Local Log-Euclidean Covariance Matrix (L²ECM) for Image Representation and Its Applications

Peihua Li, Qilong Wang

Heilongjiang University, School of Computer Science and Technology, China

Overview

Contributions of L²ECM :

1.We propose the model of Local Log-Euclidean Covariance Matrix (L²ECM) for representing the neighboring correlation of multiple image cues. By L²ECM, we produce a novel vector valued-image which captures the local structure of the original one .

2.The benefits of the L²ECM are that it preserves the manifold structure of the covariance matrices, while enabling efficient and flexible operations in the Euclidean space instead of in the Riemannian manifold.

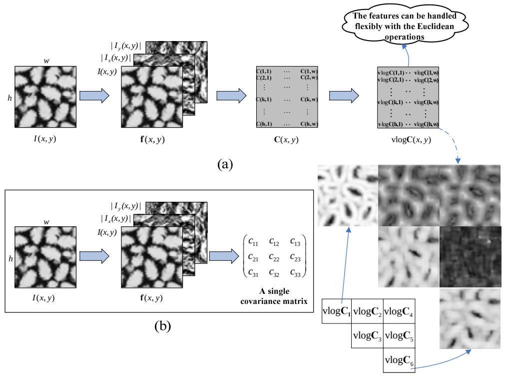

Fig. 1. Overview of L²ECM (3-D raw features are used for illustration). (a) shows the modeling methodology of L²ECM. Given an image $I(x,y)$, the raw feature image $\mathbf{f}(x,y)$ is first extracted; then the tensor-valued image $\mathbf{C}(x,y)$ is obtained by computing the covariance matrix for every pixel; after the logarithm of $\mathbf{C}(x,y)$, the symmetric matrix $\log\mathbf{C}(x,y)$ is vectorized to get the 6-D vector-valued image denoted by $\text{vlog}\mathbf{C}(x,y)$, slices of which are shown at the bottom-right. (b) shows the modeling methodology of Tuzel et al. –only one *global* covariance matrix is computed for the overall image of interest.

Motivation

Inspired by the structure tensor which computes the second-order moment of image gradients for representing local image properties, and the Diffusion Tensor Imaging (DTI) which produces tensor-valued image characterizing the local tissue structure, our motivation is to represent the local image properties via covariance matrices capturing the correlation of various image cues.

Structure Tensor	DTI	L²ECM
$\begin{bmatrix} I_x^2 & I_xI_y \\ I_xI_y & I_y^2 \end{bmatrix}$	$\begin{bmatrix} D_{11} & D_{12} & D_{13} \\ D_{12} & D_{22} & D_{23} \\ D_{13} & D_{23} & D_{33} \end{bmatrix}$	$\mathbf{C}=\begin{bmatrix} C_{11} & \cdots & C_{1n} \\ \vdots & \vdots & \vdots \\ C_{n1} & \cdots & C_{nn} \end{bmatrix} \xrightarrow{\text{vlog}} \begin{bmatrix} \text{vlog}\mathbf{C}_1 & \cdots & \text{vlog}\mathbf{C}_{m-n+1} \\ & \ddots & \vdots \\ & & \text{vlog}\mathbf{C}_m \end{bmatrix}$
2nd-order moment of partial derivatives of image *I* w.rt *x,y*	3×3 symmetric matrix describing molecules diffusion	Logarithm of nxn (n=2~5) covariance matrix **C** of raw features followed by half-vectorization ($m=(n^2+n)/2$) due to symmetry.

Table 1. Comparision of Tensor-valued (Matrix-valued) images

Log-Euclidean Framework on SPD Matrices

The Log-Euclidean framework [8] establishes the theoretical foundation of our methodology, in which we compute the logarithms of SPD matrices which are then handled with Euclidean operations.

The briefly description is given below:

Let $S(n)$ and $SPD(n)$ be the spaces of n by n symmetric matrices and SPD matrices, respectively .

1) The Lie group of $SPD(n)$ is isomorphic and diffeomorphic to $S(n)$.

2) $SPD(n)$ with the bi-invariant metrics is isometric to $S(n)$ with the associated Euclidean metrics.

3) The Lie group isomorphism exponential mapping from the Lie algebra of $S(n)$ to $SPD(n)$ can be smoothly extended into an isomorphism of vector spaces.

The key matrix operators: Matrix exponential and logarithm

By eigen-decomposition $\mathbf{S}=\mathbf{U}\Lambda\mathbf{U}^T$, the exponential of a $\mathbf{S}\in S(n)$ can be computed as :

$$\exp(\mathbf{S})=\mathbf{U}\cdot Diag(\exp(\lambda_1),\cdots,\exp(\lambda_n))\cdot\mathbf{U}^T$$

For any SPD matrix $\mathbf{S}\in SPD(n)$, there exists a unique logarithm in $S(n)$:

$$\log(\mathbf{S})=\mathbf{U}\cdot Diag(\log(\lambda_1),\cdots,\log(\lambda_n))\cdot\mathbf{U}^T$$

Lie group structure on SPD(n)

$SPD(n)$ with the associated logarithmic multiplication has Lie group structure:

$$\mathbf{S}_1\odot\mathbf{S}_2\triangleq\exp(\log(\mathbf{S}_1)+\log(\mathbf{S}_2))$$

where $\mathbf{S}_1,\mathbf{S}_2\in SPD(n)$.

Vector space structure on SPD(n)

The commutative Lie group $SPD(n)$ admits a bi-invariant Riemannian metrics and the distance between two matrices $\mathbf{S}_1$; $\mathbf{S}_2$ is

$$d(\mathbf{S}_1,\mathbf{S}_2)=\|\log(\mathbf{S}_1)-\log(\mathbf{S}_2)\|$$

where $\|\bullet\|$ is the Euclidean norm in the vector space $S(n)$. This bi-invariant metrics is called Log-Euclidean metrics, which is invariant under similarity transformation. For a real number α, define the logarithmic scalar multiplication between α and a SPD matrix **S** :

$$\alpha\otimes\mathbf{S}\triangleq\exp(\alpha\log(\mathbf{S}))=\mathbf{S}^\alpha$$

Provided with logarithmic multiplication and logarithmic scalar multiplication , the $SPD(n)$ is equipped with a vector space structure.

[8]Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Geometric means in a novel vector space structure on symmetric positive-definite matrices. SIAM J. Matrix Anal. Appl. (2006)

L²ECM Feature Image

Provided with the raw feature vectors, we can obtain a tensor-valued image by computing the covariance matrix $\mathbf{C}(x,y)$ at every pixel:

$$\mathbf{C}(x,y)=\frac{1}{N_{S_r}-1}\sum_{(x',y')\in S_r(x,y)}(\mathbf{f}(x',y')-\bar{\mathbf{f}}(x,y))(\mathbf{f}(x',y')-\bar{\mathbf{f}}(x,y))^T$$

$$\bar{\mathbf{f}}(x,y)=\frac{1}{N_{S_r}}\sum_{(x',y')\in S_r(x,y)}\mathbf{f}(x',y')$$

where $\mathbf{f}(x,y)$ which, for example, has the following form:

$$\mathbf{f}(x,y)=\left[I(x,y)\,|I_x(x,y)|\,|I_y(x,y)|\,|I_{xx}(x,y)|\,|I_{yy}(x,y)|\right]^T$$

Because of its symmetry, we perform half-vectorization of $\log\mathbf{C}(x,y)$, denoted by $\text{vlog}\mathbf{C}(x,y)$ i.e., we pack into a vector in the column order the upper triangular part of $\log\mathbf{C}(x,y)$. The final L²ECM feature descriptor can be represented as

$$\text{vlog}\mathbf{C}(x,y)=\left[\text{vlog}\mathbf{C}_1(x,y)\ \text{vlog}\mathbf{C}_2(x,y)\ldots\text{vlog}\mathbf{C}_{n(n+1)/2}(x,y)\right]$$

The covariance matrices can be computed efficiently via the Integral Images.

The L²ECM may be used in a number of ways:

– It may be seen as "imaging" technology by which various novel multi-channel images are produced. When $n=2$, by combinations of varying raw features, e.g. two components of gradients, we obtain different 3-D "color" images that may be suitable for a wide variety of image or vision tasks.

– Statistical modeling of the L²ECM features is straightforward by probabilistic mixture models, e.g. Gaussian mixture model (GMM), principal component analysis (PCA), etc. This way, the geometric structure of covariance matrices is preserved while avoiding directly computational expensive algorithms in Riemannian space.

– We can straightforwardly apply L²ECM features to a variety of machine learing methods, such as, SVM, adaboost, random forest, in the same manner of conventional vector.

Applications of L²ECM Features

Statistical modeling by the second-order moment

Human Detection

For performance evaluation, we exploit the INRIA person dataset , a challenging benchmark dataset . It includes 2416 positive, normalized images and the 1218 person-free images for training, together with 288 images of humans and 453 person-free images for testing.

For a normalized image (96×160), we first compute the L²ECM feature image. Then we divide the vector valued image into 12 overlapping, 32×32 blocks with a stride of 16 pixels. We compute for each block the second-order moment (covariance matrix) which is again subject to matrix logarithm and half-vectorization. The resulting feature for the whole, normalized image is a 1440-dimensional vector. We exploit the linear SVM with default parameters for classification.

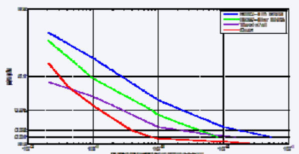

Fig. 3. DET curves of human detection on the INRIA person dataset.

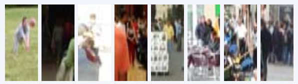

Fig.2. Some samples on the INRIA person dataset.

[29] Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C.: Local features and kernels for classification of texture and object categories: A comprehensive study. Int. J. Comput. Vision 73 (2007) 213–238

Texture Classification

The Brodatz database and KTH-TIPS database are used for performance evaluation. The Brodatz dataset contains 111 textures (texture D14 is missing); KTH-TIPS database has 10 texture classes each of which is represented by 81 image samples.

For each image, we first compute the L²ECM feature image; the feature image is then divided into four patches the covariance matrices of which are computed; KNN algorithm ($k=5$) is used for classification in our method. The votes of the four matrices associated with this testing image determine its classification.

For comparison, Lazebnik's method , Varma&Zisserman method, Hayman's method , global Gabor Filters (Manjunath, B.) , and Harris detector+Laplacian detector+SIFT descriptor+SPIN descriptor((HS+LS)(SIFT+SPIN))[29] are used.

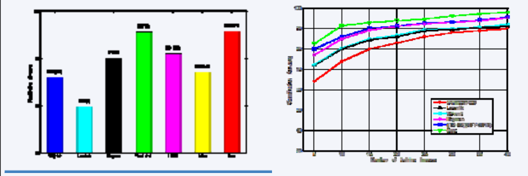

Fig. 4. Texture classification on the Brodatz (left) and KTH-TIPS (right) databases.

Object Tracking

We use similar tracking framework as Tuzel et al. except covariance matrix computation and model update. The difference of covariance matrix computation is shown in Fig.5.

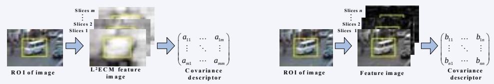

Fig. 5. Covariance matrix computation in the L²ECM (left) and Tuzel (right) trackers.

Image Seq.	Method Dist. err (pixels)	Succ. frames
Car seq.	Tuzel 10.76 ± 5.72	190/190
	L²ECM 7.55 ± 3.38	190/190
Face seq.	Tuzel 20.44 ± 13.87	370/370
	L²ECM 4.45 ± 3.87	370/370
Mall seq.	Tuzel 30.92 ± 16.58	116/190
	L²ECM 17.67 ± 10.45	190/190

Table 2. Comparison of average tracking errors (mean ± std) and number of successful frames vs total frames.

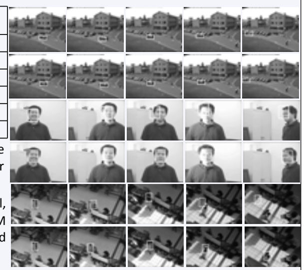

Fig. 6. Tracking results. In each panel, the results of Tuzel tracker and L²ECM tracker are shown in the first and second rows, respectively.

Contact* : e-mail: peihualj@hotmail.com , wangqilong.415@163.com ; website :http://peihuali.org/Publications.htm