Global Structure Constraint Model for Object Representation

Miao Liu and Dongwei Guo and Jie Ma and Yanbin Liu and Chunguang Zhou
Colledge of Computer Science and Technology, Jilin University, Changchun, China 130012
Email: cgzhou@jlu.edu.cn

***Abstract*— We present a novel object representation model based on the Global Structure Constraint (GSC). In our approach, the object is described as a constellation of points which are placed at all the representative patches with small color variations. The color information and spatial relations of these patches are reserved by these points to build the color model and shape model. To demonstrate the ability of its representation, we use it together with the searching algorithm to locate target objects in images. The experimental results demonstrate that it is a simple, effective and efficient representation.**

I. Introduction

Given digital images, choosing an appropriate image representation will greatly facilitate the consequently image processing methods, such as image localization, recognition and retrieval systems. There is broad agreement upon the suitability of representing objects as collections of local parts and their mutual spatial relations. Restricting the description to local parts of the image leads to higher robustness against clutter and partial occlusion than traditional global representations (for example, global appearance projected by the principal components analysis (PCA)), whereas incorporating spatial relations between parts adds significant distinctiveness to the representation. The common approach is to employ graph-based representations that deal separately with local parts (nodes of the graph) and spatial relations (arcs of the graph) [1], [2], [3], [4], [5]. The problem with this approach is the complexity of considering both local properties and spatial relations described separately in nodes and arcs, where learning and matching graph representations are known to be very expensive.

As for the contents presented by images, two clearly separated domains are the socalled ”things” and ”stuff” domains[6], [7], which correspond respectively to entities with discriminant geometry (object) and texture-rich materials with loose geometry (for example, natural landscapes such as forest, river, dessert, and so forth). Entities in these two domains seem to be represented in a fundamentally different way in human vision[7], suggesting the need for different models representing each of them[1].

In this paper, in the GSC model the object is represented as several landmark points which locate in the main patches with small color variations. The color information and spatial relations of these points are reserved and used to build two models, shape and color. The dimension of GSC model can be very low, such as only 26 points for the human face. To demonstrate the ability of this model, we adopt Genetic Algorithm (GA)[8], [9] to optimize the transformation parameters for the GSC model to locate the target.

II. Global Structure Constraint Model

A. Sample

By digital image capture device, the real-world scenarios can be discretized to digital images with some specific resolution, which can be processed by computers. Digital images with large resolution can be compressed to low resolution. These two kinds of processes both can be considered as sampling and compressing representation procedures, which cannot ruin the understanding to the image contents by human brain as long as the result resolutions are not too low. Here, we give an analogously sampling and compressing representation of images.

B. Patches and Landmarks

The patches with small texture variations can be segmented by statistical means. However, in practice we can line out the functional areas of the target object and then mark several key points on each area to depict its main color structure.

C. Statistical Model of GSC

For a 2D image, we represent the n landmark points $\{p_1, \ldots, p_n\}$ for a single example as two vectors, which are the shape vector $\mathbf{x} = (x_1, \ldots, x_n, y_1, \ldots, y_n)^T$ and the color vector $\mathbf{c} = (c_1, \ldots, c_n)^T$. c_i is the mean or representative color value of the neighbor region of point $p_i = (x_i, y_i)$. Given the training set of s examples, we can label landmarks and generate the learning vectors $\mathbf{x}_j$ and $\mathbf{c}_j$. Before analyzing statistically, we transform the shape vectors into a model coordinates in which the shapes of objects are normally considered to be independent of their original position, orientation and scale.

Suppose now we have sets of points $\mathbf{x}_j$ which have been transformed into the model coordinate system. We can model the shape model of GSC, $\mathbf{S}_{GSC}$, as the mean of training set simply.

$$\mathbf{S}_{GSC} = \frac{\sum_{j=1}^{s} \mathbf{x}_j}{s} \tag{1}$$

However, the color model of GSC,$\mathbf{C}_{GSC}$, at landmarks should be modeled representatively to the target class and independently to the imaging environment.

$$\mathbf{C}_{GSC} = \frac{\sum_{j=1}^{s} \left((\mathbf{c}_j - [\Downarrow_n MinC_j])^T \times (SpanRatio_j \times \mathbf{I}) \right)}{s} \tag{2}$$

where $MinC_j = \min_{i=1}^{n}(c_{j,i})$ and $\Downarrow_n$ concatenate copies of its operand vertically. $SpanRatio_j = \frac{SupC}{\max_{i=1}^{n}(c_{j,i}) - \min_{i=1}^{n}(c_{j,i}) + 1}$ with $SupC$ is the superior limit of image color value.

III. Searching with GSC Model

A. GSC Model Parameters

Having GSC models, what to do is to find out the optimal transformation for fitting GSC model to the target image well. To deal with rigid transformation, function T_{stat} with parameter vector $stat = (T_x, T_y, S_x, S_y, \theta)^T$ designates the Euclidean transformation defining the position, (T_x, T_y), the scale, (S_x, S_y), and orientation, θ, from the model coordinate system to the image coordinate system. For in-stance, if applied to a single point (x, y),

$$T_{stat}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} T_x \\ T_y \end{pmatrix} + \begin{pmatrix} S_x\cos\theta & -S_y\sin\theta \\ S_x\sin\theta & S_y\cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{3}$$

B. Searching Algorithms

Many searching algorithm can be adopted. The key things about GA [8], [9], which is exploited here as searching algorithm, are described as follow.

1) Chromosome Coding: Chromosomes are represented by bit strings. The way to represent T_x, T_y, S_x and S_y are to use a bit string of length eight, and θ is coded as a twelve-bit string. The transformation $T_{C\to I}$ from chromosomes of bit string $C$to their interpretation I and the inverse transformation $T_{I\to C}$ are formulated as,

$$I = T_{C\to I}(C) = \frac{[C]_D}{2^{len} - 1} \times (I_{sup} - I_{inf}) + I_{inf} \tag{4}$$

$$C = T_{I\to C}(I) = \left[\frac{I - I_{inf}}{I_{sup} - I_{inf}} \times \left(2^{len} - 1\right)\right] \tag{5}$$

where len is the length of bit string, I_{sup} and I_{inf} are the superior and inferior limits of the interpretation, $[\ \ ]_D$ and $[\ \ ]_B$ mean to get the decimal and binary value.

2) Fitness Function: The fitness function is defined as followed.

$$fitness = \frac{1}{1 + A} \tag{6}$$

with

$$A = \sqrt{\frac{1}{n}\left(\left(\sum_{i=1}^{n-n_{NII}} \left(\mathbf{M}(i) - \mathbf{C}(i)\right)^2\right) + k \times n_{NII}\right)}$$

where M is the color model of GSC and C is the color information of GSC, after being transformed to the image coordinates with transformation parameters using (3), in target image. There are also some processes for C before computing $fitness$ with (6), such as the color information must be uniformed to the range of M. n is the total number of the landmarks of GSC model and n_{NII} is the number of GSC landmarks not in the target image after transformation. k is a constant.

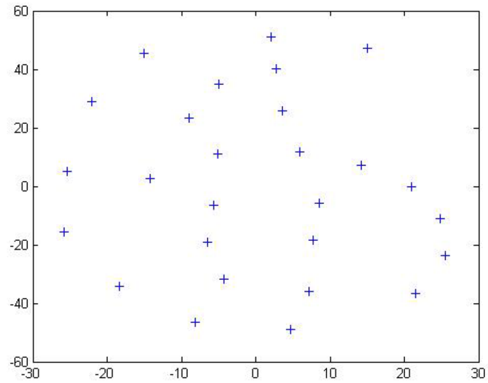


Fig. 1. 27 Landmark points of GSC shape model for number '6' in the model coordinates

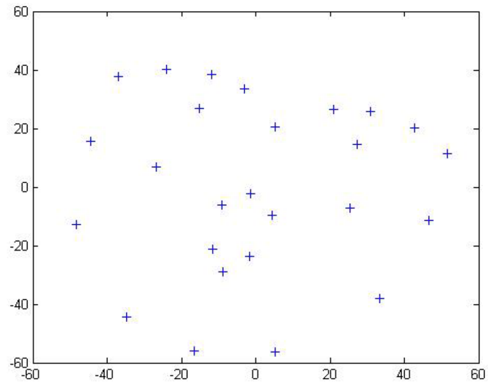


Fig. 2. 26 Landmark points of GSC shape model for human face in the model coordinates

IV. Experimental Results

First we take a simple digits test on number 6 and 9 because these two numbers are the same if the rotation is not considered. Fig. 1 is the shape of GSC model for the digit, number 6. Table I are results of the 10^{th},20^{th},30^{th} and 40^{th} iterations, the numbers under each image is the transformation parameters. Table II lists the final results for different kinds of '6' and '9'. Results describe that GSC can find out the proper parameters under the variation of scale and rotation.

To obtain a quantitative evaluation of the performance of the representation we trained GSC for human face on 20 hand labeled face images, from FERET database [10] in which the resolution of image is 256×384, and tested it on a different set of 100 labeled images with the population size 50 and iteration times 100. The shape model of GSC for human face is shown in Fig. 2. In Table III, the first two rows D_x and D_y means the average differences between the searching results S

TABLE I

TEST RESULTS ON NUMBER '6'

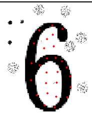
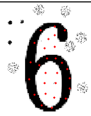
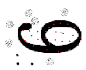
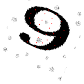
	iterations=10	iterations=20	iterations=30	iterations=40
Image				
T_x	116.813	120.268	105.065	104.374
T_y	131.435	138.496	133.452	134.461
S_x	1.375	1.506	1.686	1.686
S_y	1.421	1.504	1.594	1.649
θ	-0.963	-0.932	-1.570	-1.509

TABLE II

TEST RESULTS ON DIFFERENT KINDS OF NUMBER '6' AND '9'

Image				
T_x	150.156	139.504	165.529	220.103
T_y	140.556	96.081	169.255	191.547
S_x	1.756	1.657	1.787	1.750
S_y	3.258	1.654	1.524	3.256
θ	1.923	-2.982	-0.740	1.958

and the labeled landmarks L on x and y axis respectively. The third row gives the result $fitness$ values. The columns mean, minimum and maximum means the statistical results for the set of the 100 images. In these 100 test images, the algorithm can find the proper displacement status for 89 images and the result is not satisfied for the other 11 images yet. Table IV shows some result images after 25, 50 and 75 iterations respectively.

V. DISCUSSION AND CONCLUSION

We have demonstrated that the Global Structure Constraint model has the ability to represent target objects. The advantages of this kind representation are as followed. And it has several advantages such as that the representation is simple, the computational resumption is decreased and can deal with translation, scale and rotation intrinsically without other strategies, such as multi-resolution and model rotation operations. Besides, GSC model has other assistant functions.

TABLE III

THE STATISTICAL RESULTS ON THE SET OF 100 LABELED IMAGES

Unit: Pixels	mean	minimum	maximum
$D_x = \frac{1}{n}\sum_{i=1}^{n} \lvert S_{i,x} - I_{i,x} \rvert$	7.518	1.956	35.986
$D_y = \frac{1}{n}\sum_{i=1}^{n} \lvert S_{i,y} - I_{i,y} \rvert$	15.313	2.594	92.668
$fitness$	0.0182	0.0009	0.0574

TABLE IV

SOME TEST RESULT IMAGES OF HUMAN FACE

Iter. 25	Iter. 50	Iter. 75
 (136.240, 214.963, 1.208, 1.378, -2.844)	 (136.240, 226.338, 1.227, 1.617, -2.910)	 (133.480, 216.385, 1.347, 1.336, -2.898)
 (122.440, 131.066, 1.326, 1.435, 3.243)	 (125.200, 133.910, 1.326, 1.580, 3.312)	 (138.188, 184.894, 1.211, 1.035, 3.467)

However, the GSC model has its disadvantages which we are attempting to overcome.

ACKNOWLEDGMENT

The authors would like to thank the support from the National Natural Science Foundation of China under Grant No. 60433020, 60673099 and 60773095.

REFERENCES

[1] Jaume Amores, Nicu Sebe and Petia Radeva: *Context-Based Object-Class Recognition and Retrieval by Generalized Correlograms*. IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 29, no. 10, pp. 1818-1832 (2007)

[2] D. Crandall, P. Felzenszwalb, and D. Huttenlocher: *Spatial Priors for Part-Based Recognition Using Statistical Models*. Proc. IEEE International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 10-17 (2005)

[3] P. Hong and T.S. Huang:*Spatial Pattern Discovery by Learning a Probabilistic Parametric Model from Multiple Attributed Relational Graphs*. Journal of Discrete Applied Math, vol. 139, nos. 1-3, pp. 113-135 (2003)

[4] R. Fergus, P. Perona, and A. Zisserman:*Object Class Recognition by Unsupervised Scale-Invariant Learning*. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 264-271 (2003)

[5] M. Weber, M. Welling, and P. Perona:*Towards Automatic Discovery of Object Categories*. Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, pp. 101-108 (2000)

[6] C. Carson, S. Belongie, H. Greenspan, and J. Malik:*Blobworld: Image Segmen-tation Using Expectation-Maximization and Its Application to Image Querying*. IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 8, pp. 1026-1038 (2002)

[7] A. Pentland, R.W. Picard, and S. Sclaroff:*Photobook - Content-Based Manipulation of Image Databases*. International Journal on Computer Vision, vol. 18, no. 3, pp. 233-254 (1996)

[8] Tom M. Mitchell:*Machine Learning*. McGraw-Hill Education (Asia) Co. and China Machine Press (2003)

[9] Dongwei Guo, Chunguang Zhou and Dayou Liu:*A Niching Genetic Algorithm Based on Clustering and Resource Competition*. Journal of Computer Research and Development, vol. 40, pp. 1424-1430 (2003)
[10] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss:*The FERET database and evaluation procedure for face recognition algorithms*. Image and Vision Computing Journal, Vol. 16, No. 5, pp. 295-306 (1998)

Miao Liu received the MSc degree in Computer Science in 2003 from Jilin University, China. Now she is a PhD candidate and her research interests include biometrics and image processing.

Dongwei Guo received his PhD degree from Jilin University in 2002. Now he is an associate professor in the College of Computer Science and Technology, Jilin University. His research interests include computational intelligence and networking security.

Jie Ma is a PhD candidate in the College of Computer Science and Technology, Jilin University. His research interests include pattern recognition and image processing.

Yanbin Liu is a PhD candidate of Artificial Intelligence. He has published several papers. The current research regions include multi-agent system and optimization algorithm.

Chunguang Zhou is a professor in the College of Computer Science and Technology , Jilin University. His research interests include the computational intelligence theory and its application in image processing, business intelligence, bioinformatic, biometrics, machine olfaction and gustation. He has published more than 100 papers in the areas of his interests.