of figure 7 shows the real speed. All units are in $[\Delta X \cdot s^{-1}]$, that is the number of times the distance between the two acquisition lines is covered during one second.

Looking at these results, we can see that a sharp improvement over the measurement accuracy has been obtained. The correspondence between the real speed v_r and the measured speed is good. The thresholding operation has been able to reduce the standard deviation by a factor of 3.

3.3. RESULT: LINEARITY

The last step is to evaluate the accuracy under very good conditions. We resort to a special moving object, depicted in figure 8.This particular pattern is a space-varying sinusoidal pattern, that was mounted on the tender models. Such a pattern is well suited to optical flow measurements, because its intensity gradient is non-zero almost everywhere. We made a lot of measurements, varying the motor speed between each one.

The data obtained are plotted against the real velocity in figure 9. We see that the linearity of the measurements is excellent, over more than one decade range. The mean square best fit line has a slope m = 1.0453. The difference from the true unity can be explained by some errors made during the calibration

process of real speed measurements. The offset of this line is 0.0603 [$\Delta X \cdot s^{-1}$], which is quiet negligible. The normalized correlation coefficient is $\rho_{xy} = 0.9757$, which ascertains that the results show good linearity.

4. Conclusion

We have developed a method based on optical flow for measuring velocity with a set of two line sequence cameras, assuming unidirectional motion. The characteristic of this method is that the speed measurement is based on the image gradient, and its advantage is that of a local measurement which does not require knowledge about the shape of the objects to measure. This method has proven to be usefull in most cases where the moving objects offer sufficiently smooth and large intensity variations. In practice and under very good conditions, the linearity is excellent over more than one decade range, while the error is quiet low. In common practical situations, the measurement is noisy but can be significantly reduced by confidence thresholding. Typically, the standard deviation is reduced by a factor of three.

5. Acknoledgements

This work was supported by the swiss *Commission pour l'Encouragement de la Recherche Scientifique* under contract number 1511 and C^{ie} des Montres Longines SA.

6. References

[1] Aoki M., "Detection of Moving Objects Using Line Image Sequence", *Proc. Seventh Int. Conf. Pattern Recognition, pp.784-786,* 1984.

[2] Arking A. A., Lo R. C., Rosenfeld A., "An Evaluation of Fourier Transform Techniques for Cloud Motion Estimation", *TR-351 (January 1975) Computer Science Departement, Univeristy of Maryland, College Park/MD.*

[3] Cafforio C., Rocca F., "Methods for measuring small displacements of television images", *IEEE Trans. IT-22, pp. 573,* 1976.

[4] Haskell B. G., "Frame-to-Frame Coding of Television Pictures Using Two-Dimensional Fourier Transforms", *IEEE Trans. Information Theory IT-20, pp. 119-120,* 1974.

[5] Horn B. K. P., Schunk B. G., "Determining Optical Flow", *Artificial Intelligence, Nº 17, pp. 185-203,* 1981.

[6] Kahn P., "Local Determination of a Moving Contrast Edge", *IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. PAMI-7, Nº 4, pp. 402-409,* July 1985.

[7] Kearney J. K., Thompson W. B., Boley L. D., "Optical Flow Estimation: An Error Analysis of Gradient-Based Methods with Local Optimisation", *IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. PAMI-9, Nº 2, pp. 229-244,* March 1987.

[8] Lawton D. T., "Processing Translational Motion Sequences", *Computer Vision, Graphics and Image Processing, Nº 22, pp. 116-144,* 1983.

[9] Limb J., Murphy J., "Estimating the Velocity of Moving Images in TV Signal", *Comp. Graph. Image Proc., Nº 4, pp. 311-327,* 1975.

[10] Murray D. W., Buxton B. F., "Scene Segmentation from Visual Motion Using Global Optimisation", *IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. PAMI-9, Nº 2, pp. 220-228,* March 1987.

[11] Prager J. M., Arbib M. A., "Computing the Optic Flow: The MATCH Algorithm and Prediction", *Computer Vision, Graphics and Image Processing, Nº 24, pp. 271-304,* 1983.

[12] Roese J. A., Pratt W. V., "Combined Spatial and Temporal Coding of Digital Image Sequences", *Proc. SPIE 66, pp. 172-180,* 1975.

[13] Thompson W. B., Barnard S. T., "Lower-Level Estimation and Interpretation of Visual Motion", *Computer,* August 1981.

[14] Ullman S., "Analysis of Visual Motion by Biological and Computer Systems", *Computer,* August 1981.

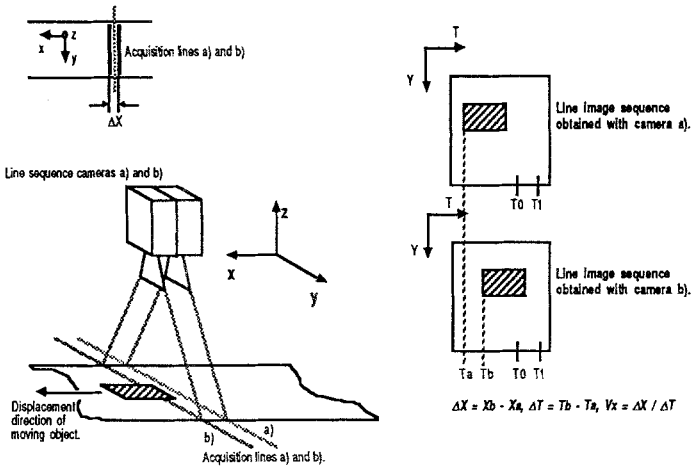

Figure 1: A set of two line cameras is used to implement speed measurements.

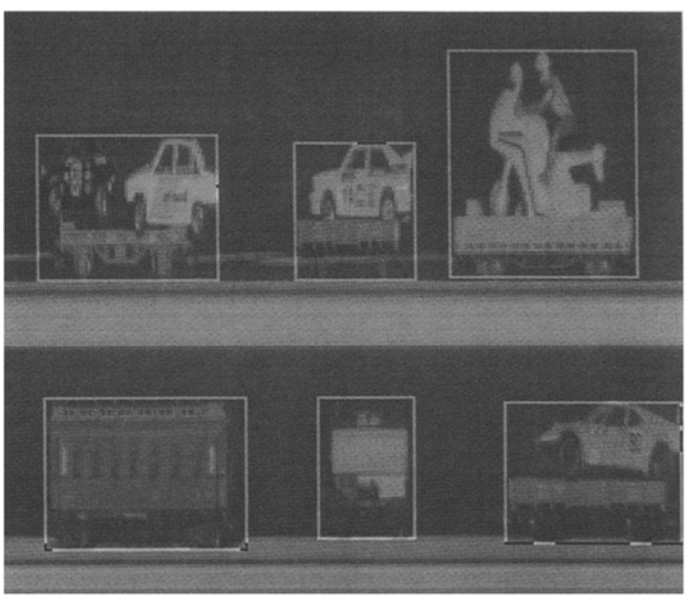

Figure 2: Line image sequence $i_1(Y, T)$, showing six models with a superimposed framing.

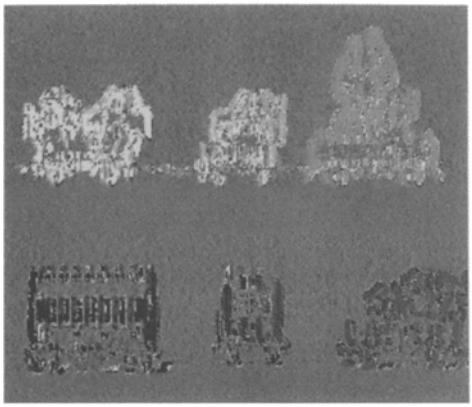

Figure 3: Optical flow field over the full image. Positive velocities are light and negative are dark. The neutral grey level codes zero-velocities as well as undefined velocities.

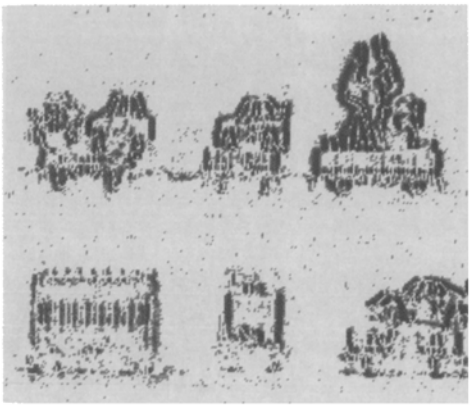

Figure 4: Relative sensitivity field. A high relative sensitivity (low confidence) is lighter than a low one (high confidence). Undefined relative sensitivities are mapped to white (lowest confidence).

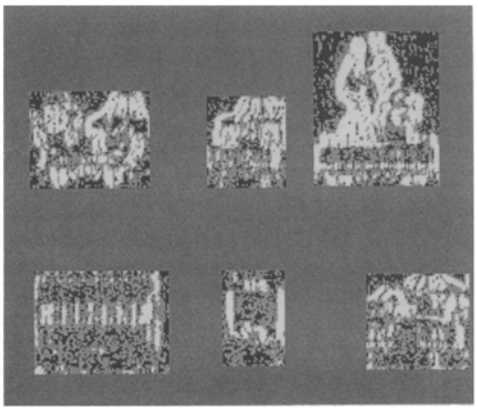

Figure 5: Result of thresholding. The white pixels encode the retained velocity measures.

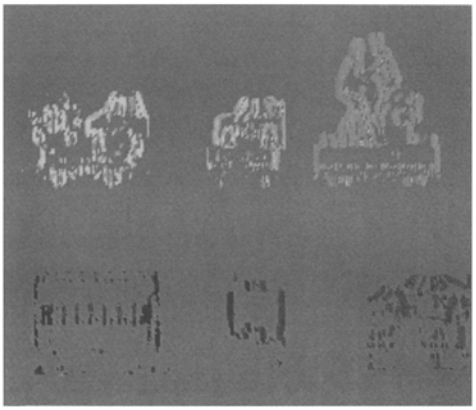

Figure 6: Figure 3 masked with figure 5.

Object	v	σ	v_t	σ_t	v_r
Two cars	27.29	29.30	31.90	11.07	32.72
Upper car	21.29	19.34	25.18	6.71	27.03
Motorcyclist	14.00	19.34	16.97	3.59	16.72
Coach	-19.75	28.29	-29.66	11.66	-32.72
Tank	-14.76	17.40	-23.08	6.14	-27.03
Lower car	-13.60	11.71	-16.63	4.20	-16.72

Figure 7: Comparison of unthresholded average velocity v, thresholded velocity v_t and real velocity v_r. The corresponding standard deviations σ and σ_t are also given.

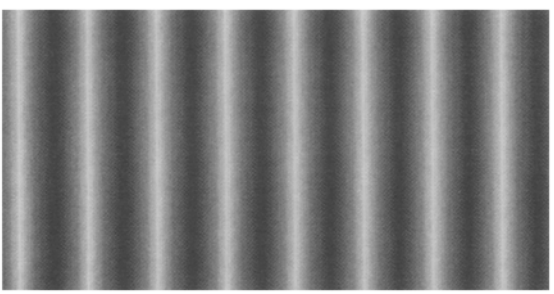

Figure 8: Special pattern used to establish figure 9. The direction of movement was perpendicular to the sinus crests.

FIGURE 9: OPTICAL FLOW LINEARITY