Research Journal of Applied Sciences, Engineering and Technology 4(24): 5469-5473, 2012
ISSN: 2040-7467

Submitted: March 18, 2012 Accepted: April 14, 2012 Published: December 15, 2012

Review Article: An Overview of Template Matching Technique in Image Processing

T. Mahalakshmi, R. Muthaiah and P. Swaminathan
School of Computing, SASTRA University, Thanjavur, Tamil Nadu, India

Abstract: Template matching is one of the areas of profound interests in recent times. It has turned out to be a revolution in the field of computer vision. Template matching provides a new dimension into the image-processing capabilities, although there have been many attempts to resolve different issues in this field there have always been newer concepts emerging in this ever challenging field. In this study, various techniques are discussed and a new FPGA based spectral architecture is proposed to achieve a fast similarity measure between template and input image.

Keywords: FPGA, image processing, similarity measure, spectral architecture, template matching

INTRODUCTION

Template matching is a technique used in classifying an object by comparing portions of images with another image. One of the important techniques in Digital image processing is template matching. Template matching is widely used for processing images and pictures. Some of its wide-spread applications object to location, edge detection of images, to plot a route for mobile robot and in image registration techniques. In general, a technique includes its unique algorithm or method, which compares the template image with input image and finds similarity between them. Some advanced methods will have a feedback path to verify its correctness.

Template matching approaches: The choice of matching depends on the nature of the image and the problem to be solved. General classifications of template or image matching approaches are: Template or Area-based approaches and Feature-based approaches. The discussion below gives an outline of both the approaches mentioned above and an idea to deal with motion tracking and occlusion handling.

Featured-based approach: Featured-based approach is well suited when both reference and template images had more correspondence with respect to features and control points. Features include points, curves, or a surface model that have to be matched. Here, the aim is to locate the pair wise connection between reference and template using their spatial relations or descriptors of features. Sub-categories of the above approach is spatial relations, invariant descriptors, pyramids and wavelets and relaxation methods.

Area-based approach: Area-based methods are sometimes called as correlation-like methods or template matching: Fonseca and Manjunath (1996) which is the combination of feature detection and feature matching. This method is best suited for the templates which have no strong features with image, since they operate directly on the bulk of values. Matches are estimated based on the intensity values of both image and template. Following are the techniques under this category: squared differences in fixed intensities, correction-based methods, optimization methods and mutual information.

Motion tracking and occlusion handling: For the templates which may not provide a direct match, then Eigen spaces are used, which gives the detail of matching image under various conditions, as illumination, colour contrast or acceptable matching poses. Example, if the user was searching for a specimen, the Eigen spaces may consist of templates of specimen in different positions to camera with different lighting conditions or expression: Luis *et al*. (2009). There are possibilities for the matching image to be occluded by an object or problems involved in motion become ambiguous. One of the possible solutions for this is to split the template into multiple sub-images and perform matching on them by Jurie and Dhome (2002)

SIMILARITY MEASURES

Number of techniques has been developed and still developing to measure similarities between input/reference image and the template. Matching accuracy of the measure depends on the type of method or algorithm selected; kind of problem which is to be solved, type of template and for what application it has to be done. Even though numbers of comparison techniques

Corresponding Author: T. Mahalakshmi, School of Computing, SASTRA University, Thanjavur, Tamil Nadu, India

have been developed, we can‘t produce the best result in all situations with a single method. However, each method works better than other; depending on the above criteria. We will discuss some of the techniques and algorithms developed regardless of application. Assume the reference Image (I) is of size m×m×m and Template (T) of size n×n×n. When (m>n), the intermediate image or Similarity image (S) will be of size of (m−n+1) × (m−n+1) × (m−n+1) at location (x, y, z).

Matching technique not only takes the similarity measure but also calculates the error between images depending on its difference using Mean Squared Error (MSE) metric:

$$MSE = \frac{\sum\sum\left(Temporary(x,y) - T\arg et(x,y)\right)}{Number\ of\ \ pixels}$$

where, Temporary (x, y) and Target (x, y) are the intensity values of Template and input image, respectively. The Similarity or the matching between them is inversely proportional to the MSE. Let‘s discuss some of the techniques and algorithms developed on the basis of TM.

Sum of Absolute Differences (SAD):

$$s(x,y,z) = \sum{}_{i=1}^{n}\sum{}_{i=1}^{n}\sum{}_{ki=1}^{n}$$
$$\left|f_1(i,j,k) - f_2(x+t-1, y+f-1, z+k-1)\right|$$

Sum of absolute intensity differences of order one is defined by Devijver and Kittler (1982).

where, x, y, z = 1...m-n+1. This measurement requires an addition in the order of n^3 for each search location. Barnea and Silverman (1972) proposed an algorithm, which speed up this calculation by an assumption that smin is the minimum of ‘s' calculated till now then the algorithm compares each ‘s’ calculated by above equation with smin, if it is equal or greater than smin then the measurement is stopped. Else smin is replaced with any one of the lowest calculated s values which speeds up the measure by eliminating unnecessary calculations but still the complexity is in order of O (n^3m^3) additions, when m>n. Since this similarity, measure finds the SAD between image intensities of same modality.

Penney *et al*. (1998) says that image gradients can produce accurate matches than image intensities (raw).

Geometric distance: It is a preferred technique for similarity measure, when both template and image has binary structures (image edges). SAD and CC cannot be used for matching when image and template are taken with different sensors. In this case, images can be alike to obtain similar surface structures, these surface structures are used for matching. This process is known as chamfer matching. This technique was proposed by Borgefors, (1996) for matching volumetric images.

Chamfer matching algorithm uses average distance between two binary structures as a measure. That is the image and the template are overlaid and the closest structure points between them are determined and the average of the distances between the corresponding points are used in similarity measure. The match is said to be the best when average distance between structures is minimum. Match is said to be accurate when a mean is zero and as mean increases the similarity between image and template decreases.

Computation takes in the order of m^3 additions. The time needed to find out the best match position is O (Nd^3+m^3) additions. Computational complexity of this technique is O (m^3) additions when m = n. When m>>n, computational complexity is O (m^3N) additions. Segmentation time is not included in this. This approach succeeds only when the subset of template and image matches, which is similar to surface matching method, which aligns surfaces using optimization method. Image intensities produce accurate results than surface structures since they discard intensity information‘s: Kularatna, (1991) and Fitzpatrick and West (1998).

Mutual information:

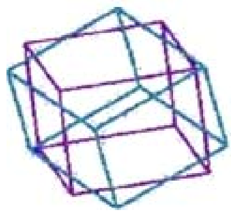
$$I(T,W) = \sum a\sum bP_{Tw}(a,b)\log\frac{P_{Tw}(a,b)}{P_T(a)\,P_w(b)}$$

Assuming a Template (T) and image (W) are to be compared for determining their similarity. If we overlie the template and image, the probability that intensity a over b will be equal to its joint probability: P_{TW} (a, b) else their probability will be P_T(a) and P_W(b), where, P_T(a) is the probability of intensity in template and P_W (b) is the probability of intensity in image. If both correspond to each other than joint probability will be high else the probability will be small. Computations of mutual information will be I (T, W) described by Maes *et al*. (1997)

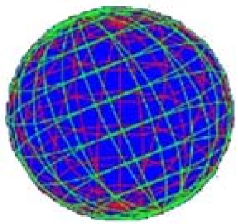
$$P_T(a) = \sum{}_{a=0}^{255}P_{Tw}(a,b);\ P_w(b) = \sum{}_{a=0}^{255}P_{Tw}(a,b)\ P_{TW}\ (a,\ b)$$

is evaluated from the histogram. To obtain the histogram, an array (for example 256*256) is allocated and all its entries are initialized to zero. It is known that the intensities vary from 0 to 255. After dealing out all voxels in the template and the image, an array will be obtained with (a, b). The entry of obtained array gives P_{TW} (a, b) for different values of (a) and (b). We can estimate P_T(a) and P_W(b) from the above array using

High mutual information is obtained when intensities of both image and template correlate. This method is more sensitive to image noise. Noise either in image or



(a)



(b)

Fig. 1: (a) A cubic template and a cubic window whose centers coincide, (b) A spherical window whose centers coincide

template will quickly corrupt the similarity measure. In such cases image is smoothened which reduces noise and improves accuracy of mutual information. For each shift position, computational complexity is n^3 additions and 2562 multiplications. Matching time, when m>>n is O (n^3m^3) additions. If m = n computation time will be O (n^3m^3) additions. This algorithm is more accurate when, there exists a rotational difference between image and template, but invariant moment should be used to align them.

Invariant moments: It is a very helpful method, when matching images have rotational differences. Since this difference makes all the above mentioned approaches ineffective in template matching. This method is free from the orientation of a pattern. Invariant of the position and orientation of a pattern can be obtained by normalizing moments: Hu (1962):

$$\mu_{pqr} = \sum_{t=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(t-\bar{t})p(f-\bar{f})q(k-\bar{k})rf_1(t,f,k)$$
$$m_{pqr} = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}t^p f^q k^r f_1(t,f,k)$$

Figure 1a shows a cubic template on a cubic window whose centres coincide; it is not possible to have such a view unless the orientation difference between them is 90 degrees. Even if you have that view the similarity measure is not enough to produce accurate match. But as a remedy for this spherical templates are used as shown in Fig. 1b since we can have the same pattern irrespective of their orientations. Yang *et al*. (1997) described the computation of geometry moments using discrete divergence theorem. The pqrth moment of f_1 (m_{pqr}) and central moment (μ_{pqr}) of the template is defined by:

Axes of a 3-D pattern are defined by eigenvectors of the inertia matrix which is real and symmetric and the matrix is:

$$M = \begin{bmatrix} \mu_{200} & \mu_{110} & \mu_{101} \\ \mu_{110} & \mu_{020} & \mu_{011} \\ \mu_{101} & \mu_{011} & \mu_{002} \end{bmatrix}$$

The amount of rotations required aligning the principal axes of the template and window depends on the rotational difference between them. Largest matching eigenvectors are aligned first and the second largest values are aligned followed by the third eigenvectors till the rotational difference between them is eliminated, their similarity can be found using any of the above mentioned approaches. For the match to be accurate, it is essential that no two eigenvalues should have the same magnitude and only the template should have large eigenvalues, if not, wrong axes could be aligned, missing the correct match.

The computational complexity is O (n^3m^3) multiplications if m>>n and O (m^3d^3) multiplications if m = n. Real estimation of the computational complexity is O (n^3m^4) or O (m^4d^3), since constant multiplication is involved in finding the eigenvalues.

CORRELATION METHODOLOGY

The conventional method of the area-based template matching approach is the normalized CC:

$$CC(i,j) = \frac{\sum w[w - F(w)]\left[I_{i,j} - F\left(I_{i,j}\right)\right]}{\sqrt{\sum w[w - F(w)]^2}\sqrt{\sum_{I(i,j)}\left[I_{i,j} - E\left(I_{i,j}\right)\right]^2}}$$

Similarity measure is computed between them to determine the maximum match. If there is a demand on sub pixel accuracy then interpolation of the CC values are to be used. This method is applied when there is a slight rotation and scaling other than this there are various generalized CC methods for geometrically deformed images, they compute CC for each geometric transformation of template image: Hanaizumi and Fujimura (1993). The computational complexity grows with the increase in transformation complexity. For this type of problems extended cross correlation based on increment sign correlation is suitable: Kaneko *et al*. (2003). Although in Sequential Similarity Detection Algorithm (SSDA), a sequential search approach and simpler distance measure which has been discussed

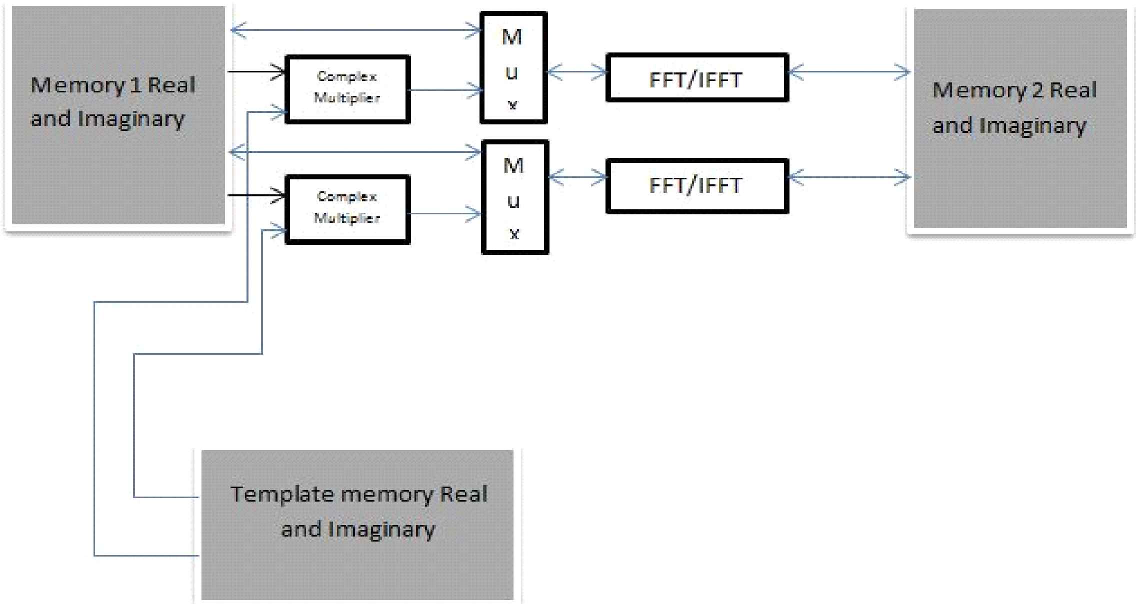


Fig. 2: Spectral architecture for template matching

above; likewise sum of squared differences measure was used which is faster but accuracy is likely to be less than cross correlation and is applied by Anuta (1970) for the evaluation of perspective deformation using piecewise for images that has been decomposed into small patches. The limitations of correlation methods are the flatness of the peaks and high computational complexity, the peaks can be sharpened by vector correlation: Chen *et al*. (1994). But still CC based methods are used for the easy implementation of hardware and processing image for real time applications.

Fourier methods: Fourier type methods are preferred than correlation when we require computational speedup and when the images get corrupted by frequency-dependent noise. That is a treatment in frequency domain. Fourier Shift Theorem is used in phase-correlation methods: Reddy and Chatterji (1996) which computes the power spectrum of the images from its cross-correlation and it looks for the peak in its inverse to be measured.

This method is proposed for translated images. It shows its strong strength against noise (frequency dependent) and illumination disturbances. This shows a good computational speed up when image sizes are large. Another method was proposed: Cideciyan, (1995) is an extension of phase correlation; if the scale difference is too large then mapping is done through the combination of fourier-mellin transform and phase correlation (Chen *et al*., 1994). They are used in applications such as medical imaging and remote sensing.

PROPOSED SYSTEM

From the above discussions it could be seen that there can be a better architecture which will improve the performance of template matching. It is found that template matching using cross correlation is better than other techniques on feature- based approach, Even though this technique has some limitation; it is used in real time image processing applications. Logic utilization in FPGA, Computational speedup in template matching can be obtained in a much efficient way, when the spatial correlation architecture is replaced with spectral correlation architecture which includes three FPGA memory blocks with further no external memory usage, parallel FFT computation using FFT slices and single complex multiplier. The cross correlation is computed with two 2 dimensional FFTs (one direct and one inverse) and a complex multiplication. The size of the image decides the number of FFT slices to be used.

In this architecture (Fig. 2) images are split into rows or columns to perform parallel FFT computation. Template image is stored in M_T and input image is stored in M 1. The similarity maxima are obtained by performing four steps:

- Compute FFT for the rows of M 1 and store the result in M 2.
- Compute FFT for the columns of M 2 and store the result in M 1.
- Compute inverse FFT of the rows of M 1 multiplied by the conjugate of the corresponding elements of M_T and store the result in M 2.

- Compute the inverse FFT of the columns of M 2 and store the result in M 1.

CONCLUSION

Feature-based matching methods are usually applied, when the structural information matches than the intensity information. They can also handle image distortions. Area-based approaches are preferable when they do not have many prominent details and the characteristic information (colour/gray than shape/size). Template and input images must have either statistical dependent or intensity similarity. If they have intensity similarities, then correlation methods can also be used. In geometric point of view, only small amount of shifts and rotations are allowed. If it is too large then the computational load will be high. To speed up this, we employ pyramidal and optimization algorithms to find out similarity on software implementations. While moving to hardware implementations; it is best to use FFT based correlations but it is limited for its computational effort. This can be overcome with the help of FPGAs. The proposed architecture will reduce the correlation computation by more than two orders of magnitudes with respect to software implementations without loss of accuracy.

REFERENCES

Anuta, P.E., 1970. Spatial registration of multispectral and multi temporal digital imagery using Fast Fourier Transform. IEEE T. Geosci. Electr., 8(4): 353-368.

Barnea, D.I. and H.F. Silverman, 1972. A class of algorithms for fast digital image registration. IEEE T. Comput., 21(2): 179-186.

Borgefors, G., 1996. On digital distance transforms in three dimensions. Comput. Vis. Image Understand., 64(3): 368-376.

Chen, Q., M. Defrise and F. Deconinck, 1994. Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition. IEEE T. Pattern Anal., 16: 1156-1168.

Cideciyan, A.V., 1995. Registration of ocular fundus images. IEEE Eng. Med. Biol., 14(1): 52-58.

Devijver, P.A. and J. Kittler, 1982. Pattern Recognition: A Statistical Approach. Prentice-Hall International, UK, pp: 232.

Fitzpatrick, J.M. and J.B. West, 1998. A Blinded Evaluation and Comparison of Image Registration Methods. In: Bowyer, K.W. and P.J. Phillips (Eds.), Reprinted in Empirical Evaluation Techniques in Computer Vision. IEEE Computer Society Press, Wiley, pp: 12-27.

Fonseca, L.M.G. and B.S. Manjunath, 1996. Registration techniques for multi-sensor remotely sensed imagery. Photogram. Eng. Rem. S., 62(9): 1049-1056.

Hanaizumi, H. and S. Fujimura, 1993. An automated method for registration of satellite remote sensing images. Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS'93, Tokyo, pp: 1348-1350.

Hu, M.K., 1962. Visual pattern recognition by moment invariants. IRE Trans. Inform. Theor., 8(2): 179-187.

Jurie, F. and M. Dhome, 2002. Real time robust template matching. The 13th British Machine Vision Conference (BMVC '02), pp: 123-132.

Kaneko, S., Y. Satoh and S. Igarashi, 2003. Using selective correlation coefficient for robust image registration. J. Pattern Recogn. Soc., 36(5): 1165-1173.

Kularatna, T., 1991. Volumetric image registration using chamfer matching. M.S. Thesis, CSE Department, Wright State University.

Luis, A.M., S. Dan and K.G. Walter, 2009. Expanding irregular graph pyramid for an approaching object. Proceedings of the 14th Iberoamerican Conference on Pattern Recognition: Progress in Pattern Recognition, Image Analysis, Computer Vision and Applications, Springer-Verlag Berlin, Heidelberg, pp: 885-891.

Maes, F., A. Collignon, D. Vandermeulen, G. Marchal and P. Suetens, 1997. Multimodality image registration by maximization of mutual information. IEEE T. Med. Imag., 16(2): 187-198.

Penney, G.P., J. Weese, J.A. Little, P. Desmedt, D.L.G. Hill and D. Hawkes, 1998. A comparison of similarity measures for use in 2-D and 3-D medical image registration. IEEE T. Med. Imag., 17(4): 586-595.

Reddy, B.S. and B.N. Chatterji, 1996. An FFT-based technique for translation, rotation and scale-invariant image registration. IEEE T. Image Proc., 5(8): 1266-1271.

Yang, L., F. Albregtsen and T. Taxt, 1997. Fast computation of three-dimensional geometric moments using a discrete divergence theorem and a generalization to higher dimensions. Graph. Mod. Image Proc., 59(2): 97-108.