MVA2005 IAPR Conference on Machine VIsion Applications, May 16-18, 2005 Tsukuba Science City, Japan

8-7

Add Cartesian Differential Invariants to Minimum Description Length Shape Models

Shaoyu Wang Feihu Qi Huaqing Li
Department of Computer Science and Engineering, Shanghai Jiao Tong University, China
Room 1003a, 28#, 222 lane, Panyu Road, Shanghai 200052, China
wang@sjtu.edu.cn fhqi@cs.sjtu.edu.cn waking_lee@sjtu.edu.cn

Abstract

The minimum description length approach can automatic solve the point correspondence problem and give the better statistical shape models than those built by hand or equally spaced way. The current mdl approaches build the models only based on the segmented shapes without considering the local image structure and may place the markers at wrong places. This paper adds Cartesian differential invariants to the Minimum Description Length Shape Models and uses it to get better models.

1 Introduction

Statistical models of shape and appearance have been used widely in image segmentation and interpretation. To construct such models, we require sets of labeled training examples. The labels consist of landmark points defining the correspondences between similar structures in each example across the set. The problem of establishing the 'correct' correspondences is fundamental to the shape model building. The correspondences are often established by manual annotation, which is labor intensive and error-prone. To reduce the burden, several authors stated the correspondence problem as an optimization problem. Landmark points are altered and refined across all training examples until an objective function that measures the utility of the resulting model in minimized.

Kotcheff and Taylor [1] describe an approach where the best model is defined in terms of 'compactness', as measured by the determinant of its covariance matrix. Although the method is workable, its objective function could not be rigorously justified and it was difficult to make the optimization converge. To solve these problems, Davies [2, 3] defined a new objection function with a rigorous theoretical basis and described a new representation of correspondence. The automatic model optimization method of him is based on finding the set of dense correspondences over a set of shapes that produce the 'simplest' linear statistical shape model. A minimum description length (MDL) objective function is used to measure model complexity, and optimized numerically with respect to the correspondences. The basic idea is that 'natural' correspondences give rise to simple explanations of the variability in the data. One shape example was chosen as a reference shape and the positions of its correspondence points remained fixed throughout. The optimization process involved perturbing the locations of the correspondence points of each shape in turn optimizing the MDL objective function. Thodberg [4] proposed curvature as another salient piece of information. By adding a term to the cost function expressing the mismatch of curvature features across the data set, the MDL approach can be fine-tuned.

However, the above MDL approaches only concern on the shape characteristics without capturing the local image structure that can also be used in finding point correspondences. This paper uses the differential structure of image to construct a vector of invariants over a range of scales at each node and add a CDICost term to the cost function. After the same optimization passes, our method establishes better point correspondence and the models generated by this approach suggest more specific and generalization ability.

2 Statistical Shape Models

A statistical shape model is built from a training set of ns shapes, aligned to a common coordinate frame. Each shape, $S_i(i=1,\ldots n_s)$, can be represented by a set of n landmarks. By concatenating the coordinates of the sample points, each shape can be represented by an n_p dimensional shape vector x_i. Principal Component Analysis (PCA) is then performed to define a set of axes that aligned with the principal directions of the data. A linear model of the form can now express the ith shape

$$x_i = \bar{x} + Pb_i \quad \bar{x} \quad \sum_m p^m b_i^m \qquad (1)$$

where $\bar{x}$ is the mean shape vector, the columns of P describe a set of orthogonal modes of shape variation, and b_i is the vector of shape parameters for the ith shape.

3 Minimum Description Length Shape Models

Davies et al. [2] derived an object function with favors models that encode the training set most efficiently. The correspondence problem is regarded as an optimization problem. Point samples are altered and refined across all training shapes until an objective function is minimized.

The correspondence is optimized with respect to the following costs:

1) The MDL cost of mark positions:

$$DL = \sum_{\lambda \ge \lambda_{cut}} \left(1 + \log \frac{\lambda_m}{\lambda_{cut}}\right) + \sum_{\lambda < \lambda_{cut}} \frac{\lambda_m}{\lambda_{cut}} \quad (2)$$

2) The Node cost of the mark positions:

$$\text{NodeCost} = \sum (\alpha_i^{\text{average}} - \alpha_i^{\text{target}})^2 \quad (3)$$

λ_m are the eigenvalues of a principle components decomposition of landmark positions, λ_{cut} is preset threshold; $a_i^{average}$ and a_i^{target} are average and target parameters for the ith landmark.

Hans Henrik Thodberg [4] added the Curvature variation cost that expresses the mismatch of curvature features across the data set. The curvatures can be weighted with a factor and appended to the aligned position coordinates and included in the PCA:

$$\text{CurvatureCost} = C \frac{1}{N} \frac{1}{s} \sum_{i,r} (k_{ir} - k_i^{mean})^2 \quad (4)$$

k_{ir} is the curvature at ith landmark on rth shape, k_i mean is the mean curvature at the ith landmark across all shapes, s is the number of shapes and C is a weighting factor for this term.

4 The Cartesian Differential Invariants Cost

To get better models, we use the differential structure of the image to construct a vector of invariants over a range of scales centered at each landmark. These vectors describe the local geometry around each landmark at a particular scale independently of translation and orientation.

We use a scaled set of differential operators to extract the differential structure of the image. A complete, hierarchically ordered family of multiplicative scale-space kernels in D dimensions is given by the following set of convolution filters:

$$\left\{ G_{i_1 \ldots i_n}(x;\sigma) = \partial_{i_1 \ldots i_n} G(x;\sigma) \mid (x;\sigma) \in \mathbb{R}^D \times \mathbb{R}^+, n \in \mathbb{Z}_0 \right\} \quad (5)$$

where $G(x;\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{1}{2}\frac{x^2}{\sigma^2})$, $\partial_{i_1 \ldots i_n}$ is the linear partial derivative operator: $\frac{\partial^n}{\partial x_{i_1} \ldots \partial x_{i_n}}$

Let $\psi : \mathbb{R}^D \to \mathbb{R}$ be a given image and let σ be a physically sensible inner scale for ψ, then the local jet of ψ of order N at base point x and inner scale σ can be represented with respect to an arbitrary Cartesian coordinate system by the set:

$$J^N[\psi](x;\sigma) = \left\{ L_{i_1 \ldots i_n}(x;\sigma) \mid (x;\sigma) \in \mathbb{R}^D \times \mathbb{R}^+, n = 0, \ldots, N \right\} \quad (6)$$

where $L_{i_1 \ldots i_n}$ is given by the convolution of ψ with the Gaussian derivative $G_{i_1 \ldots i_n}(x;\sigma)$:

$$L_{i_1 \ldots i_n}(x;\sigma) = (G_{i_1 \ldots i_n} * \psi)(x;\sigma) \quad (7)$$

The certain combinations of $L_{i_1 \ldots i_n}$ are known as Cartesian differential invariants [5], these combinations are independent of the choice of coordinate frame. The flowing shows the set of 2D polynomial invariants to third order:

$$I1 = L_i L_i \qquad I2 = L_i L_{ij} L_j$$
$$I3 = L_{ii} L_j L_j - L_{ij} L_i L_j \qquad I4 = -\varepsilon_{ij} L_{jk} L_i L_k$$
$$I5 = \varepsilon_{ij}(L_{jkl} L_i L_k L_l - L_{jkk} L_i L_l L_l) \qquad I6 = L_{iij} L_j L_k L_k - L_{ijk} L_i L_j L_k$$
$$I7 = -\varepsilon_{ij} L_{jkl} L_i L_k L_l \qquad I8 = L_{ijk} L_i L_j L_k$$

For each landmark of a training example, the above Cartesian Invariants over a range of scales describe the differential structure of the example independently of the chosen Cartesian coordinate system. We can construct a vector CDI that represents the local image structure around a given landmark over 3 scales (σ=4, 5, 6). The filter's size are defined as σ*7, which will avoid significant truncation, without wasting the outer taps on near-zero values.

$$\begin{aligned} CDI(x,y) = [&I1(\sigma=4), I2(\sigma=4), \ldots, I8(\sigma=4), \\ &I1(\sigma=5), I2(\sigma=5), \ldots, I8(\sigma=5), \\ &I1(\sigma=6), I2(\sigma=6), \ldots, I8(\sigma=6)]^T \end{aligned} \quad (8)$$

For each landmark i:

$$Cost_i = \sum_{k=1}^{24} (CDI_{ik} - CDI_i^{mean})^2 \quad (9)$$

$$CDI_i^{mean} = \frac{1}{s} \sum_{k=1}^{s} CDI_{ik} \quad (10)$$

Because there are N landmarks on each example, the total cost:

$$CDICost = C_{cdi} \frac{1}{N} \sum_{j=1}^{N} Cost_j \quad (11)$$

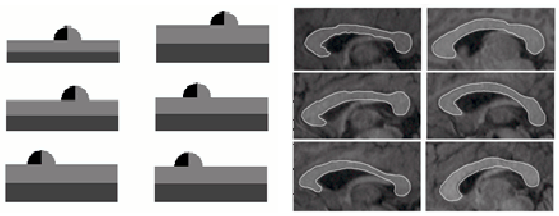
where C_{cdi} is a weighting factor for this term. After adding CDICost to MDL, the cost of MDL becomes:

$$MDLCost = DL \quad NodeCost + CurvatureCost + CDICost \tag{12}$$

MDLCost can be used in Our MDL optimization algorithm that has 3 steps. The first is Initialization. Each shape is defined by a number of landmarks and the reparameterisation functions are initially set to be equally space. The second step is rescaling and aligning shapes. To obtain a true shape representation, each iteration start by aligning and rescaling all curves according to Procrustes analysis. The third step is Landmarks position updating. The MDL optimization does not need to be done on all landmarks, but only on landmarks up to a given level. These active landmarks are called nodes that are ordered according to ascending level. Each node is associated with a step length and the parameters anode for each node and each training example are probed which run over a number of passes, until the MDLCost has stabilized.

5 Experiment Results

We tested our method on a set of 24 box-bumps and 20 MR images of the corpus callosum. The 6 training samples of Bump model and corpus callsosum model are shown in figure 1. The box-bumps are generated with a bump at a varying location and with varying aspect ratio of the box and have the similar gray-level distribution. The MR images of corpus callosum are randomly selected from Dinggang Sheng's database of normal elderly subjects participating in the Baltimore Longitudinal Study of Aging [7].



(a) (b)

Figure 1. The box-bumps used as training set (a) and the MR images of corpus collosum (b).

In the experiment of box-bumps, 8 control nodes have been used for the reparameterisation, 64 markers are sampled on each box-bump to evaluate the Description Length at the given parameterisation. We set the λ_{cut} as 0.003 and passes as 40. After 80000 (40passes*8nodes*5steps*24examples) evaluations, the results of MDL with CDICost analysis of the training set and the results of MDL with CDICost turned off are shown in fig 2. In the experiment of Corpus Callosum, the same control nodes (8) and markers (64) are used to build MDL model. We set the λ_{cut} as 0.001 and passes as 160. The results are shown in fig 3. It can be seen that MDL with CDICost can produce better result than that without CDICost.

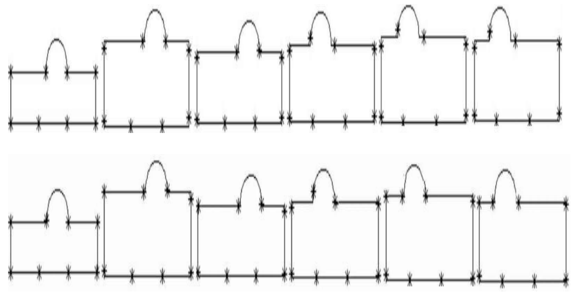


Figure 2. Result of MDL without CDICost (top) and Result of MDL with CDICost (bottom) of box-bumps.

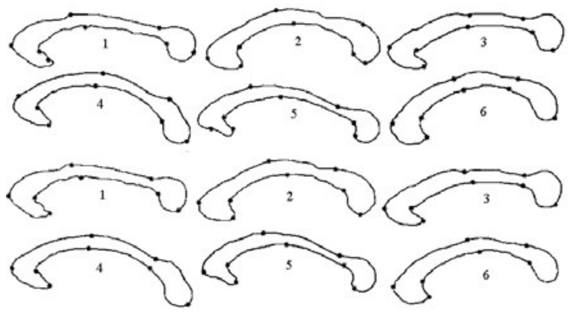


Figure 3. Result of MDL without CDICost (top) and Result of MDL with CDICost (bottom) of corpus collosum.

In figure 4 and 5, we show qualitative results by displaying the variation captured by the first two modes of each model (b^m (m=1, and m=2) varied by ±3 standard deviations over training set. About box-bumps model, the first parameter corresponds to the largest eigenvalue of the covariance matrix, which gives its variance across the training set, and the main effect is close to a horizontal displacement of the bump. For the second parameter, the main effect describes the box's aspect ration changes. The results show that the shapes generated by our method are plausible and suggest specific models.

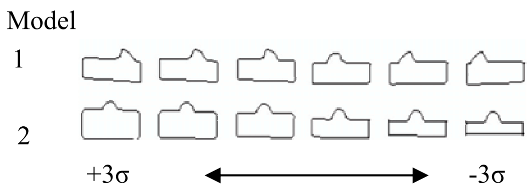


Figure 4. The first two modes (m=1, m=2) of shape variation (±3σ) of the automatically generated models of box-bumps.

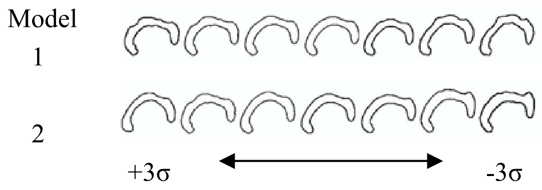


Figure 5. The first two modes (m=1, m=2) of shape variation (±3σ) of the automatically generated models of corpus callosum.

We also give the quantitative results in table 1 and 2, tabulating the value of F (CDICost subtracted from MDLCost) and the variances explained by the first 3 modes, comparing our methods with MDL method and equally spaced method. The results show that CDICost can be work as a catalyst like curvature cost for minimizing the MDL cost and produces models more compact than either the models built by equally spaced method or those obtained using the current MDL methods.

Table 1. A quantitative comparison of box-bumps showing the variance explained by first 3 modes. F is the value of CDICost subtracted from MDLCost.

Mode	Equally spaced	MDL	Our method
1	0.0071548	0.0045953	0.0036741
2	0.0032664	0.0023326	0.0016741
3	0.0007587	0.0006353	0.0003023
F	31.677	17.809	15.6021

Table 2. A quantitative comparison of corpus collosum showing the variance explained by first 3 modes. F is the value of CDICost subtracted from MDLCost.

Mode	Equally spaced	MDL	Our method
1	0.0018313	0.0016861	0.0015949
2	0.0013193	0.0013196	0.0011758
3	0.0011177	0.0008383	0.0008604
F	38.239	25.292	22.234

6 Conclusions

This paper describes how vector of image invariants can be used to the establishing correspondence between a set of training images. Different from the existing MDL approaches, our approach considers not only the shape characteristics but also the local image structure characteristics. By adding a CDICost term to MDL approach, we get more compact and specific models than that built by the existing MDL approaches.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No.60271033). We thank Dr. Dinggang Shen for his providing the data as part of his research in University of Pennsylvania, also his helpful comments and discussions. We also thank Professor Thodberg for providing his matlab code of MDL.

References

[1] Kotcheff, A.C.W. and C.J. Taylor: Automatic Construction of Eigenshape Models by Direct Optimization. Medical Image Analysis, 1998, volume 2, pages 303-314.

[2] Rhodri H Davies, Carole J Twining, Tim F Cootes, John C Waterton, and Chris T Taylor: A minimum description length approach to statistical shape modelling. IEEE Trans. on Medical Imaging, 2002, 21(5): pages 525–537.

[3] Rhodri H Davies, Carole J Twining, Tim F Cootes, John C Waterton, and Chris T Taylor: 3D statistical shape models using direct optimisation of description length. In 7th European Conference on Computer Vision, pages 3–21, 2002.

[4] H.H Thodberg, H. Olafsdottir: Adding curvature to minimum description length shape models. In 14th British Machine Vision Conference, 2003, volume 2, pages 251–260.

[5] Luc M J Florack, Bart M ter Harr Romeny, Jan J Koenderink and Max A Viergever: Scale and the Differential Structure of Images, Image & Vision Computing, 1992, vol. 10, no.6, pages 376-388.

[6] Christos Davatzikos, Xiaodong Tao, Dinggang Shen: Hierarchical Active Shape Models, Using the Wavelet Transform, IEEE Trans. on Medical Imaging, 2003, 22(3), pages 414-423.