National Conference on Innovative Paradigms in Engineering & Technology (NCIPET-2012)
Proceedings published by International Journal of Computer Applications® (IJCA)

Genetic Algorithm Based Dot Pattern Image Processing

Purshottam J. Assudani
Computer Science and Engineering
G.H. Raisoni College of Engg
Nagpur, India

Latesh G. Malik,
Professor & HOD
Computer Science and Engineering Department
G.H. Raisoni College of Engg
Nagpur, India

ABSTRACT

Dot pattern analysis and matching is necessary for many of the image analysis and pattern recognition problems. This paper uses local binary pattern for extracting the Dot pattern image features which is first pre-processed (Re-constructed, Rotated, Enhanced). It states that only the more discriminated features can be retained by discarding the less discriminated features using Genetic Algorithm. The optimized features thus obtained can be used for matching the two dot patterns for similarity using Euclidean Distance.

Keywords

Dot pattern; Local Binary Pattern; Genetic Alorithm; Euclidean Distance.

1. INTRODUCTION

A dot pattern is a set of dots or points in 2-D or 3-D space arranged to represent some physical objects or class of objects in the feature space. In processing visual information one often encounters dot patterns instead of gray level or color images. For example, objects in an image are often represented by the locations of some of their spatial features, such as spots, corners, etc. In pattern recognition dot patterns are ubiquitous. Representation of images by the values of certain features measured on the image parts provides dot patterns in the feature space. Pattern recognition procedures such as classification and clustering operate on such dot patterns. The nighttime sky is a natural dot pattern. Operations commonly performed include identifying certain star clusters and matching stored star patterns against repeated observations. An air traffic situation may be represented by specifying the point locations of aircrafts. This representation may then be used to detect potential collisions. Research on visual perception has made extensive use of dot patterns to investigate human image understanding. Any characterization or comparison of dot patterns must be in terms of the relative spatial arrangements of points.

Processing a dot pattern (DP) or point set in the plane is a useful and important problem in pattern recognition. Dot patterns encountered in various problems include points in feature space [1], pixels in a digital image [2,3], physical objects like stars in the galaxy [4] or spatial data [5–7]. Some of the sample Dot patterns are shown in Fig. 1, Fig. 2 and Fig. 3. Fig 1 shows the dot pattern of an alphabet, Fig 2 shows the dot pattern of a human being and fig. 3 shows the collection of dots forming a particular pattern. Thus, a dot pattern is a collection of dots that represents a particular object or shape in 2D or 3D.

Sample Dot Patterns:

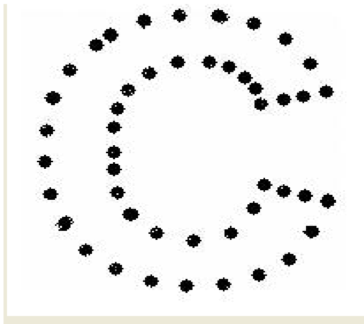

Fig 1: A Dot Pattern image

Fig 2: A Dot Pattern image.

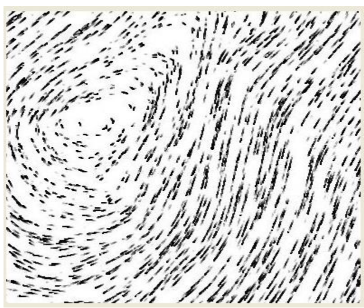

Fig 3: A Dot Pattern image.

Genetic Algorithm (GA) is one of the optimization algorithms, which is invented to mimic some of the processes observed in natural evolution. The Genetic Algorithm is stochastic search techniques based on the mechanism of natural selection and natural genetics. That is a general one, capable of being applied to an extremely wide range of problems. The GA, differing from conventional search techniques, start with an initial set of random solutions called population.

GAs attempt to find an optimal (or near optimal) subset of features for a particular problem [19-26]. First, a number of individuals or candidate Feature Subsets (FSs) are generated to form an initial population. Each FS is then function specific to the problem at hand. Parents are then selected based on fitness. New FSs are produced from the selected parents by the processes of reproduction. Survivors are selected from the previous generation and combined with the offspring to form the next generation. This process continues for user specified number of cycles evaluated and assigned a fitness obtained from the evaluation

Dot Pattern Matching (DPM) is an important step in many pattern recognition problems. Some of the application areas concern image registration, object recognition, object tracking, autonomous navigation, remote sensing, spatial information system, biomedical imaging, image texture analysis [8]-[10].

If there are two sets of points in *d*-dimensional space, one can determine whether there is a transformation that maps the first set onto or close to the second set of points. In general, the pattern matching can be categorized into two namely, *complete matching* and *approximate matching*. The complete matching occurs in an idealistic situation and in practice there exist spurious or lost points due to image degradation and binarization problem so that the exact match is not possible. Various optimization techniques have been applied to solve dot pattern matching problems. A simulated annealing technique was used in dot pattern matching [11]. A meta-heuristic particle swarm optimization algorithm [12] is also proposed in recent past. Zhang et. al. [10] employed genetic algorithm for dot pattern recognition and used the reference triplet points as the chromosome representation that reduced the search space significantly. The effective use of Genetic Algorithm (GA) in dot pattern recognition is also reported in various literatures [10],[13].

The paper is organized as follows. The following section describes about pre-processing steps used. Section 3 gives description about Local Binary pattern used for Dot pattern image feature extraction which can be optimized by Genetic Algorithm. Section 4 gives description about Euclidean distance which is used for Dot pattern matching. Section 5 describes in detail about Genetic Algorithm. Section 6 gives partial implementation details. Section 7 concludes the paper.

2. DOT PATTERN IMAGE ROTATION, RE-CONSTRUCTION, ENHANCEMENT

Controlling dot position is a key factor in controlling and effecting image quality in dot-based imaging systems. Often, halftone algorithms are developed to ensure minimal visibility. However, while the pattern is being rendered by the print engine, variations in imaging head speed or balance, jet-performance (in ink jet systems), paper handling, and other mechanical causes can disrupt the positions of the dots in the final pattern. In addition, interactions between marking and receiving media can further confound the rendering process by causing dot break-up or other effects. Dot placement errors not only effect the quality halftoned regions, but they also effect other features such as line fidelity (particularly that of fine lines) and text. Often the disparity between intended dot positions and actual dot positions is considerable.

Therefore before Extracting the Dot pattern Image features and matching the Dot pattern image, The image is first reconstructed if it is not fully constructed. The image is then rotated to align it with the image to be matched. The image is then enhanced.

3. DOT PATTERN FEATURE EXTRACTION

3.1 Local Binary Pattern

LBP [14] is a gray-scale texture operator which characterizes the spatial structure of the local image texture. Features of an image are extracted from LBP in the form of histogram. Given a central pixel in the image, a pattern number is computed by comparing its value with those of its neighborhoods.

$$LBP_{P,R} = \sum_{P=0}^{P-1} s(g_p - g_c \qquad (1)$$

$$s(x) = \begin{cases} 1, x \\ 0, x \end{cases} \qquad (2)$$

For example, LBP pattern 00000000 has a U value of 0 and 01000000 of 2. The uniform LBP pattern refers to the uniform appearance pattern which has limited transition or discontinuities (U<=2) in the circular binary presentation [14]. It was verified that only ''uniform'' patterns are fundamental patterns of local image texture.

3.2 Rotation Invariant Variance Measures (VAR)

A rotation invariant measure of the local variance can be defined as [14]

$$VAR_{P,R} = {}^{1}/_{P} \sum_{P=0}^{P-1} (g_P - \qquad (3)$$
$$\text{where } u = {}^{1}/_{P} \sum_{P=}^{P-}$$

3.3 LBP variance (LBPV)

$LBP_{P,R}$=$VAR_{P,R}$ is powerful because it exploits the complementary information of local spatial pattern and local contrast [14]. However VAR_P, has continuous values and it has to be quantized. This can be done by first calculating feature distributions from all training images to get a total distribution and then, to guarantee the highest quantization resolution some threshold values are computed to partition the total distribution into N bins with an equal number of entries. These threshold values are used to quantize the VAR of the test images. There are three particular limitations to this quantization procedure. First, it requires a training stage to determine the threshold value for each bin. Second, because different classes of textures may have very different contrasts, the quantization is dependent on the training samples. Last, there is an important parameter, i.e. the number of bins, to be preset. Too few bins will fail to provide enough discriminative information while too many bins may lead to sparse and unstable histograms and make the feature size too large. Although there are some rules to guide selection [14], it is hard to obtain an optimal number of bins in terms of accuracy and feature size.

The LPBV descriptor proposed in this section offers a solution to the above problems of $LBP_{P,R}$=$VAR_{P,R}$. The LBPV is a simplified but efficient joint LBP and contrast distribution method. LBPV gives as output the features of an image in the form of a histogram. We have computed the LBPV histogram by the following equations by dividing the pattern into blocks.

$$LBP_{P,R}(k) = \sum_{i=1}^{n} \sum_{j=1}^{m} w\left(LBP_{P,R}(i,j), k\right), k \in [(\qquad (4)$$

$$W\left(LBP_{P,R}(i,j),k\right) = \begin{cases} VAR_{P,R}(i,j), LBP_{P,R}(i,j) : \\ 0 \quad otherw \end{cases} \quad (5)$$

The ideal feature selection technique removes those features that are less discriminative and keeps those features that have high discriminatory power. We therefore will use Genetic Algorithm for retaining only the features that have high discriminative power.

4. DOT PATTERN MATCHING

After the optimized features from the Dot pattern images have been extracted. The dot pattern images can be matched for similarity scores. We use Euclidean distance to perform matching between two or more Dot patterns.

The Euclidean distance is the distance between two points that one would measure with a ruler, and is given by the Pythagorean formula

The Euclidean distance between points p and q is the length of the line segment connecting them

In Cartesian coordinates, if $p = (p_1, p_2, ...,$ and $q = (q_1, q_2, ...,$ are two points in Euclidean *n*-space, then the distance from p to q, or from q to p is given by:

$$d(p,q) = d(q,p) = sqrt((q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2 \quad (6)$$

5. GENETIC ALGORITHM

Genetic Algorithm (GA), first introduced by John Holland in the early seventies, is the powerful stochastic algorithm based on the principles of natural selection and natural genetics, which has been quite successfully, applied in machine learning and optimization problems. To solve a problem, a GA maintains a population of individuals (also called strings or chromosomes) and probabilistically modifies the population by some genetic, operators such as selection, crossover and mutation, with the intent of seeking a near optimal solution to the problem.

Genetic Algorithm will be used for retaining only the features that have high discriminative power and removes those features that have less discriminative power.

In order to describe GA for optimal feature selection, consider the feature vector in Fig. 4.

[30 16 24 161 21 212 …..]

Fig 4: Sample feature vector

Furthermore, consider the vector shown in Fig. 5 as a candidate real-coded feature mask.

[0.5 0.1 0.7 0.2 0.8 0……]

Fig 5: Real-coded feature mask

For GA based optimal feature selection a masking threshold value of 0.5 is used to create a binary coded candidate feature mask which will be used as condition for masking features. If the random real number generated is less than the threshold (0.5 in this case), then the value corresponding to the real generated number is set to 0 in the candidate feature mask vector or 1 otherwise. The candidate feature mask is used to mask out a feature set. Fig. 6 shows the candidate binary coded feature mask matrix obtained from the random real numbers generated in Fig. 5. The masking threshold value is applied on the real numbers to obtain the binary representation.

[1 0 1 0 1 0….]

Fig 6: The Resulting feature vector after feature masking

For GA based optimal feature selection, we can also use threshold value between 0 and 1 instead of using a static threshold value of 0.5. So each random number generated using a uniform distribution has a masking threshold value that determines whether the feature corresponding to features is masked out or not.

GA based feature weighting for weighting the features can be also be used for selecting the most discriminatory features, here the real-coded candidate feature mask is used to weight features within the feature matrix. The real-coded candidate feature mask value is multiplied by each feature value to provide a weighted feature. If the number generated is 0 (or approximately equal to 0) the feature value is 0, which basically means that the feature is masked. As given in Equation 8, the fitness returned by the evaluation function is the number of recognition errors encountered after applying the feature mask multiplied by 10 plus the percentage of features used. The selection of the parent is based on smaller fitness values because the optimization goal is to reduce the number of recognition errors (i.e. increasing the accuracy) while reducing the number of features

$$Fitness = (number\ of\ errors) * 10 + \%\ Features\ u \quad (7)$$

6. RESULTS

Dot-Pattern Image Reconstruction:

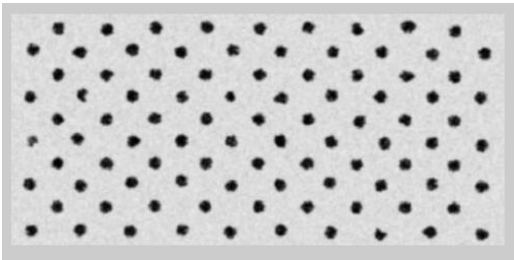

Fig 7: Distorted Dot-Pattern image

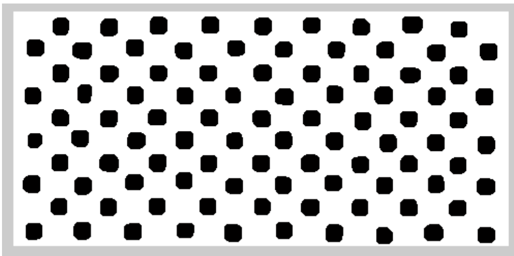

Fig 8: Re-constructed Dot-Pattern

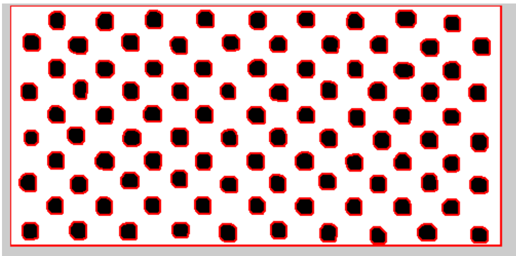

Fig9: Re-constructed Dot-Pattern with boundaries

Dot-Pattern Image Rotation:

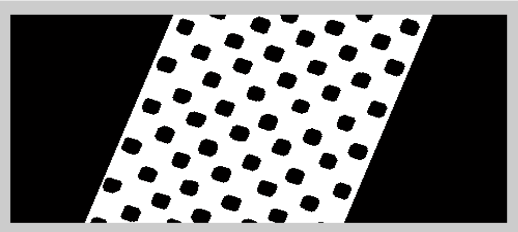

Fig 10: Rotated Dot-Pattern image

Dot-Pattern Image Enhancement:

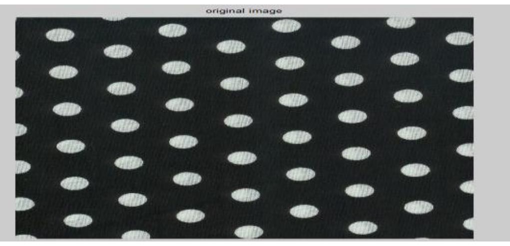

Fig 11: Original Image

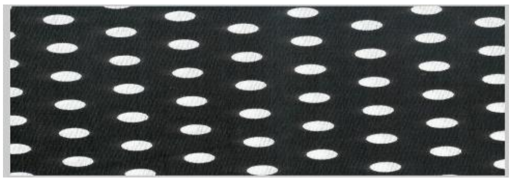

Fig 12: Enhanced Original Image

Dot Pattern Image feature Extraction using LBP based method shown in Fig. 13, Fig. 14, Fig. 15, Fig. 16.

Fig 13: Dot Pattern Image

Features of Dot Pattern image using LBP based method are first extracted in the form of histogram only as shown in Fig. 19.

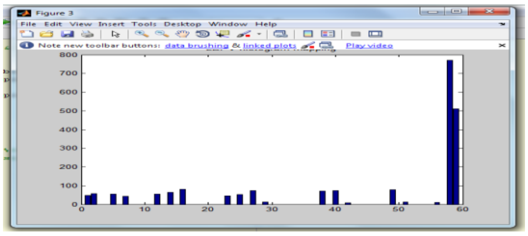

Fig 14: Histogram of extracted features of Dot Pattern image

Extracted Features of Dot-Pattern Image from histogram of the image of fig 13 :

49.04, 56.44, 0, 0, 54.04, 0, 43.03, 0, 0, 0, 0, 55.95, 0, 63.47, 0, 80.33, 0, 0, 0, 0, 0, 0, 46.39, 0, 53.30, 0, 73.59, 0, 12.66, 0, 0, 0, 0, 0, 0, 0, 0, 71.11, 0, 74.80, 0, 9.35, 0, 0, 0, 0, 0, 0, 77.88, 0, 13.23, 0, 0, 0, 0, 10.30, 0, 771.40, 512.00

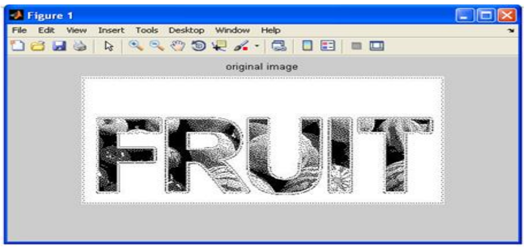

Fig 15: Dot Pattern Image

Features of Dot Pattern image using LBP based method are first extracted in the form of histogram only.

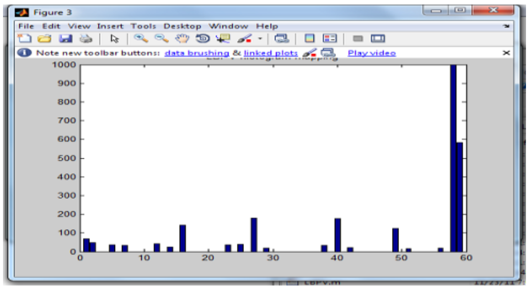

Fig 16: Histogram of extracted features of image

Extracted Features of Dot-Pattern Image from histogram of the image of fig 15:

68.66, 47.10, 0, 0, 36.49, 0, 34.96, 0, 0, 0, 0, 42.62, 0, 26.20, 0, 142.43, 0, 0, 0, 0, 0, 0, 37.73, 0, 39.13, 0, 180.08, 0, 18.92, 0, 0, 0, 0, 0, 0, 0, 0, 33.62, 0, 175.73, 0, 20.56, 0, 0, 0, 0, 0, 0, 124.14, 0, 16.43, 0, 0, 0, 0, 18.56, 0, 998.83, 582.90.

Euclidean Distance of the two images gives based on their features which are not yet optimized by GA

For First image,

Euclidean distance with itself=0

Euclidean distance with second image=46.9918

For Second image,

Euclidean distance with itself=0

Euclidean distance with First image=46.9918

7. CONCLUSION AND FUTURE WORK:

The dot pattern matching approach is important in image analysis and pattern recognition problems. The technique depends on the matching score of two dot patterns by finding an optimal transformation. To extract the dot-pattern image features this paper used LBP scheme and used Euclidean distance for finding image similarity. The features will be optimized by the application of GA and then the matching between the two Dot patterns will be performed using Euclidean distance for image similarity.

8. REFERENCES

[1] R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973.

[2] O. Faugeras, Three-Dimensional Computer Vision: A Geometric Viewpoint,The MIT Press, 1993.

[3] J.M. Chassery, A. Montanvert, Geometric Representation of Shapes and Objects for Visual Perception, Geometric

Reasoning for Perception and Action, C. Laugier (Ed.), selected papers presented in a Workshop held at Grenoble, France, September 16–17, 1991, Springer–Verlag, pp. 163–182.

[4] H. Ogawa, Labeled pattern matching by Delaunay triangulation and maximal cliques, Pattern Recognition 19 (1986) 35–40.

[5] R. Laurini, D. Thompson, Fundamentals of Spatial Information Systems, The A.P.I.C. Series, No. 37, Academic Press, London, 1992.

[6] P.J. Taylor, Quantitative Methods in Geography: An Introduction to Spatial Analysis, Houghton Mifflin Company, Boston, 1977.

[7] A. Okabe, B. Boots, K. Sugihara, Spatial tessellations: Concepts andApplications of Voronoi Diagrams, John Wiley and Sons, 1992.

[8] J. Sprinzak and M. Werman, Affine point matching, Pattern Recog. Letters, vol. 15, no. 4, pp. 337-339, 1994.

[9] L. Zhang and W. Xu, Point-pattern matching using irreducible matrix and relative invariant, Tsinghua Sci. Tech., vol. 4, no. 4, pp. 1602-1605, 1999.

[10] L. Zhang, W. Xu and C. Chang, Genetic algorithm for point pattern matching, Pattern Recog. Letters, vol. 24, pp. 9-19, 2003.

[11] J.P. Starink and E. Backer, Finding point correspondences using simulated annealing, Pattern Recogn., vol. 28, pp. 231-240, 1995.

[12] P.-Y. Yin, Particle swarm optimization for point pattern matching, J.Visual Commu. & Image, vol. 17, pp. 143-162, 2003.

[13] M. Melanie, An introduction to Genetic Algorithms, Cambridge, Massachusetts, MIT Press, 1998.

[14] T.Ojala, M.Pietik¨ ainen, T.T.M¨aenp¨a¨ a, Multiresolution gray-scale and rotation invariant texture classification with localbinarypattern, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (7) (2002) 971–987.

[15] M. Kokare, P.K. Biswas, B.N. Chatterji, Rotation-invariant texture image retrieval using rotated complex wavelet filters, IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics 36 (6) (2006) 1273–1282. [16] V. Kyrki, J.K. Kamarainen, Simple Gabor feature space for invariant object recognition, Pattern Recognition Letter 25 (3) (2004) 311–318.

[16] V. Kyrki, J.K. Kamarainen, Simple Gabor feature space for invariant object recognition, Pattern Recognition Letter 25 (3) (2004) 311–318.

[17] N.G. Kingsbury, Rotation-invariant local feature matching with complex wavelets, in: 14th European Signal Processing Conference, 2006.

[18] J. Adams, D. L. Woodard, G. Dozier, P. Miller, K. Bryant, and G. Glenn. Genetic-based type II feature extraction for periocular biometric recognition: Less is more. In Proc. Int. Conf. on Pattern Recognition, 2010. to appear.

[19] Huang C. L. and Wang C. J. “GA-based feature selection and parameters optimization for support vector machines”,. C.-L. Huang, C.-J. Wang / Expert Systems with Applications. Vol. 31(2), 2006, pp231–240.

[20] Adams, J., Woodard, D. L., Dozier, G., Miller, P., Glenn, G., Bryant, K. "GEFE: Genetic & Evolutionary Feature Extraction for Periocular- Based Biometric Recognition," Proceedings 2010 ACM Southeast Conference, April 15-17, 2010, Oxford, MS.

[21] Dozier, G., Adams, J., Woodard, D. L., Miller, P., Bryant, K. "A Comparison of Two Genetic and Evolutionary Feature Selection Strategies for Periocular-Based Biometric Recognition via XTOOLSS", Proceedings of the 2010 International Conference on Genetic and Evolutionary Methods (GEM'10: July 12-15, 2010, Las Vegas, USA).

[22] Simpson, L. , Dozier, G., Adams, J., Woodard, D. L., Dozier, G., Miller, P., Glenn, G., Bryant, K.. "GEC-Based Type-II Feature Extraction for Periocular Recognition via X-TOOLSS," Proceedings 2010 Congress on Evolutionary Computation, July 18-23, Barcelona, Spain

[23] Dozier, G., Bell, D., Barnes, L., and Bryant, K. (2009). "Refining Iris Templates via Weighted Bit Consistency", Proceedings of the 2009 Midwest Artificial Intelligence & Cognitive Science (MAICS) Conference, Fort Wayne, April 18-19, 2009.

[24] Dozier, G., Adams, J., Woodard, D. L., Miller, P., Bryant, K. "A Comparison of Two Genetic and Evolutionary Feature Selection Strategies for Periocular-Based Biometric Recognition via XTOOLSS", (to appear in) The Proceedings of the 2010 International Conference on Genetic and Evolutionary Methods (GEM'10: July 12- 15, 2010, Las Vegas, USA).

[25] Tamirat Abegaz, Gerry Dozier, Kelvin Bryant Joshua Adams, Khary Popplewell, Joseph Shelton, ,Karl Ricanek, Damon L. Woodard” “Hybrid GAs for Eigen-Based Facial Recognition”, accepted for IEEE Symposium Series in Computational Intelligence 2011 (SSCI 2011)