MACHINE VISION IN THE GRAPHIC ARTS: DIGITAL REGISTRATION OF FOUR COLOR HALFTONE SEPARATIONS

By Monti R. Wilson *

Abstract: Opti-Copy, Inc. has recently developed the RegiStar™ digital registration machine, a machine that quickly and accurately pin registers four color halftone separations through the use of digital imaging coupled with high speed computation. In the machine, the computed translational misregistration at two points on the separations is the input to the machine movement computations that accomplish film registration.

The heart of the digital registration machine is the computation of the translational misregistration between the operator's first reference film and a given film to be registered to the reference film, at each of the two points that locate halftone picture detail within a small area. This paper describes the image processing algorithms required to compute the translational movement of one small halftone picture into registration with the corresponding reference halftone picture. Basically, the highly resolved halftone pictures are converted to interpolatively resampled continuous tone pictures that are subsequently enhanced and then registered by a hierarchical resolution correlation method.

Background

One of the most ubiquitous prepress operations is the manual registration of four color halftone separations. Manual registration, however, is time consuming, requires intensive

---

* OPTI-COPY, INC., 1986

skilled labor and is difficult to perform repetitively with uniform high accuracy. By automating manual registration with machine vision technology, productivity and quality gains in printing can be realized. A registration machine is faster than manual registration, requires greatly reduced labor and skill and provides uniform, repetitive high registration accuracy essentially independent of human performance.

In recent years, several registration machines have been introduced. These machines utilize only special fiducial or register marks to achieve alignment. Frequently, register marks can not be standardized, are incorrectly placed or are missing entirely, leaving only halftone picture detail to be registered. The recently developed Opti-Copy RegiStar™ digital registration system not only can use register marks, but also registers solely from most halftone picture detail.

THE REGISTAR™ REGISTRATION PROCESS

Figure 1 is a photograph of a RegiStar™ beta site test unit. To the left is a registration picture point selection station and to the right is the film punching station. Prior to registration, the operator selects one color as the reference color and strips the separation onto an unpunched carrier sheet. The remaining colors are then each stripped onto separate carrier sheets in rough register with the reference color. The rough registration need only be accurate to + or - 5 rows of halftone dots, a requirement that is quickly met by even inexperienced operators.

To initiate the registration process, the operator places the reference carrier in the picture point selection station and selects two points, on the carrier mounted separation, with the electronic cursor. The two points need to be reasonably well separated; each point locates the center of a small area (about .3 inch on a side) of sufficient halftone picture detail (an area devoid of detail would offer no possibility of registration). The reference carrier is then placed in the film punching station. The machine positions the digital camera at each of the two selected points and captures digital pictures. The pictures are processed by digital computer algorithms, reduced to digital pictures suitable for registration and stored; the reference carrier is then punched.

The other color separations are then sequentially fed to the machine which, for each separation, again captures digital pictures at the selected points, performs digital computations and punches each carrier in register with the punched reference carrier. In the registration process for each non-reference separation, a reduced digital picture was created at each of two points. At each point, the two reduced pictures (one stored reference and one non-reference) are digitally correlated to determine a registration move at that point. The registration move is the xy translation of the small non-reference picture into register with the small reference picture at the given selected point. With a registration move computed for both of the two selected points, a final machine motion registration algorithm computes the actual mechanical movements required to position the non-reference carriers prior to punching.

Scope

The key data in the RegiStar™ registration process are the registration moves computed for two small halftone pictures at each of two selected points. In the case of register marks, one or both of the two moves could be computed for two corresponding register marks. Register marks typically require an algorithm for the registration move that is radically different from the algorithmic process in the case of small halftone pictures.

The remainder of this paper is devoted to a presentation of the algorithms originally used in the RegiStar™ to predict the registration move for two small halftone pictures. The paper concludes with some comments.

Overview of the Algorithms for Small Halftone Picture Registration

At the microscopic level, halftone images are essentially binary or two-level pictures (black and white). If the halftone dots overlayed each other in registration, one could easily register halftones by first thresholding the gray level pictures into binary pictures and using classical correlation (similar to the correlation registration described later) to predict the registration move between pictures. Unfortunately, due to the varying halftone screen angles and possibly interdot spacings between colors, the halftone dots of the combined colors form a rosette pattern, rendering direct correlation registration inaccurate at best.

The human stripper registers the apparent continuous tone picture detail of a halftone image first coarsely, then to fractional dot accuracy with the aid of magnification. The RegiStar™ imitates this process by first capturing high resolution digital pictures with about 13 (1/2 mil) pixels spanning a full size 150 line screen halftone dot in one direction (13 x 13 cell resolution). The high resolution corresponds to the stripper's use of magnification. The further correspondence with manual registration is that continuous tone pictures are constructed from the raw data with digital picture tones based on the varying size of halftone dots. Interpolation of the constructed pictures provides the fractional dot accuracy inherent in the high resolution pictures. Interpolation is also used in the final registration, as the last stage of a hierarchical process that first locates registration to pixel (1/2 row of dots) accuracy then uses subpixel interpolation to bring out the remaining fractional pixel registration move information. Prior to registration, the registration pictures are enhanced to render the registration algorithm more accurate.

The following sections describe seven steps to a registration move computation. The first five steps are applied to each small halftone picture. The last two steps operate between pairs of pictures to be registered. The seven steps are the following:

1) Optimum thresholding of the raw data picture to a binary picture.
2) Extraction of halftone screen angle and interdot spacing.
3) Construction of a dot centered pixel, continuous tone array picture.
4) Interpolation of the dot centered picture to a reduced continuous tone picture.
5) Edge enhancement of the reduced picture.
6) Tone normalization of the moved picture to the reference picture.
7) Hierarchical correlation registration.

1. Optimum thresholding of the raw data picture to a binary picture

To determine continuous tones from dot areas, the dot areas must be determined from counting black pixels within a halftone cell. To identify black pixels from the white background, the raw data (gray level picture) is thresholded into a binary picture with tone 1 for black and tone 0 for white (this scheme facilitates counting arithmetic). First, the global maximum and minimum gray levels are found and the average yields the initial threshold. Next, the initial threshold divides pixels into black (below the threshold) and white (above the threshold). The average black and average white gray level is computed for each category. The optimum threshold [1] is (intuitively) the average of the two average gray levels. A final thresholding is performed using the optimum threshold.

2. Extraction of halftone screen angle and interdot spacing

To compute dot centered continuous tones, the regular array of black halftone dot centers is calculable from knowledge of the interdot spacing and the halftone screen angle. The halftone screen angle here is only the angle between horizontal and the first set of parallel lines of dot centers encountered. A simple pattern recognition approach yields the desired results.

First one finds the centers of about 40 dots centrally located in the thresholded raw data picture. These so called data dot locations are the data in a least squares fit to the best screen angle. The interdot spacing appears as a by-product of the screen angle computation.

The data dots are found by scanning rasterwise (Figure 2) to locate runs of pixels that intersect dots idealized here as circular. The runs either start white then run black then terminate white (black dots on a white background) or start black then run white then terminate black (white dots on a black background). Both types of runs are sought and either 40 black dot centers are data or 40 white dot centers are data depending on the picture. Ultimately, black dot centers are predicted, but the equations based on white dot data that predict the centers need only be offset by 1/2 of a cell spacing to yield black dot centers.

The run centers yield the initial data dot centers. These results are refined by chord bisection (Figure 3) to obtain the final data dot centers. The dot center algorithms perform adequately even for non-circular dots.

The least squares formula for the screen angle θ is

$$\theta = \tan^{-1} \left[\frac{-\sum_{i=1}^{N} y_i'(x_i-h) + \sum_{i=1}^{N} x_i'(y_i-k)}{\sum_{i=1}^{N} x_i'(x_i-h) + \sum_{i=1}^{N} y_i'(y_i-k)} \right], \qquad (1)$$

where h and k locate the origin dot in the x_i, y_i frame. This equation involves pairs of coordinates for data dot centers, namely x_i', y_i' locating the i th center in angled or prime frame of the dot centers and x_i, y_i locating the i th center in the frame fixed in the picture. The data dot centers found were x_i, y_i. To determine x_i', y_i', one performs an incremental stepping in a square spiral search pattern about one data dot chosen as the origin dot to a known x_i', y_i' location. This prime frame location is then transformed to the picture frame and the resulting x_i, y_i checked against the data collection to determine if a data dot is indeed at that location. The search begins with an estimate for the screen angle and is refined from (1) self-consistently as the solution proceeds by accumulating more and more confirmed coordinate pairs. The by-product mean interdot spacing d is

$$d = N^{-1} \sum_{i=1}^{N} \left[\frac{(x_i-h)^2 + (y_i-k)^2}{x'^2_{iin} + Y'^2_{iin}} \right]^{\frac{1}{2}}, \qquad (2)$$

where x'_{iin} and y'_{iin} are integer (spacing divided out) prime frame coordinates.

Tests have shown that the screen angle from (1) is accurate to about .1 degree. The interdot spacing is accurate to a few percent or better.

3. Construction of a dot centered pixel, continuous tone array picture

The first constructed picture is the regular array of black dot centers with a continuous tone computed for each centered black dot. The regular array of dot centers in the frame of the thresholded raw data picture is computed from the familiar coordinate transformation equations with input data of regular array prime frame coordinates (specified), screen angle , interdot spacing d and raw data frame location of the data dot selected as origin dot (closest to picture center). At each dot center in the raw data frame, a d by d square cell is erected with center at dot center and oriented according to the screen angle. The black pixels n_b in each cell are counted and the cell fractional dot area % is

$$\% = \frac{n_b}{n_{tot}}, \qquad (3)$$

where n_{tot} is the total pixels in the cell. The optical transmission t is

$$t = 1 - \% \,. \qquad (4)$$

Since transmissions multiply, the optical density of t is taken:

$$O = -\log_{10}(t). \qquad (5)$$

The continuous tone lies in the range 0 (black) to 1 (white) and is defined using (5) by

$$\text{continuous tone} = \frac{O_{max}-O}{O_{max}} \qquad (6)$$

.O_{max} is the cutoff optical density (blackest black that need be considered). Typically, a 95% dot yields a satisfactory cutoff O_{max}.

4. Interpolation of the dot centered picture to a reduced continuous tone picture

The final reduced continuous tone picture is about 80 x 80 pixels and always centered the same way (no relative offsets) square within the continuous tone picture of section 3. The pixels of the reduced picture are always spaced 3.33 mils apart (1/2 row of dots spacing for a 150 line standard) and form a square grid of known locations in the standard xy frame of the section 3. picture. The pixels do not lie on the dot centered continuous tones of the angled (the frame of the screen angled dots) or primed frame of the section 3. picture, hence interpolation is used to compute the reduced picture pixel tones.

The interpolation is 81 point bicubic spline interpolation [2] that is accurate to several parts per thousand. For each pixel in the reduced picture, the pixel center is transformed to the prime frame of the dot centered tones and truncated to integers to locate a central dot centered data tone. The truncation difference yields the interpolation point as an offset from the central data point. Symmetrically about the central data point lie the 9 x 9 grid of dot centered interpolation data points. The dot product of the data array and the interpolation coefficient array yields the interpolated tone. The resulting reduced continuous tone picture is thus an interpolatively resampled, standard registration picture. The spline interpolation yields an accurate yet computationally tractable approximation to bandlimited "exact" reconstruction interpolation of the two dimensional Whittaker-Shannon sampling theorem [3]

5. Edge enhancement of the reduced picture

The reduced continuous tone picture of section 4. is not yet quite suitable for registration. Edge enhancement (and also tone normalization to be discussed) of the reduced picture sharpens the accuracy of the registration. Pratt [4] suggested the convolution of pictures with an edge mask to sharpen correlation registration. A problem is that this technique also heightens local (near edge) discrepancies at the expense of global similarities between two pictures in correlation registration. The RegiStar™ solution to this problem is to retain the original picture as base global structure and to add to it an edge based enhancement, thereby merely highlighting edges in the original picture. The qualitative notion of original picture plus edge enhancement also has some recent physiological support [5] in studies of eye receptor profiles in which a better fit to data is obtained with a Gaussian (original picture) plus the Lapacian of a Gaussian [6] (edge enhancement) rather than with the Laplacian of a Gaussian alone.

The choice of edge operator is facilitated by the fact that first derivatives respond with a rather broad peak near edges while second derivative operators respond with a sharp zero crossing at edges. Better edge localization dictates operators that cross zero at edges. The suggested form of the picture plus enhancement is therefore

$$P(j,i) + 1 - \frac{P_c(j,i)}{P_{cmax}} , \qquad (7)$$

where $P(j,i)$ is the j, i th pixel of the reduced picture and $P_c(j,i)$ refers to the convolution of P with the edge mask of the edge operator. P_c is modified to $P_c - P_{cmin}$ if P_{cmin} is negative, and P_{cmax} is the maximum of the resulting P_c. Near an edge, (7) will result in a large enhancement or highlighting.

Many edge masks are candidates for use in the convolution P_c in (7), such as a 5 x 5 Laplacian of a Gaussian mask. The best choice of edge mask for the RegiStar™ was somewhat novel and was validated empirically.

6. Tone normalization of the moved picture to the reference picture

So far, the reference picture P_{ref} and the picture P_{mov} to be moved into registration with the reference picture have been computed from equation (7). The tones in these two pictures can still differ enough to cause problems in high registration accuracy. Fortunately, a simple linear transformation significantly improves the quality of registration results by making the tones between pictures more similar. The linear transformation is applied only to picture P_{mov} and the form of the transformation is

$$P_{mov\ new} = aP_{mov\ old} + b \ . \tag{8}$$

The coefficients a and b are computed (in two distinct cases) to realize the linear mapping illustrated in Figure 4. The result of the mapping is that the minimum, average and maximum tones of P_{mov} are mapped into the respective minimum, average and maximum tones of P_{ref}. This is a global technique to render the two pictures more similar prior to registration.

7. Hierarchical correlation registration of P_{mov} into P_{ref}

The final step is the computation of the registration move of P_{mov} into register with P_{ref}. In a two-stage process, the registration move is first located to within 1 pixel, then a refinement stage using an interpolatively modified error criterion provides the final result to within a fraction of a pixel.

The first stage registration is accomplished with the correlation error criterion suggested by Barnea and Silverman [7] and later widely adopted by the machine vision community, namely

$$\text{error} = \sum_{i=i_s}^{i_e} \sum_{j=j_s}^{j_e} \left| P_{ref}(j,i) - P_{mov}(j-n_y,\ i-n_x) \right| . \tag{9}$$

For each of all possible trial moves (Figure 5) n_x and n_y, the error is computed from (9). The minimum error yields the best n_x and n_y. In practice, better results are obtained by multiplying the error in (9) by a normalization factor that is dependent on the overlap region of Figure 5, bounded by i_s, i_e, j_s and j_e, similar to a factor previously suggested [7].

In the final, interpolative stage of registration, equation (9) is modified to

$$\text{error} = \sum_{j=j_s+4}^{j_e-4} \sum_{i=i_s+4}^{i_e-4} \left| P_{iref}(j+d_y,\ i+d_x) - P_{mov}(j-n_y, i-n_x) \right| , \tag{10}$$

where d_x and d_y are fractional pixel offsets, P_{iref} is a 81 point bicubic spline interpolation [2] of P_{ref} and i_s, i_e, j_s, j_e, n_x, and n_y are the minimum error values from (9). The final registration move to a fraction of a pixel is $n_x + d_x$ and $n_y + d_y$. The normalization factor on (10) has been suppressed (it is similar to the factor for (9). Equation (10) is exact to within the interpolation error. A straightforward but rigorous proof of (10) is omitted for brevity.

Concluding Remarks

Given suitable halftone picture detail, the final registration accuracy in the finished product typically is tightly clustered about 1 mil error with rare excursions as high as about 2 mil error. A human stripper can match or exceed the machine performance on any one registration by spending sufficient time on it. Day in and day out, however, the machine performance will register significant productivity and quality gains over the human stripper. The machine never tires.

The halftone registration algorithms of section 1 through 7 are not without limitations. A tacit assumption in correlation registration is that the pictures to be registered are basically similar. Crudely, the algorithm expects to match black to black and white to white. Inaccuracy will result if the registration pictures are qualitatively tonally reversed in appearance. The operator must take care not to select such pictures for registration. Research is in progress to overcome the limitations of the small halftone registration algorithm presented in this paper.

The computational requirements of the RegiStar™ are rather severe for a machine vision application. With algorithms executed in software, a computational power of about 20 Megaflops is required. This situation will rapidly improve over the next decade as computer and hardware performance climbs simultaneously with declining cost per performance.

LITERATURE CITED

1. Gonzalez, Rafel C. and Wintz, Paul
1977 Digital Image Processing, Addison Wesley Publishing Co., 1977, p. 327.

2. Powell, M. J. D.
1981 Approximation Theory and Methods, Cambridge University Press, 1981, p. 226 and chap. 19.

3. Pratt, William K.
1978 Digital Image Processing, Wiley-Interscience Publication, 1978, p. 119.

4. Pratt, William K.
1978 Digital Image Processing, Wiley-Interscience Publication, 1978, p. 563.

5. Young, Richard A.
1985 "The gaussian derivative model for machine and biological image processing", Proceedings of Electronic Imaging 85, 1985, pp. 64-70.

6. Marr, D. and Hildreth E.
1980 "Theory of edge detection", Proc. Roy. Soc. Lond. B 207, 1980, pp. 187-217.

7. Barnea, Daniel I. and Silverman, Harvey F.
1972 "A class of algorithms for fast digital image registration", IEEE Trans. on Computers, Vol. C-21, No. 2, 1972, pp. 179-186.

Figure 1. A RegiStar™ beta site unit.

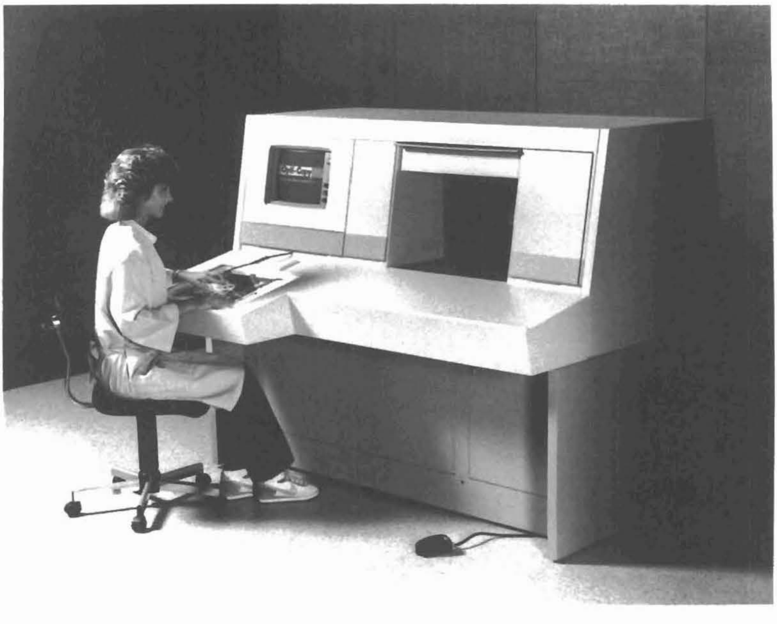

Figure 2. Horizontal runs initially locate data dot centers.

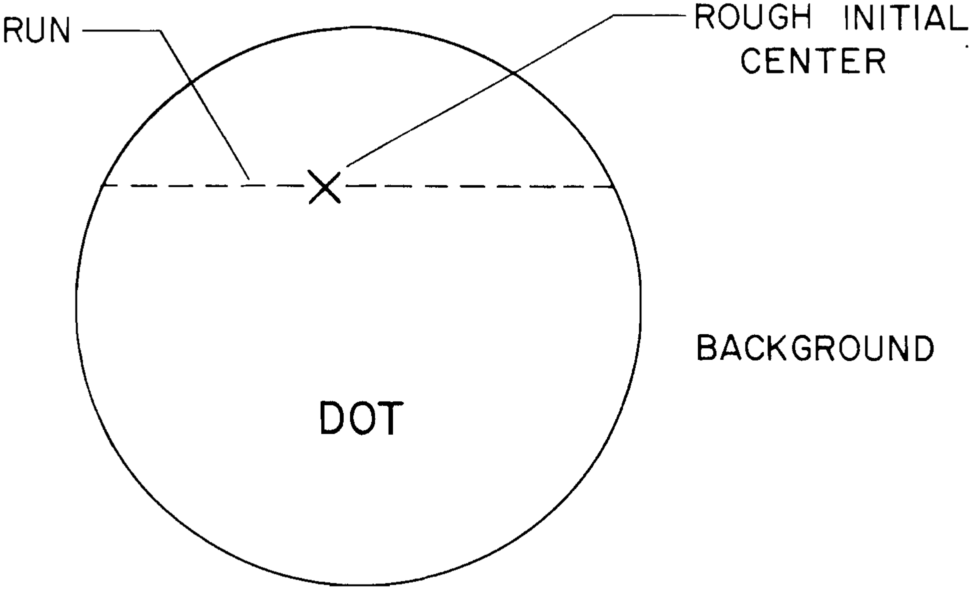

Figure 3. Chord bisection refines data dot centers.

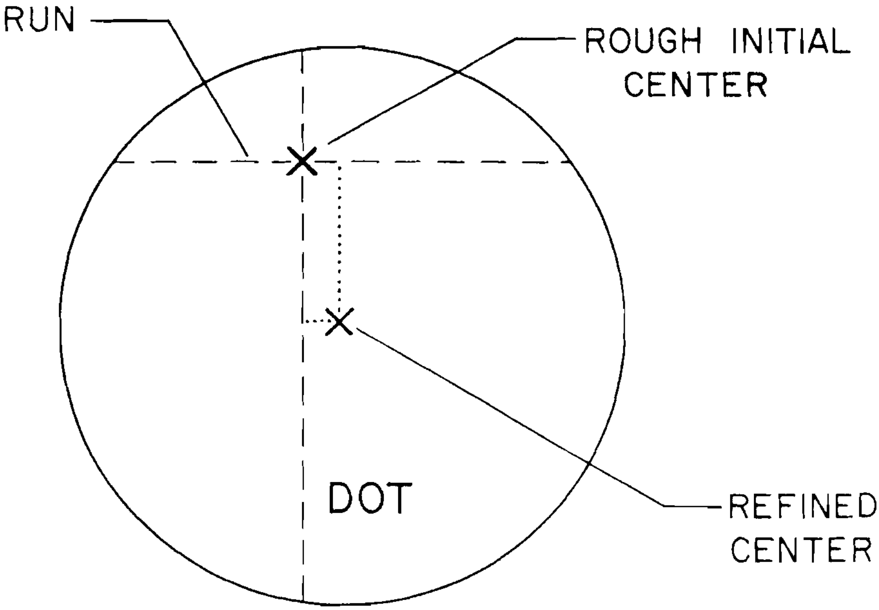

Figure 4. Tone normalization of P_{mov} to P_{ref}.

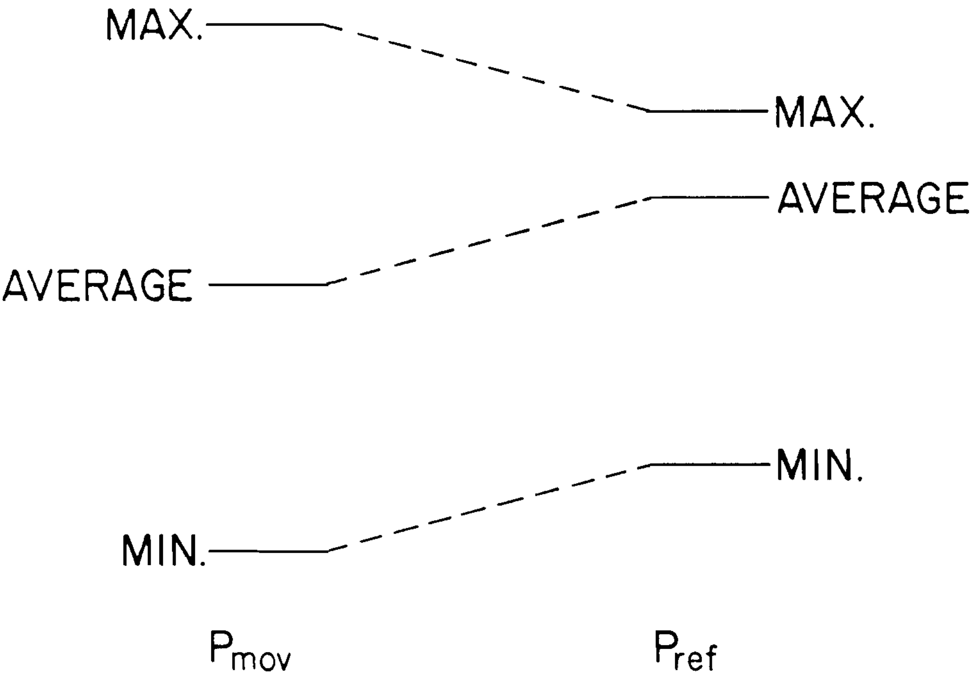

Figure 5. The minimum error trial move n_x, n_y yields registration.