Prof. dr. ing. Radu Dobrescu
Prof. dr. ing. Gabriel Ionescu
Prof. dr. ing. Dan Popescu
POLITEHNICA University, Department of Control & Computers
Bucharest

IMPROVING PATTERN RECOGNITION THROUGH IMAGE MATCHING

Abstract: *The paper presents an imbedded method for recognition of obstacles in a robot scene using the Image Matching Method (IMM). The first section presents the fundamentals of IMM, especially the possibility to match attributes of correspondent points in order to accomplish the correspondence from pixel to pixel to domain to domain between two static images taken from different angles or with different illumination. The second section is dedicated to the determination of interest points, classified on criteria of: discrimination, invariance, stability and uniqueness. Actually, the procedure is limited to find corner points and centers for circular symmetric structures. The implementation was performed by using representations in the Hough space; a complete procedure for robust estimation of the IMM procedure is also discussed. The third section presents the solution used for the environment development and comments some representative simulation results.*

1. INTRODUCTION

In several control applications with external adaptive mobile robots the incomplete definition of the environment creates inconvenience for decision about the avoidance of obstacles. Sensor fusion is the most common method used to improve information on the robot scene [1], but it implies heterogeneous sensors (i.e. CCD camera and sonar) and consequently difficulties in matching data. We propose now a matching method which uses information taken from the same kind of sensors: CCD cameras.

In remote sensing , matching can be defined as the establishment of the correspondence between various data sets; so the matching problem can be referred as a correspondence problem. The data sets can represent images, maps or object models. In our work we consider two matching steps: first, parts of an image are matched with parts of another image in order to define representative object points; then parts of the image are matched with object models in order to identify and localize the depicted scene models. Of course, we have been using only digital image matching, which automatically establishes the correspondence between primitives extracted from two digital images depicting at least partly the same scene. We have considered two situations: a single CCD placed on the mobile robot which offers successive images when the robot is moving to the target or two fixed CCD cameras placed at different angles but focussed on the same window scene.

The paper deals mainly with the algorithms which ensure data processing for image matching, but also some consideration on the development environment are finally proposed.

2. IMAGE MATCHING BACKGROUND

Let be two digital images I' and I". Let be the points P' and P" of the images I' and I" with the coordinates (r',c') and (r",c") and the light intensities g' and g". We assume that if (r',c') and (r",c") are corresponding points, then the coordinates are connected by the relationship:

$$(r',c') = T_g(r'',c'';p_g) \qquad (1)$$

where Tg is a geometric function which reflect the geometrical relations between images and p is a set (vector) of unknown parameters. The intensity level of one image can be deduced from the intensity level of the other image through:

$$g' = T_i(g'';p_i) \qquad (2)$$

where the mapping function T_i offers the intensity level correspondence and the vector p_i is unknown. So, the complete model for Image Matching appears as :

$$g'(r',c') = T_i\left\{g''\left[T_g(r'',c'',p_g)\right];p_i\right\} \qquad (3)$$

The functions T_g and T_i can be deterministic or stochastic. An arbitrary pair of points (P',P") can have two states: the first, when P' and P"are corespondent, the second when the points don't match. Thus, the Image Matching problem can be split in two parts:

1. Finding all the corespondent points
2. Determination of parameters p_g and p_i of the functions T_g and T_i.

The solution are always based on the intensity functions g' and g" and on the attributes a' and a" of the points in the vicinity of the points (r',c') and (r",c") such as:

P' = P'(r',c';a')

P" = P"(r",c";a")

and consists in a three steps procedure:

1. selection of the interest points in both images using lines of interest in each image
2. finding the corespondent pairs of interest points P' and P" using either the criteria of the similarity based on the attributes of intensity (which is essentially an *area based matching* (ABM) procedure) or the consistence criteria based on the mapping functions (which is essentially an *feature based matching* (FBM) procedure)
3. interpolation between the interest points, only for stereo images

In order to explain the determination of the similarity by an ABM procedure, we present in detail a method of differential recognition for one-dimensional signals expressed in terms of variances, which is a relevant procedure for the location of an object when the image is affected by noise. The model (3) can be redefined as:

$$g(x) = T_i\left\{f\left[T_g(y;p_g)\right];p_i\right\} + n(x) \qquad (4)$$

where g',g", (r',c') and (r", c") was replaced with g, f, x and y and n(x) is the observable component of noise. We assume that the signals represent two images of the same scene and there are no changes in intensity on the direction of the image caption. In this case:

$$g(x_i) = f[x_i - \hat{u}(x_i)]$$

where $\hat{u}(x_i)$ is the unknown distortion in determining the correspondent point. A nonlinear model valid for the m observed values of $g(x_i)$ can be expressed as:

$$g(x_i) = f[x_i - \hat{u}(x_i)] + n(x_i); i = 1..m$$

$n(x_i)$ is the additive noise in position x_i with the variance σ_n^2. If we already have an approximation u_0 of the function $\hat{u}$, then $\hat{u}(x_i) = u_0(x_i) + \Delta\hat{u}(x_i)$ where $\Delta\hat{u}(x_i)$ is the correction for $\hat{u}(x_i)$. The nonlinear model becomes:

$$g_i = f(x_i - \hat{u}_i) + n_i = f(x_i - u_{0i} - \Delta\hat{u}_i) \quad ; i = 1....m \qquad (5)$$

and with the general notation $p_i = p(x_i)$, a linearization around the point $x_i - u_{0i}$ gives:

$$g_i = f(x_i - u_{0i}) - f'(x_i - u_{0i}) \cdot \Delta\hat{u}_i + 0{,}5 \cdot f''(x_i - u_{0i} - \xi \cdot \Delta\hat{u}_i) \cdot (\Delta\hat{u}_i)^2$$

with $\xi \in [0,1]$.

If we neglect the second order terms, we obtain a linear model where Δg_i and f_i' are known and $\Delta\hat{u}_i$ is unknown:

$$\Delta g_i = -f_i' \cdot \Delta\hat{u}_i + n_i \quad ; \quad i = 1....m \qquad (6)$$

where $\Delta g_i = \Delta g(x_i) = g(x_i) - f(x_i - u_{0i})$

Since the variance $\sigma_{\Delta g}^2$ of Δg is equal with the variance of the noise σ_n^2, we obtain for the standard deviation σ_{ni} of the estimate of:

$$\hat{u}_i = u_0(x_i) + \Delta\hat{u}_i$$

$$\sigma_{ni} = \sigma_n / f_i' \qquad (7)$$

The precision of the estimate depends of the derivative f_i' of the given function.

Because the supposition on the invariance of the intensity affects strongly the estimate, supplementary restrictions are necessary to embed the result.

The method presented above, with appropriate changes, was utilized in order to:

- estimate an unknown displacement, assuming $\hat{u}(x)$ only as an uniform displacement
- estimate an unknown displacement and a scaling factor, introducing in the model a compensation parameter s for the unknown scale, i.e. $g_i = s \cdot f(x_i - x_0) = s \cdot x_{\_i} - \hat{u}$
- introduce compensation parameters for illumination and contrast, if this values are different in the two images, by using a nonlinear model $g(x_i) = a \cdot f(x_i - \hat{u}_i) + b + n(x_i)$ where a represents the change in contrast illumination and b the change in illumination
- recognize to contours affected by noise
- recognize two bi-dimensional signals f(r,c) and g(r,c)

3. DETERMINATION OF INVARIANT POINTS

The following properties are essentially to define interest points, which can later be used as invariant points:

- discrimination, i.e. such a point is distinct comparing to its neighbors

- invariance, i.e. the selected position is invariant to geometric transforms .
- stability, i.e. all the selected points are invariant to the focus angle; as an example, all the corner points are stable
- uniqueness, i.e. this points don't have repetitive attributes

The interest points have a special meaning, being usually: corners, lines junctions, centers of circles, rings etc. The following three-step procedure can be utilized in order to obtain such points:

1. selection of the optimal window, using as measure the mean value of the gradient for a windows of specified size, which is invariant to rotation;
2. classification of the image function from the selected window based on isotropic features;
3. estimation of the optimal points of the windows

A typical procedure to determine invariant points when they are transferred in other windows consist in the use of a model with only three parameters: two displacements u and v and a scaling parameter having the following symmetric matrix of the normal equation:

$$N = \begin{pmatrix} N_{11} & N_{12} & N_{13} \\ N_{21} & N_{22} & N_{23} \\ N_{31} & N_{32} & N_{33} \end{pmatrix} = \begin{pmatrix} \sum_i f_{ri}^2 & \sum_i f_{ri} \cdot f_{ci} & \sum_i f_{ri} \cdot (f_{ri} \cdot \bar{r}_i + f_{ci} \cdot \bar{c}_i) \\ & \sum_i f_{ci}^2 & \sum_i f_{ci} \cdot (f_{ri} \cdot \bar{r}_i + f_{ci} \cdot \bar{c}_i) \\ & & \Sigma (f_{ri} \cdot \bar{r}_i + f_{cci} \cdot \bar{c}_i) \end{pmatrix} \quad (8)$$

using $\bar{r}_i = r_i - r_0 and \bar{c}_i = c_i - c_0$ with the unknown reference point (r_0, c_0). If we desire that the estimates of the two displacements *u* and *v* are independent of the estimated scale, we must determine r_0 and c_0 from:

$$N_{13} = 0 = \Sigma f_{ri} \cdot \left[f_{ri} \cdot (r_i - r_0) + f_{ci} \cdot (c_i - c_0) \right]$$

$$N_{23} = 0 = \Sigma f_{ci} \cdot \left[f_{ri} \cdot (r_i - r_0) + f_{ci} \cdot (c_i - c_0) \right]$$

which lead to the following 2x2 equation system:

$$\begin{pmatrix} \sum_i f_{ri}^2 & \sum_i f_{ri} \cdot f_{ci} \\ \sum_i f_{ri} \cdot f_{ci} & \sum_i f_{ci}^2 \end{pmatrix} \cdot \begin{pmatrix} r_0 \\ c_0 \end{pmatrix} = \begin{pmatrix} \sum_i (f_{ri}^2 \cdot r_i + f_{ri} \cdot f_{ci} \cdot c_i) \\ \sum_i (f_{ri} \cdot f_{ci} \cdot r_i + f_{ci}^2 \cdot c_i) \end{pmatrix} \quad (9)$$

The point (r_0, c_0) has the following important properties:

1. The transferred points $(r_0 - \hat{r}, c_0 - \hat{c})$ have minimum variance
2. The invariance of (r_0, c_0) is typical for corner points, because in every junction of lines all the gradients $\nabla f_i = (f_{ri}, f_{ci})$ are orthogonal with ($r_i - r_0, c_i - c_0$) leading to the condition: $f_{ri} \cdot \bar{r}_i + f_{ci} \cdot \bar{c}_i = 0$ and so $N_{13} = N_{23} = 0$

The same procedure can be applied to determine the point (r^*_0, c^*_0) which is invariant to rotation when is moving in other windows.

Now we can discuss the procedure to estimate corner points. We assume a windows of size m_r x m_c which contains a corner point p_0 = (r_0,c_0). We want to obtain an estimate

$\hat{p}_0 = (\hat{r}_0, \hat{c}_0)$. Because we don't know how many edges intersects in p_0, we take every individual element of an edge as representative for a line which crosses $p_i = (r_i, c_i)$ with $i = 1...m$. We assume that the estimated point $\hat{p}_i$ is closest to all the lines intersects by edges, and let: $p'_i = (r_i, c_i)'$ a pixel in the window with the gradient:

$$e'_i = e(r_i, c_i) = \nabla f'_i = (f_{ri}, f_{ci}) = \left[f_{ri}(r_i, c_i), f_{ci}(r_i, c_i)\right]$$ and with

$$\frac{\nabla f'_i}{|\Delta f_i|} = (\cos\Phi_i, \sin\Phi_i) = \left[\cos\Phi(r_i, c_i), \sin(r_i, c_i)\right]$$

the unit vector on the gradient direction. The line passing through $p_i = (r_i, c_i)$, parallel with the edge direction is given by:

$$(p - p_i)' \cdot e_i = 0 \text{ with } p' = (r, c)'$$

In order to obtain a noniterative solution, we consider the Φ_i of the edge element (the normal on ∇f_i in (r_i, c_i) without error. The original estimate is the distance $l_i = p_i * e_i$. The weight of the independent edge elements is assumed to be $w_i = \|\nabla f_i\|^2$. The linear model becomes:

$$l(r_i, c_i) = \cos\Phi(r_i, c_i) \cdot \hat{r}_0 + \sin\Phi(r_i, c_i) \cdot \hat{c}_0 + n(r_i, c_i) \qquad (10)$$

If now we consider (a_i, s_i) as a representation of the edge element in a Hough space we can assume the equation of the model (10) as a line in this space with unknown parameters $(\hat{r}_0, \hat{c}_0)$ which in the image correspond to the intersection of the edge elements. By choosing the right value for the estimate a_i, or for $\|\nabla f_i\|^2$ we obtain the same normal system of equations. The same procedure can be extended in the case when the window contains a symmetric circular structure (as a circle or a ring) The main idea is also to use the lines crossing the points $p'_i = (r_i, c_i)$ with the gradient direction ∇f_i.

4. THE DEVELOPMENT ENVIRONMENT AND THE EVALUATION OF THE SIMULATION RESULTS

In order to simulate IMM procedure we have conceived a GUI Windows environment which offer the following facilities:

1. representation of the robot scene, with the possibility to implement simulated obstacles in this scene, but also with the possibility to transpose images of the real environment taken with CCD camera
2. positioning of a mobile robot in this scene, with the possibility to choose a trajectory between the departure and respectively the arrival position
3. simulation of noise on contours of the patterns represented in the robot scene
4. embedding pattern recognition using IMM and generation of motion control commands, as a consequence of the decision based on IMM estimates.

The following program modules are offered:

a. *for the user interface*

- MAIN, which create windows to represent images; several dialog boxes to launch IMM operations as establishing invariant points, compensating parameters determination and also buttons which allow to optimize IMM by choosing the adequate number of iterations, the interpolation method and the image dimension.

- DIALOG to exports data from resource files.
- WGRAYIMG to define function for image processing.
- WINIT to initialize the application.
    b. *for image analysis*
- POINT which allows the construction of the matrix for corner points determination.
- SYSTEM to solve the normal equation system through gaussian triangulation

The simulation confirm that in many cases the spatial variation of the cross-correlation coefficient makes rather difficult to find its maximum and so ambiguous solutions may arise when comparing IMM results. We try to propose more refined procedures which guarantee that a solution exists, is unique and is stable with respect to small variation in the input data.

Because the least mean square method is a particular form of the maximum likelihood principle for errors with Gaussian distribution, in order to evaluate the final results we can use the true density function of the estimation process, in a four-step procedure:

1. A global check of how consistent are data and model using Ω (the sum of square of residues for all accepted correspondence). The ratio Ω / σ_0^2 is distributed with X_{2m-6} degree of freedom. It can be compared with an expected value of 1/2 pixels using a Fisher test (4). If the test is rejected, we conclude that the model was too much simplified
2. The precision of parallaxes can be determinate as: $\sigma_{\hat{p}k} = \sigma_{\hat{q}k} = \sigma_0 \sqrt{h_k}$
3. We check if we have enough corresponding points. If we renounce at a corresponding point, the maximum influence of a parameter a_i is limited to: $\Delta_k a_i \leq \frac{1}{1-h_k} \sqrt{n'_k n_k h_k q_{ii}}$
4. We check the final cross-correlation coefficient in order to verify if all the information of the original images was used.

5. CONCLUSIONS

We tried to obtain several matching algorithms using both ABM and FBM. The main criteria to evaluate similarity in ABM procedures was the covariance or cross-correlation between the corresponding values of intensity (gray levels). We have also used FBM procedures in order to detect contours of obstacles in a robot scene when these contours where affected by noise and with this aim we have used points of interest which are invariant to translation or rotation as point corners. The conclusion is that a well designed matching algorithm, i.e. the proper combination of various of the mentioned solutions associated with internal self evaluation can be successfully used as a powerful technique of artificial vision for mobile robots which works in a weak-structured environment.. Further researches are focussed on the improvement of the real time use of the IMM procedures

REFERENCES

[1] D. Popescu, G. Ionescu, R. Dobrescu - Sensor Fusion for Mobile Robot Control, Proceedings of ISCMR'95
[2] G.I. Shapiro, R.M. Haralick – Relational matching, Applied Optics, vol.26, no.10, 1987
[3] C. Heipke – Digitale photogrammetrische Arbeitsstationen, DKG – C, 1995
[4] R. Duda - Pattern classification and scene analysis, J. Wiley & Sons, NY, 1996