MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

TARGETTING NUMBER PLATES EFFECTIVELY USING SPARSE/FULL TEMPLATES AND COARSE/FINE TEMPLATE MARCHING

Harvey A. Cohen and Alan L. Harvey
Department of Computer Science
and Computer Engineering
La Trobe University

Bundoora, Victoria, Australia 3083

ABSTRACT

A coarse-to-fine approach to object location via template matching is described that entails the progression from sparse to full templates, and the progression from large step to small step "marching" of the templates through an image. Our "coarse-to-fine" approach may be loosely described as "multigrid" or "multi-resolution" although it differs in significant detail from previous work. The approach was developed in the context of an assessment of possible fully automated analysis of police camera films of traffic and speed violators. In this paper we describe how our approach has been applied to the location of the border of number plates in selected images. Elsewhere, we shall discuss the application of the "coarse-to-fine" approach to the actual reading of the characters on the number plates so "targetted", allowing for tilt, perspective and scale change. Herein we describe the ndividual strategies that have been implemented and present theoretical speed-up and actual timing data for performance on a UNIX system. The individual results are generally impressive, although gathered only on a very limited set of images. In this assessment project a complete integrated package has not been developed.

INTRODUCTION

This paper arose from a concept development project designed to develop and test practical means for the automatic reading of number plates within images captured by police cameras. Our starting point was the use of template matching applied to the two stage process of firstly locating the numberplate within the scene, and secondly reading the alphanumeric characters on the numberplate. In order to cover a range of possibilities of orientation angle and size, as well as actual variation of numberplate format between various states, it is well known [2][3] that in a developed system multiple applications of templates from generic families would be required. Classic template matching is computationally expensive, and multiple application of a number of templates further exacerbates the situation. Possible alternate strategies involving feature selection could offer some scope for scale and rotation dependence, and may well be more akin to human methods of reading numberpates; however in the very cluttered environment of a street scene these techniques would undoubtedly be more computationally expensive. Hence the need to speed-up template matching provided the basic framework. From this work evolved a general coarse to fine strategy of template matching that utilised a gradient change criterion for switching from coarse to fine. Our approach is related to earlier work on "multi-resolution" and hierarchical template matching, and to "multi-grid" methods in numerical analysis. The basic goal was to develop software methods appropriate to a conventional supermicrocomputer so that hardware strategies utilising DSP processors, or processor arrays were not considered.

The plan of this paper is to first review template matching, and introduce our coarse/fine strategy, then present the data. Finally we present our conclusions and discuss the relationship to other approaches

TEMPLATE MATCHING

Template matching involves moving a template across the image (or window of interest). In this paper we refer to the method of scanning through the image as "marching". The template is marched through the image until at some image location the template matching error is less than some threshold. The simplest such measures of match are:

• the Chebyshev distance, the sum of the absolute errors between the image I and the template T:

$$E_C = \sum\sum |I_{i+r,j+s} - T_{r,s}|$$

• and the Euclidean distance:

$$E_S = \sum\sum(I_{i+r,j+s} - T_{r,s})^2$$

The Euclidean norm tends to 'penalise' outliers far more than the Chebyshev, with prospects of causing large variations from match due to minor erros in scale and orientation. The computational costs of these two norms depends on the processor architecture (and instruction set), as the Chebyshev norm computation implicitly implies a test and branch operation. For the SUN workstations we utilised the Chebyshev norm is indeed cheaper and this was used in all the calculations reported here.

We must mention here the potential use of correlation functions as norms:

$$E_{cor} = -\sum\sum I_{i+r,j+s} \cdot T_{r,s}$$

Correlation functions are in general more expensive than the simple Chebyshev norm. Within the context of grey-scale imagery, they probably offer the simplest means for compensating for lighting and the like variation across the targetted object. Within the ambit of our assessment, we restricted attention to the use of the Chebyshev norm as matching error function.

Note, that for some processors, such as the DSP TMS 320, the Euclidean norm can be faster to compute than the Chebyshev.

SPEED-UP ISSUES

The basic scheme of object location using templates is

(i) to apply the template to the image at an image location and to determine a figure of merit for the local match

(ii) to march the template about the image until a satisfactory figure of merit is obtained.

For an image size N*N, template size M*M, a complete passage through the image will involve (N-M)*(N-M)*M*M operations (dependent on the merit function utilised). In order to speed-up the matching process a multi-resolution approach was developed and given a preliminary evaluation. In this approach, which can be characterised as a coarse/fine approach we implemented and evaluated the following:

Coarse-Fine Marching: To reduce the number of points where an (unsuccessful) match is made, the template is marched across the image by "coarse" raster strides. When at a given test point, the rate of change in the merit function indicates proximity of the match point, thenceforth "fine" marching is applied. The maxum stride is related to the width of the "dip" in the merit function about the location of match (of number plate templates):

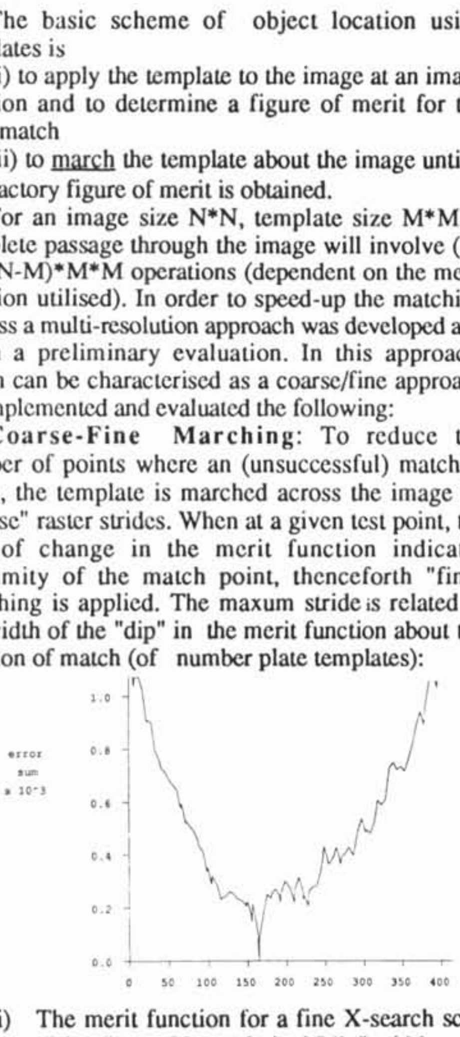

Fig (i) The merit function for a fine X-search scan along the "right" row. Note 16 pixel "dip" width.

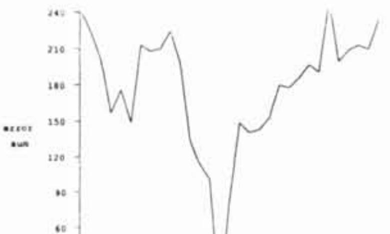

Fig (ii). Expansion of "dip" region of Fig (i)

Sparse Templates: To reduce the template size, a reduced (or sparse) template is applied until close to match, when a full(er) template is utilised. In the work reported here we report on the time to locate using a range of reduced templates, applied, without switching on a noise -free image.

COARSE-FINE TEMPLATE MARCHING

As a first approach to object location in an image, the template to be matched to the image may be moved over the image, pixel by pixel and the matching error computed. Clearly if the matching calculations could be made at a reduced number of locations, i.e., in a coarse search, with a switch to a finer search when close to match, then a large computational saving may be made. (See Fig(i)

Coarse-fine X Marching: The plots of Fig(i) and (ii) support the strategy of narching the (full) template along rows by multiple column steps until approaching match location, then switching to scolumn by column dvance. The simplest criteria to use for switching is a threshold level for the matching error function. However, in a cluttered environment, as in street scenes, a more robust measure of approach to match-ing location is provided by the rate of change of the error function, so that in effect the gradient of the error function is used as the trigger for change. The specific criterion we have used has been to switch from coarse to fine when the figure of merit decreases by a fixed percentage $\alpha\%$ in successive steps. Data follows:

Coarse-fine X-Marching
Columns Coarse to Fine

Column Spacing	Search time secs	Speed-up Factor
1	28.7	1.0
2	14.3	2.0
3	9.9	2.9
4	7.1	2.9
5	6.2	4.0
6	4.9	5.9
7	4.7	6.1
8	fail	fail

These results are for coarse-fine searching using full template that is marching along consecutive rows, starting from top left corner of window of interest with indicated step size, followed by unit step size marching to match location.

In assessing these results for coarse/fine X marching, recourse should be had to Fig (ii) which indicates the width of the dip in the figure of merit near the match point. When step size is more than half the width of the "dip" coarse/fine search fails.

Coarse-fine Y Marching: In coarse/fine X Marching, an $\alpha\%$ decrease in successively evaluated merit function triggers the switch from coarse to fine. This criterion is NOT appropriate for Y Marching, where one requires a measure of the Y gradient for switching. Our algorithm for Y-marching requires the maintaining of a record of the smallest figure of merit in a row, and when the figure of merit in the next row is more than $\alpha\%$ smaller, then there is a switch to fine Y marching. Using this switching criterion, the evaluated search time, starting from top left of the image, to locate the number plate, using a range of row steps from 1 to 8:

Coarse-fine Y-Marching Rows Coarse to Fine		
Row Spacing	Search time secs	Speed-up Factor
1	22.9	1.0
2	12.3	1.9
3	9.2	2.5
4	7.9	2.9
5	6.6	3.5
6	4.6	5.0
7	4.6	5.0
8	fail	fail
These results are for searching using full template that are marched column by column along rows with indicated row spacing, as explained in the text.		

The above results show that a law of diminishing returns applies to increasing row size, and in fact scans with the coarse row steps 6 and 7 (prior to adoption of single row advance at the end of a line scan) took the same time; while for row step size of 8 our coarse-fine algorithm failed.

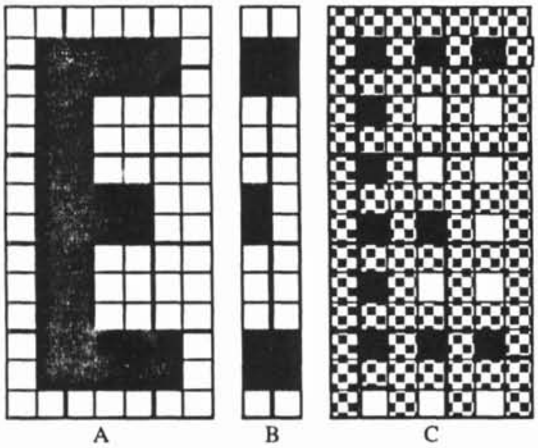

Fig (iii). Illustrating partial and sparse templates: For simplicity, a binary template is shown. Checkerboard pixels are ignored.

A Regular template 7*14
B Partial template 2*14
C Sparse Template 3*7

SPARSE TEMPLATES

Instead of using the entire template to match pixel by pixel against the image it is reasonable to sub-sample the template, but to continue to utilise the entire image. Consider the example of templates for the letter E in Fig (iii) (above).

Note that in Fig (iii) the essential features of the letter E are clear for both a sparse template, denoted by C, and a ("smart") reduced template, denoted by B. Clearly the sparse template would reduce the signal to noise ratio of the matching function as compared with a complete templay array so that method should not be used with very noisy images.

We have examined the speed up in the "ideal" case where the template exactly matched the number plate on a particular image. Thus there was no noise, and exact matching was possible. Experimental results verified the expected speed-up.

Sparse Template		
Row and Col Spacing	Search time secs	Speed-up Factor
1	346.1	1.0
2	86.1	4.0
4	21.8	15.9
8	6.1	56.7
16	<1	>346
These results are for searching using sparse template that are marched column by column along successive rows until match achieved.		

In the pyramidal approach of Rosenfeld and Vanderberg [1] both template and image are (essentially) averaged over pixel blocks. For images of significant variation, such as alphanumeric characters, the averaging process applied to the template might well be "out of phase" to an extreme extent making close matching unlikely. Nevertheless, the essential feature of their approach [1] warrants following, particularly where generic templates have been applied. What would be entailed is that as a near match is achieved, by whatever criteria, a less sparse template should be utilised.

DISCUSSION

In the coarse-fine approach reported here, we have adopted a "conservative" figure to determine when to switch from coarse to fine. The data presented here refer solely to forward marching. Were one to adopt a less conservative measure: when it appears that the region of near-match has been reached, the system would have backtracks to the previous test point, and from there resume marching from pixel to pixel along the raster. For arbitary images, using generic templates, this sort of back-tracking would presumeably become obligatory.

We have restricted our analysis to the use of simple templates derived from the grey-scale images actually studied. These templates were 96*46 pixel size (except where sparse templates were used), and included the border region.

Caelli and Liu [3] have pointed that only a finite number of templates to required cover ALL possible rotations and scale changes of an image. In our application,there is in fact only a limited possible range of scale, and a limited range of angles of rotations, so that their point is even more true. Implementation will involve the determination of appropriate generic set of templates for the numberplates of various states.

The work reported here has been concerned with the targeting of number plates, for which there is not really an appropriate reduced template (see fig(iii) B)

Unfortunately, some number plates have light characters on a dark background. Thus the dark and light number plates are essentially twop different classes of objects, for which some form of hierarchical analysis, such as has been given by Liu and Caelli, [2] and [3] is appropriate.

A complete package was not developed, but each speed-up strategy was individually confirmed in a limited study, Results are highly encouraging, but major further development is required.

MULTI-RESOLUTION TEMPLATE MATCHING

It is salient to compare our multi-resolution approach to that contained in a classic paper of Rosenfeld and Vanderberg, and a more recent scheme propounded by Caelli and Liu.

Rosenfeld and Vanderberg [1] proponded a multi-resolution approach where are a reduced image is produled by summing over blocks of cells, e.g., for a 2*2 block reduction the new image $I_{i,j}$, is given by

$$I_{2i,2j} + I_{2i+1,2j} + I_{2i,2j+1} + I_{2i+1,2j+1}$$

and a similar reduction process is applied to the template (extended if necessary). This smearing if applied "out of phase" to significant changes in template (and taget object) can lead to match failure.

In contrast, in our approach, we work solely with the original image, and "turn-off" i.e., ignore, most template elements in the sparse state.

Caeli and Liu [2] have considered the use of generic prototypes for templates of three very limited classes of object. However, they assume the object is already targetted, so that translation issues paramount in this paper are not considered. Nevertheless, this work is pertinent to extension of our work to the more general task of detecting the number plate frame of a greater range of number plates subjected to scale and orientation differences.

ACKNOWLEDGEMENT

Partial support from the Victorian Police Ministry is acknowledged.

REFERENCES

[1] Rosenfeld, A and G.J. Vanderberg, *Coarse-fine Template Matching* IEEE Trans, Systems, Man and Cybernetics, SMC-7 No 2, pp 104-107 (1977)

[2] Caelli, T.M. and Z.Q. Liu, *Multi-object Pattern Recognition and Detection in Noisy Backgrounds Using a Hierarchical Approach* Computer Vision Graphics and Image Processing Vol 44 pp 296-306 (1988)

[3] Caelli, T.M. and Z.Q. Liu, *On the Minimum Number of Templates Required for Shift, Rotation and Size Invariant Pattern Recognition* Pattern Recognition Vol 23m No 3, pp 205-216 (1988)

[4] A.L. Harvey and H.A. Cohen, *Speed up Techniques for Image Object Recognition* Proceedings of the Int. Conference on Robotics and Computer Vision, ICARV'90, Singapore, September 18-21, pp 965-968.(1990)

[5] Ikeuchi,K. and T. Kanade, *Automatic Generation of Object Recognition Programs* Proc IEEE August 1988, pp 1016-35 (1988)

[6] Gromson, W.E.L., *Recognition of object families using parametrised models* Proc, IEEE Intl. Conf. Computer Vision, ICCV'87, London, pp 93-101 (1987)

Fig(iv) From police speed camera a window of interest showing car with number plate to be targetted for reading