World Journal of Engineering Research and Technology
WJERT
www.wjert.org
SJIF Impact Factor: 5.218

PERFORMANCE EVALUATION OF SHAPE-SPECIFIC AND GENERAL DEFINITION OF MODELS USING DYNAMIC DEFORMABLE TEMPLATE FOR IMAGE-BASED OBJECT RECOGNITION

Bilkis Jamal Ferdosi*[1] and Muhammad Masud Tarek[2]

[1]Associate Professor, Department of CSE, University of Asia Pacific, Dhaka, Bangladesh.

[2]Associate Professor, Department of CSE, State University of Bangladesh, Dhaka, Bangladesh.

Article Received on 22/02/2018 **Article Revised on 15/03/2018** **Article Accepted on 05/04/2018**

***Corresponding Author**
Bilkis Jamal Ferdosi
Associate Professor,
Department of CSE
University of Asia Pacific,
Dhaka, Bangladesh.

ABSTRACT

Model-based segmentation and recognition can be invaluable if a prior knowledge about the object's shape is available. It enables a model description that may even override data information in case of inaccurate and missing information. However, achieving proper model description and its in-class variation is not a trivial task. In literature there exist several approaches. Some require enough representative samples, some fails in case of large size and shape variation etc. In our work, we propose a template combining deformable templates and physically based models together. Using the proposed template we study i) whether object's shape-only-template is sufficient or information about the space surrounding the object should be incorporated, iii) whether there exist a useful quality measure for matching templates. We found that the proposed dynamic deformable template works successfully. Incorporation of the surrounding space information in the template significantly improves the recognition rate which also enables a better template matching criteria.

KEYWORDS: *Segmentation; recognition; deformable template; physically based model.*

I. INTRODUCTION

To extract and recognize an object from an image of medical data or similar application areas where shape and structure of the region of interest such as organs, bones, tissues etc. are available or known a priori, efficient use of model-based image segmentation and recognition is very common. However, in such case, main obstacles to overcome is to obtain an efficient model description and its shape variation in the object class.

In free-form models, no global structure is used where the model is constrained only by local continuity and smoothness constraints.[1–3] On the other hand, in,[4–7] contour or skeleton of the object is utilized to obtain the model description. To obtain the shape and shape variation of a class different approaches are used, such as in[4] statistical distribution of shape and its variation is learned from training on enough representative samples. However, the success of the approach severely depends on the availability of enough samples that truly represent the object variation.

Along with the shape description, it is also essential to incorporate the in-class shape variation in the model. Deformable templates that incorporate prior geometric shape information and shape variation in terms of non-parameterized bitmap images,[8] or parameterized curves,[9–11] can be useful. In the first type, separate deformation mechanism is needed and in the second type, setting a large number of parameters is not trivial.

Another solution can be the use of physically based deformable models[5,12] that also incorporate prior shape information in the model. The model deformation is achieved by changing the physical properties of the model itself. Being intuitive in nature, physically based models provide the advantage of not requiring any prior training and incorporation of the shape variation into model parameters. Therefore, combining a template-based and a physically-based strategy seems to be reasonable. A template could be modeled using a physically based model (PBM) and utilize the inherent quality of PBM for model deformation instead of using a separate deformation mechanism.

Therefore, in this paper, we propose a Dynamic Deformable Template (DDT) using a Stable Mass-Spring Model (SMSM). DDT is a template whose deformation is governed by its physical properties. Unlike the state of the art approaches, our approach incorporates object's neighborhood information as a part of shape information. Our model description also enables better measure for template matching. We use two different strategies for generating DDTs.

- **Handcrafted DDTs** are specific to the shape of the object class and vary among the classes. A different object class is represented by a different number of masses and springs. Thus, the inter-class comparison is possible only in terms of global energy (total energy experienced by all of the masses of a class model).
- **Rectangular grid DDTs** are structurally equal for all the shape classes and capable of sampling space around the object. Object-specific shape information is introduced into the templates as model forces. In this model inter-class comparison is possible both in terms of global energy and local energy (energy experienced by mass, ***m*** of the model, ***x*** can be compared with mass, ***m*** of the model, ***y***).

We also experimented with two different matching quality measures: **global** and **local** measures.

- The first one measures the quality of global fitting of the templates. It compares total energy required for a template of class, ***x*** and a template for class, ***y*** to fit in to object, ***m.*** If the class, ***x*** template requires less energy to fit into an object, ***m***, then the object will be classified as a class, ***x,*** and vice versa. It can be applied to both types of templates.
- The second one measures the quality of fitting of the templates at the local level. In this measure, we compare two templates in terms of energy experienced by corresponding mass points of the templates. It can only be applied to the rectangular grid DDTs as it requires a comparison between corresponding parts of templates.

We investigate the capability of proposed DDTs of capturing shape variations using handwritten digits because of their high shape variation.

II. THE STABLE MASS-SPRING MODEL (SMSM)

The mass-spring model is a geometric entity that receives its shape from physical constraints. It is usually represented by a graph where nodes are mass points and edges are springs connecting two mass points. Using a prior knowledge of the object shape, mass positions and springs rest lengths are estimated. The model dynamics are governed by internal forces, which are defined within the model graph, and external forces that are computed from the image data. The model deforms only under external force. The external force acts as some kind of attraction force that attracts the model towards the features in the image. The magnitude of the attraction force increases with decreasing distance to the attraction point. The model deformation continues until internal and external force balance.

Consider a mass-spring model consisting of mass points with masses mi and initial position vector, $\boldsymbol{p_i}$ and connected with springs of spring constant, $\boldsymbol{k_{ij}}$. The rest length, $\boldsymbol{Rl_{ij}}$ that are spatially associated with two mass points ***i*** and ***j*** can be calculated using equation 1.

$$Rl_{ij} = p_i - p_j \qquad (1)$$

The dynamics of such a system can be described by Newtonian mechanics. The movement of the model is influenced by the forces acting on the mass points.

A. Internal Forces

1) Spring Force: The elastically deformed springs exert spring forces, $\vec{F}_{S_i}$ on the mass points as they deviate from their rest lengths i.e. mass points gets new position vector, p_{new_i}. The spring force, $\vec{F}_{S_i}$ acting on mass point ***i*** can be calculated using equation 2.

$$\vec{F}_{S_i} = \sum_j \frac{k_{ij}.(\left\| p_{new_i} - p_{new_j} \right\| - Rl_{ij}).(p_{new_i} - p_{new_j})}{\left\| p_{new_i} - p_{new_j} \right\|} \qquad (2)$$

2) Torsion Force: For model stability, another internal force called torsion force is included in the model dynamics.[12] To introduce torsion force in the model, rest directions $\boldsymbol{rd_{ij}}$ of the springs starting from one mass point ***i*** to all adjacent mass point ***j*** along the lines of the spring rest length is calculated.

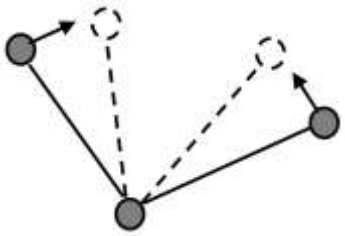


Figure 1: Torsion Force: rest directions of the springs are shown in dotted lines; springs deviated from their rest direction under the influence of the external force (shown in solid line).Thus the torsion force is exerted on the nodes in the direction of their rest direction.

Springs deviations from their rest direction exert a torsion force to the mass points in the direction of their rest directions to restore the original position. The torsion force, $\vec{F}_{T_i}$, exerts on mass point ***i*** can be calculated using equation 3.

$$\vec{F}_{T_i} = \sum_j \frac{t_j \cdot \left| \angle(p_{ji}, r_{d_{ji}}) \right|}{\left\| p_{ji} \right\|} \cdot \frac{n_{ji}}{\left\| n_{ji} \right\|} \qquad (3)$$

Where, $p_{ji} = p_{new_i} - p_{new_j}$ and t_j = torsion constant of the mass point j

The working direction of the torsion forces, n_{ji} is calculated using equation 4.

$$n_{ji} = r_{d_{ji}} - \frac{\left\langle p_{ji}, r_{d_{ji}} \right\rangle \cdot p_{ji}}{\left\| p_{ji} \right\|^2} \quad \text{for all } j \qquad (4)$$

B. External Force

Image-Force The model is constrained to lie on the image allowing it to deform according to the image topography. Image feature collected by each mass point act as a force on that mass and the image force, $\vec{F}_{I_i}$ exerted on mass point ***i*** can be calculated using equation 5.[7]

$$\vec{F}_{I_i} = image(loc(i)) \qquad (5)$$

Where ***loc(i)*** gives the coordinates of the current location of the sensor ***i*** in the image and ***image( )*** gives the image information at that position.

1) ***Damping Force:*** Damping force is used to prevent oscillatory behavior of the model. It can be defined as a function of velocity as in[5] or the speed of motion of the model can be damped by a factor, ***d*** as in.[12]

2) Discrete Equations of Motion

Given the forces discussed above, applying Newton's second law of motion can be simulated in a discrete time step of distance Δt . According to Newton's second law of motion, acceleration a_i of mass point ***i*** can be calculated using equation 6.

$$a_i = \frac{\vec{F}_i}{m_i} \qquad (6)$$

Where $\vec{F}_i$ is a total force acting on mass point ***i*** and can be calculated using equation 7.

$$\vec{F}_i = w_{S_i}.\vec{F}_{S_i} + w_{T_i}.\vec{F}_{T_i} + w_{I_i}.\vec{F}_{I_i} \quad (7)$$

Where,

w_{S_i} = weight for the spring force acting on mass point ***i***,

w_{T_i} = weight for the torsion force acting on mass point ***i***,

w_{I_i} = weight for the image force acting on mass point ***i***.

Following the calculation of the mass point forces we can compute the new velocity, $\boldsymbol{v_i}$ and position, $\boldsymbol{p_i}$ of each mass point given the old velocity, v_i^{old} and old position values, p_i^{old} as follows:

$$v_i = (v_i^{old} + a_i.\Delta t).(1-d) \quad (8)$$

$$p_i = p_i^{old} + v_t.\Delta t \quad (9)$$

The model adaptation to the image will continue until equilibrium when external forces cancel out internal forces i.e., $\vec{F}_{\text{int}ernal} = \vec{F}_{external}$.

III. APPLICATION OF SMSM AS DDT

A. Handcrafted DDTs (hcDDTs)

The hcDDTs are generated by placing the mass points along the medial axis of the digit as we are interested in the actual shape of the digits. In some of the hcDDTs, for example in the templates for digit 0, 6, 8 or 9, we include some mass points that have dark intensity sensors associated with them for a better description of their shape (see fig. 2a).

B. Rectangular Grid DDTs (rgDDTs)

This template has similarity with the 2D planar rubber sheet used by Jain et al.[6] for deformation transformation of the bit map image templates. However, in Jain's approach, a parametric statistical mapping is needed to generate random variations in the template. In rectangular grid DDT deformation is guided by forces, which can be determined by features in the image to be segmented and by geometric constraints within the model. Therefore, statistical mapping of template deformation can be avoided in our proposed rectangular template. The template used by Jain et al. is a bit map image of the object to be found. On the other hand, the template used in this paper is based on a mass-spring model that can be considered as a discretization of a bit map image.

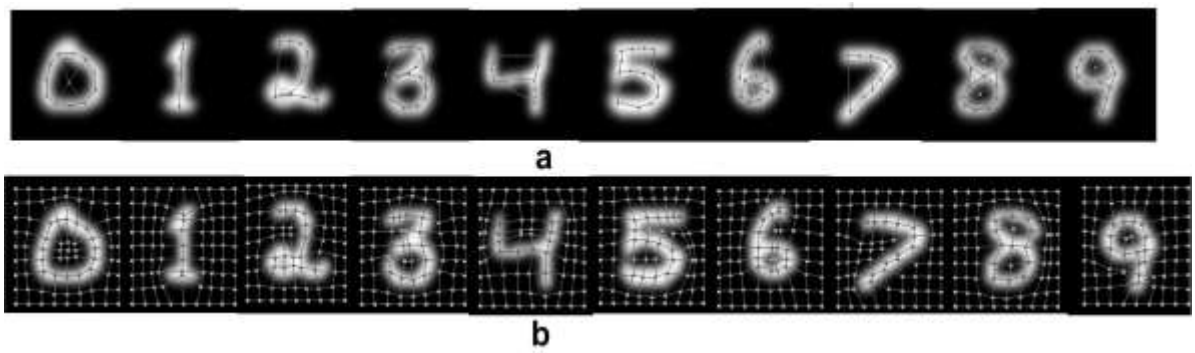


Fig. 2: (a) Handcrafted DDTs (b) Rectangular grid DDTs. Masses in grey are associated with bright intensity sensors, masses in white are associated with dark intensitv sensors.

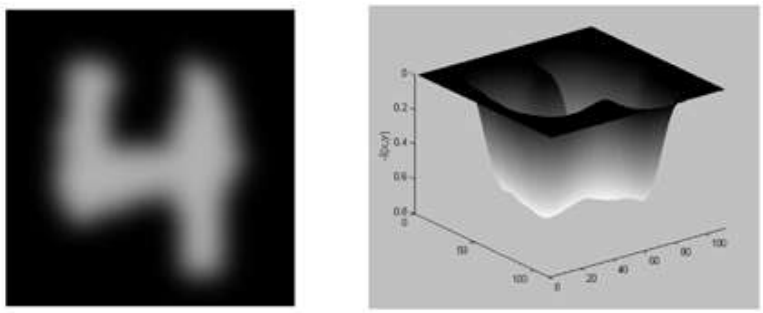
IV. MNIST DATABASE

All of our investigations related to DDTs are based on the handwritten digits from the MNIST database.[17] The digits are size-normalized and centered in a 28x28 grey level image. We are interested in the template's ability to catch essential aspects of a digit's shape and not so much in improving the quite impressive recognition rates of trained classification on all the samples. Therefore, in our experiments we took a subset of 18 samples for each digit for training, considering their morphological similarity in shape, so that the variation among the samples can be learned with a single template. For the test data set, we selected 850 digit images from the first 1000 images of MNIST test data set. In our work, we super-sampled the image into 112x112 pixels with the resolution of 600 dpi using the nearest neighbor method.

In order to allow the deformation process some room for adjusting the template, a 32-pixel-wide border was placed around each image. This increased the actual image size to 144x144 pixels.

V. SENSORS

Each mass point in the model can be made sensitive to different sensors. Sensors are defined application-specific and may be sensitive, e.g., to corners, edges, or specific intensities. In our work we used (dark/bright) intensity sensor convolved with distance kernel (fig. 3).



a b

Figure 3: Image seen as 3D surface. In (a) is shown a convoluted image of intensities of digit 4. In (b) is shown its representation as a 3D surface. This representation corresponds to the graph of the function -I(x,y).

VI. TEMPLATE MATCHING MEASURES FOR CLASSIFICATION

A. Global Energy Measure (GEM)

In GEM classification, we use the adapted model energy normalized by the expected model energy, as the decision measure. The expected model energy, $E_{expected}$ of template i is obtained from the training data according to the equation 10, where E_k is the energy of the template on the k^{th} sample.

$$E_{\text{expected}}(i) = \frac{1}{NumberOfSamples} \sum_{k=1}^{NumberOfSampes} E_k \qquad (10)$$

and

$$E_k = \left\| \vec{F}_k \right\| \qquad (11)$$

All the 10 templates for 10 digits are applied to the image of the digit to be classified one after another. After adaptation, the rest energy, $E_{adapted}$ of template i is normalized by its expected energy, $E_{expected}$ that yields global decision measure, $E_{decision}$ (as in equation 12).

$$E_{decision}(i) = \frac{E_{adapted}(i)}{E_{\text{expected}}(i)} \qquad (12)$$

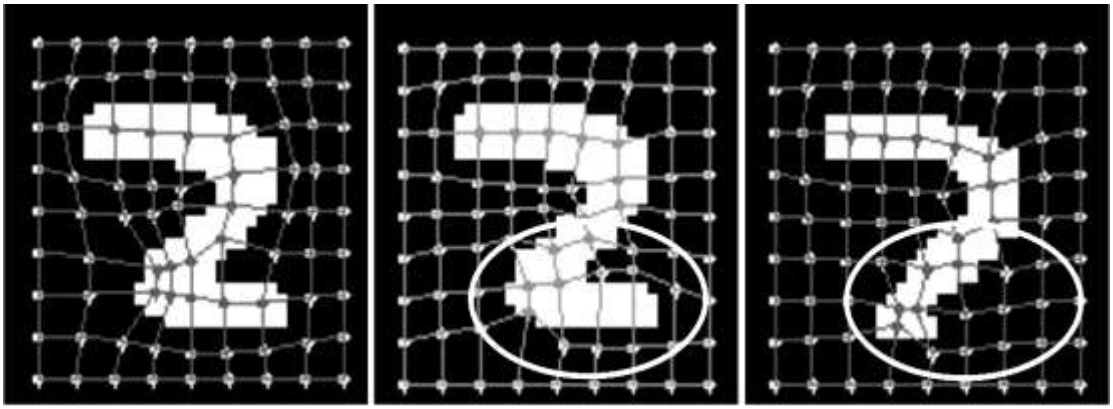


Figure 4: (a) Digit 2 is adapted with the rectangular grid DDT for 2 (b) Digit 2 is adapted with the template for 7 (c) Digit 7 is adapted with the template for 7.

It is likely that the template on wrong digit image will have higher $E_{decision}$ measure than the template on the correct digit. Therefore, the template that will have the lowest $E_{decision}$ measure should be declared as a winner (equation 13).

$$Winner = \arg\min(E_{decision}(i)) \quad (13)$$

B. Local Energy Measure (LEM)

LEM bears some similarity to correlation measures used in rigid template matching and requires templates for every class that have the same number of elements (springs and masses). LEM is calculated using equation 14.

$$E_{correlation}(i) = \sum_{D,S,I} \sum_{j=1}^{NumberOfNodes} (E_{adapted}(j) - E_{expected}(j)) \quad (14)$$

VII. RESULTS AND DISCUSSION

A. hcDDTs vs. rgDDTs using Global Measure

Using global measures hcDDTs achieved 69:44% success rate on the training set and 31.84% success rate on the test data set. On the other hand, rgDDT achieved 72.78% success rate on training data which is 3.34% improvement over hcDDTs. On test dataset rgDDT achieved 49.82% success rate which is 17.98% improvement compared to the performance of hcDDTs. It indicates the capability of rgDDT in generalizing training results to unknown samples of the same type.

B. hcDDTs vs. rgDDTs using Local Measure

In hcDDT different models are represented by different numbers of masses and springs, thus it is not possible to use local measures for hcDDT. Using correlation measure for classification, rgDDTs gained 96.67% success rate on the training data set which is 23.89% improvement if compared to its own performance using global measure and 27.23% improvement if compared to the performance of hcDDTs. For the test data set rgDDT obtained 74.69% success rate using a local measure which is 24.87% improvement compared to global measure, and 42.85% if compared to hcDDTs.

The success of rgDDT using local measure may lie in the fact that the rgDDT incorporates more information about the object and it enables comparison of the templates at the local level. For example, in fig. 4a digit 2 is adapted with the template for digit 2 that results in global decision measure $\boldsymbol{E_{decision}}$ (2) = 1.13415 and in fig. 4b digit 2 is adapted with the template for digit 7 that results in global normalized $\boldsymbol{E_{decision}}$ (7) = 1.09002. Thus the digit 2 is wrongly classified as digit 7. If the local node by node energy measure is considered then it would not be the case. As indicated in figure 4b the nodes in the region encircled have higher variation than the region in fig. 4c.

However, the performance of such models severely depends on the initial placement. If the model is not placed sufficiently close to the target, finding shapes cannot be ensured. In most of the cases, user interaction and expert knowledge are needed for the placement of the model.

VIII. CONCLUSION

Deformable models are used in different application areas such medical imaging[15,16] where achieving proper model definition is crucial. In our experiments, we used handwritten digits of MNIST database for investigating the potential of DDTs. The success rate of using match quality for classification is far from results of state-of-the-art statistical methods for classifying digits, where error rates of less than 1% can be obtained. Jain et al.[13] achieved similar results using deformable templates on a 2000 character training set. Using syntactic pattern recognition Lopez et al.[14] obtained 83.83% recognition rate on a 300 character training set. On the other hand, using our rgDDTs we achieved 96.67% recognition rate using only 18 character training set. The power of DDTs is their ability to provide a meaningful match quality based on very few samples and little apriori information by directly incorporating shape information into representation and matching measure. The proposed DDTs has huge potential in the application areas where large training data sets are not available. The results

obtained from our experiments are sufficient to support our hypothesis of "better representation of a template that incorporates additional scope for template matching is necessary to improve the object recognition". Therefore, our proposed DDTs qualify for further investigation to find out more about its use for recognizing objects. In future, we will investigate its potential for general object recognition.

REFERENCES

1. Kass, M., Witkin, A., Terzopoulos, D.: Snakes (1988): Active contour models. INTERNATIONAL JOURNAL OF COMPUTER VISION, 1(4): 321–331.
2. Cohen, L.D. (1991): On active contour models and balloons. CVGIP: Image Underst, 53(2): 211–218.
3. Xu, C., Prince, J.L. (1998): Snakes, shapes, and gradient vector flow. IEEE TRANSACTIONS ON IMAGE PROCESSING, 7(3): 359–369.
4. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J. (1995): Active shape models their training and application. Comput. Vis. Image Underst, 61(1): 38–59.
5. Hamarneh, G., McInerney, T. (2003): Physics-based shape deformations for medical image analysis. In: SPIE Electronic Imaging, 5014: 354–362.
6. Jain, A.K., Zhong, Y., Dubuisson-Jolly, M.P. (1998): Deformable template models: A review. Signal Processing, 71(2): 109–129.
7. T¨oennies, K.D., Benedix, P. (2005): A weight-adaptive dynamic model for shape segmentation. In: IEEE International Conference on Image Processing ICIP. (11.-14. September 2005).
8. Jain, A.K., Zhong, Y., Lakshmanan, S. (1996): Object matching using deformable templates. IEEE Trans. Pattern Anal. Mach. Intell, 18(3): 267–278.
9. Yuille, A., Hallinan, P., Cohen, D. (1992): Feature extraction from faces using deformable templates. International Journal of Computer Vision, 8: 99–111.
10. Li, C., Xu, C., Gui, C., Fox, M.D. (2010): Distance regularized level set evolution and its application to image segmentation. Trans. Img. Proc., 19(12): 3243–3254.
11. Wang, Q., Boyer, K.L. (2012): The active geometric shape model: A new robust deformable shape model and its applications. Comput. Vis. Image Underst, 116(12): 1178–1194.
12. Dornheim, L., Tonnies, K.D., Dornheim, J. (2005): Stable dynamic 3d shape models. In: In ICIP05: International Conference on Image Processing.

13. Jain, A.K., Zongker, D. (1997): Representation and recognition of handwritten digits using deformable templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19: 1386–1391.
14. L´opez, D., Pinaga, I. (2000): Syntactic pattern recognition by error correcting analysis on tree automata. In Ferri, F.J., Quereda, J.M.I., Amin, A., Pudil, P., eds.: Advances in Pattern Recognition, Joint IAPR International Workshops SSPR 2000 and SPR 2000, Alicante. Volume 1876 of Lecture Notes in Computer Science, Springer, 133–142.
15. Camara, M., Mayer, E., Darzi, A. et al. Int J CARS, 2016; 11: 919.
16. Yanni Zou, Peter X. Liu, A. (2017): high-resolution model for soft tissue deformation based on point primitives, Computer Methods and Programs in Biomedicine, 148; Pages 113-121, ISSN 0169-2607.
17. http://yann.lecun.com/exdb/mnist/