K-CENTROIDS-BASED SUPERVISED CLASSIFICATION OF TEXTURE IMAGES: HANDLING THE INTRA-CLASS DIVERSITY

Aurélien Schutz, Lionel Bombrun and Yannick Berthoumieu

Université de Bordeaux, ENSEIRB-Matmeca, Laboratoire IMS, Groupe Signal et Image
{*aurelien.schutz, lionel.bombrun, yannick.berthoumieu* }*@ims-bordeaux.fr*

ABSTRACT

Natural texture images exhibit a high intra-class diversity due to different acquisition conditions (scene enlightenment, perspective angle, . . .). To handle with the diversity, a new supervised classification algorithm based on a parametric formalism is introduced: the K-centroids-based classifier (K-CB). A comparative study between various supervised classification algorithms on the VisTex and Brodatz image databases is conducted and reveals that the proposed K-CB classifier obtains relatively good classification accuracy with a low computational complexity.

***Index Terms*—** Supervised classification, texture, Jeffrey divergence, information geometry.

1. INTRODUCTION

In research area devoted to machine vision, texture analysis is still a challenging issue for applications using classification, segmentation or indexing approaches. One of the major problem is the diversity in appearance of the image samples coming from the same class of natural texture. Diversity relates jointly to phenomena such as the near-stationary content but not strict, the changes in illumination or in viewing position (as shown in Fig. 1) and so on.

Fig. 1. Natural diversity inside Wood class of the Brodatz database

The consequence is that intra-class diversity leads to possible misclassification if we work with just one training sample to extract features. A single image sample does not allow the understanding of the underlying geometry of the cluster in the feature space. Diversity induces a complex geometric expansion of the cluster and leads to a "macro-cluster". In order to provide invariant approaches to the intra-class diversity with a robust modeling, the geometry of the macro-cluster must be characterised. Considering the above remarks, an obvious approach for taking into account the diversity is to consider supervised approach based on a large training database. Samples form a set of representative instances of the diversity.

Thanks TOTAL for funding.

A possible solution for classifying samples, without any prior on the class probability, is to propose instance-based approach such as K-NN algorithm. However, a main disadvantage of the K-NN method is the required time to estimate the class of a test image. Let us consider N_{Tr} training samples in a space of dimension d. Then applying the method to one test sample requires $\mathcal{O}(N_{Tr}d)$ time. If the trained features are stored in a sophisticated data structure, e.g. kd-tree, then finding nearest neighbours can be done much faster if the dimensionality d is small, typically less than 20, which is generally too low in practice [1].

In the field of texture analysis, diversity has been considered by few authors. Varma and Zisserman [2] have proposed a K-means algorithm to partition the macro-cluster into a finite number of sub-clusters. This finite partition enables authors to reduce the complexity against K-NN. However the partitioning is done directly on the magnitude of the filter responses using a quadratic norm. The issue is that data draw a non-euclidean high dimensional. For this manifold, the Euclidean metric is not adapted, leading to algorithm with limited performances and implies to work with a high number of sub-clusters. In the context of texture modeling, various authors have proposed the use of jointly scale-space approaches and statistical modeling characterize the textural content [3, 4, 5, 6, 7, 8, 9]. The modeling process consists in fitting the histogram of each sub-band with a given parametric probability density function (pdf). The modeling process results in homogeneous textured samples summarized by a limited number of parameters, inducing directly a dimension reduction in classifying. In [8], Do and Vetterli use a centred generalized Gaussian distribution (GGD). In this seminal work, it is also shown that probabilistic similarity measure such as Kullback-Leibler divergence, which can be derived in closed-form in terms of pdf parameters, can be used. Thereby, the parametric space forms a smooth Riemannian manifold for which well founded processing can be derived. In this way, Choy and Tong proposed in [9] to compute a centroid from several instances of parameter vectors from each sub-band for a given class. Even if the proposed algorithm exploits the geometrical properties of the manifold, the main drawback of the last proposal is that classifying with only one centroid is not enough robust to handle diversity.

In this paper, our contributions are two fold. First, an extension of the Choy's algorithm, which takes into account the direction of the derivatives in order to speed-up the convergence of the steepest descent, is derived. We also give the expressions of the update in terms of manifold constraints. Second, we propose to use the K-means algorithm inside the K-centroid-based (K-CB) supervised classification. This supervised classification is new because Varma and Zisserman does not use the parametric point-of-view. This new supervised classification, enhances the Choy and Tong algorithm by using a higher number of centroids according to the natural diversity found in texture images. Finally, numerical application of K-CB shows an increase of overall accuracy against Choy and Tong algorithm and a computation time significantly lower than the 1-NN algorithm.

2. BARYCENTRIC REPRESENTATION

Let I be a texture image. Let N_o and N_s be respectively the number of orientation and scale of a multi-scale decomposition. I is hence decomposed into $N_o \times N_s$ sub-bands. Let us consider the parametric vector $\lambda_{s,o}$ of the pdf associated to each sub-band. The collection T of those parametric vectors will represent the texture image I.

$$T = \{\lambda_{s,o} | s = 1, \ldots, N_s, o = 1, \ldots, N_o\}. \quad (1)$$

The components $\lambda_{s,o}$ of the vector T form a parametric Riemannian manifold. In the sequel of the paper, we call $\mathcal{M}$ the corresponding manifold.

2.1. Computing a centroid

Let $(T_{c,n})_{n=1}^{N_{Tr}}$ be N_{Tr} training samples from the same class c. In [9], Choy and Tong have introduced an iterative algorithm to estimate the barycentric sample $\bar{T}_c$ (also called centroid) from this collection of samples. Let $l_c(T)$ be the cost function defined by:

$$l_c(T) = \frac{1}{N_{Tr}} \sum_{n=1}^{N_{Tr}} m(T \| T_{c,n}), \quad (2)$$

the centroid is obtained as the solution of the following optimization problem:

$$\bar{T}_c = \underset{T \in \mathcal{M}}{argmin}\ l_c(T). \quad (3)$$

The dissimilarity measure m between two instances of T is computed as the sum of the dissimilarity measures SIM between all sub-band distributions at each scale and orientation:

$$m(T_{c,n} \| T_{c',n'}) = \sum_{s=1}^{N_s} \sum_{o=1}^{N_o} \text{SIM}(p(x; \lambda_{c,n,s,o}) \| p(x'; \lambda_{c',n',s,o})); \quad (4)$$

where $p(x; \lambda_{c,n,s,o})$ is the probabilistic distribution which model the sub-band coefficients x at scale s and orientation o.

This paper introduces the natural gradient algorithm [10, 11] to solve the optimization problem defined in (3). Let $\bar{T}_c$ be the solution of (3), *i.e.* the minimizer of the cost function $l_c(T)$. To speed-up the convergence, the Fisher information matrix $G(T)$ is included with the gradient $\nabla l_c(T)$ of the cost function in the optimization step. Then, the sequence $\left(\bar{T}_{c,i}\right)_{i=1}^{\infty}$ defined by:

$$\bar{T}_{c,i+1} = Proj_{\mathcal{M}}\left(\bar{T}_{c,i} - G^{-1}(\bar{T}_{c,i}) \nabla l_c(\bar{T}_{c,i})\right), \quad (5)$$

converges to the centroid $\bar{T}_c$. The operator $Proj_{\mathcal{M}}$ representing the projection on the manifold $\mathcal{M}$ assures that $\bar{T}_c$ belongs to the manifold $\mathcal{M}$. Practically, on the VisTex database, the projected gradient descent algorithm of [9] converges in 170 iterations whereas the proposed projected natural gradient converges in only 9 iterations to the same solution.

In the following, the computation of the centroid will be applied to the generalized Gaussian distribution (GGD) since this model has been successfully validated for the modeling of wavelet coefficients of texture images [8, 7, 12]. Note that it can be generalized to any other stochastic models provided that a closed-form expression of the similarity measure m and the Fisher information matrix exist.

2.2. Application to the generalized Gaussian distribution

The probability density function of an univariate GGD is

$$p(x; \lambda) = \frac{1}{\alpha \Gamma(1/\beta + 1)} \exp\left\{ - \left(\frac{|x|}{\alpha} \right)^{\beta} \right\}, \quad (6)$$

where $\Gamma(z) = \int_{\mathbb{R}^+} t^{z-1} e^{-t} dt$ is the Gamma function, α and β being respectively the scale and shape parameters. In the following the parameter space for one sub-band is represented by $\lambda = \{\alpha, \beta\}$.

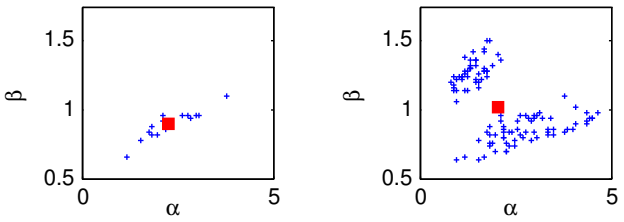
The Jeffrey divergence (i.e. the double sided Kullback-Leibler divergence) is considered as a dissimilarity measure SIM between two GGDs, its expression is given by [8]:

$$\begin{aligned} \text{JD}(p(x;\lambda) \| p(x;\lambda')) &= A + A' - \frac{1}{\beta} - \frac{1}{\beta'} \\ A = \frac{\Gamma((\beta'+1)/\beta)}{\alpha^{-\beta'}(\alpha')^{\beta'}\Gamma(1/\beta)}, &\quad A' = \frac{\Gamma((\beta+1)/\beta')}{(\alpha')^{-\beta}\alpha^{\beta}\Gamma(1/\beta')}. \end{aligned} \quad (7)$$

Next, by combining (4) and (7), one obtains the dissimilarity measure m between two samples. Then, to derive the cost function and obtain its gradient $\nabla l_c(T)$, the partial derivatives of the Jeffrey divergence are needed. It yields:

$$\begin{aligned} \frac{\partial \text{JD}}{\partial \alpha}(\lambda, \lambda') &= \frac{\beta'}{\alpha} A + \frac{\beta}{\alpha} A', \\ \frac{\partial \text{JD}}{\partial \beta}(\lambda, \lambda') &= -\left(A' \left[\ln\left\{ \frac{\alpha'}{\alpha} \right\} + \frac{1}{\beta'} \Psi\left(\frac{\beta+1}{\beta'} \right) \right] - \right. \\ &\quad \left. A \left[\Psi\left(\frac{1}{\beta} \right) - (\beta'+1) \Psi\left(\frac{\beta'+1}{\beta} \right) \right] \right) + \frac{1}{\beta^2}, \end{aligned} \quad (8)$$

where $\Psi(z)$ is the digamma function.



(a) (b)

Fig. 2. Scatter plot in the α/β plane: (a) for textures with low intra-class diversity and (b) for textures with different orientations.

To compute the natural gradient, the 2×2 Fisher information matrix of a GGD is also required. Its components are:

$$g_{\alpha\alpha}(\lambda) = \frac{\beta}{\alpha^2}, \quad g_{\alpha\beta}(\lambda) = \frac{1}{\alpha\beta}\left(\Psi\left(\frac{1}{\beta}\right) - \beta + 1\right),$$
$$g_{\beta\beta}(\lambda) = \frac{\beta+1}{\beta^4}\Psi'\left(\frac{1}{\beta}\right) + \frac{1}{\beta^2} + \frac{1}{\beta^3}\left[\Psi^2\left(\frac{1}{\beta}\right) + 2(\beta+1)\Psi\left(\frac{1}{\beta}\right)\right]. \quad (9)$$

Finally, by injecting (8) and (9) in (5), one can iteratively estimate the centroid for generalized Gaussian distributed sub-bands.

2.3. Capabilities and limits of an unique centroid

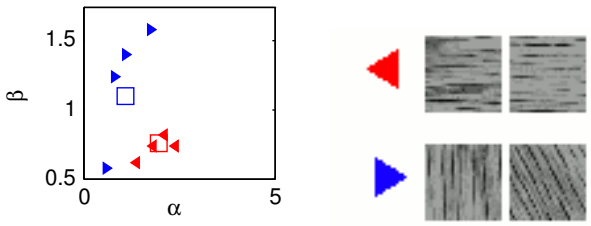
In this section, some experiments are conducted on real texture images to evaluate the potential and limits of the centroid definition. Fig. 2 draws a scatter plot of the training sample (blue cross) in the α/β feature space. The red square corresponds to the location of the estimated centroid. As observed on Fig. 2.(a), when the scatter plot is compact, the centroid represents well the cluster. Nevertheless, when the intra-class diversity is large, an unique centroid is not able to capture this diversity. This is the case for Fig. 2.(b) where the samples are issued from the same texture class but with different orientations. The natural diversity inside a class of texture images are due to many reasons such as differences in the scene enlightenment, differences in the scale considered, differences in the perspective, ... This diversity modifies the shape of the scatter plot by stretching or splitting the clusters. To capture this intra-class diversity, we propose in the next section a multi-barycentric approach based on an adaptation of the K-means algorithm to the centroid definition introduced in Section 2.

3. K-CENTROIDS-BASED SUPERVISED CLASSIFICATION (K-CB)

3.1. Principle

For each class $c = 1, \ldots, N_c$, K centroids $(\bar{T}_{c,k})_{k=1}^{K}$ are computed according to the K-means classifier described in Algorithm 1. Let T_t be a test sample. This sample is labeled to the class $\hat{c}$, corresponding to the class of the closest centroid, *i.e.*

$$\hat{c} = \underset{c}{argmin}\left\{m(T_t \| \bar{T}_{c,k}) | \forall k = 1, \ldots, K\right\}. \quad (10)$$



(a) (b)

Fig. 3. (a) Scatter plot in the α/β plane for the 3^{rd} sub-band and (b) example of samples from each cluster.

In the following, this supervised classification algorithm will be referred to K-CB. Note that the algorithm of Choy and Tong presented in Section 2 is a particular case of K-CB when $K = 1$. Note also that when K is equal to the number of training samples per class, N_{Tr}, each training sample is a centroid and the K-CB classifier reduces to the nearest neighbour (1-NN) classifier.

Algorithm 1 Pseudo-code of the proposed K-means algorithm

Require: A collection of N_{Tr} training samples $(T_n)_{n=1}^{N_{Tr}}$ and K the number of clusters
Ensure: A collection of K centroids $\left(\bar{T}_k\right)_{k=1}^{K}$

Initialization of K centroids $\left(\bar{T}_k\right)_{k=1}^{K}$
Computation of the intra-cluster inertia I_{intra}
repeat
  $I_{\text{old}} \leftarrow I_{\text{intra}}$
  for each training sample n **do** ▷ Assignment
    $k_n \leftarrow \underset{k}{argmin}\, m(\bar{T}_k \| T_n)$
  end for
  for each cluster k **do** ▷ Update the centroids
    $N_k \leftarrow \sum_{n=1}^{N} \delta(k_n, k)$
    $\bar{T}_k \leftarrow \underset{T}{argmin}\left\{\frac{1}{N_k}\sum_{n=1}^{N} m(T \| T_n)\delta(k_n, k)\right\}$
  end for
  $I_{\text{intra}} \leftarrow \sum_{k=1}^{K}\frac{1}{N_k}\sum_{n=1}^{N} m(\bar{T}_k \| T_n)\delta(k_n, k)$
until $I_{\text{old}} \geq I_{\text{intra}}$

This algorithm deals with the Kroneckner symbol $\delta(a, b)$ that equals 1 if $a = b$ and 0 otherwise.

3.2. Validation

Fig. 3(a) draws a projection of the scatter plot of training samples (triangle) in the α/β plane for a sub-band. The training samples are issued from the same texture class but with different orientations. The proposed K-means classifier is applied with $K = 2$. The red and blue squares correspond to the location of the two estimated centroids. In these experiments, the K-means classifier is applied on a manifold $\mathcal{M}$ of dimension 24[1]. Only one projection is displayed in Fig. 3(a).

[1]For the decomposition we use the Steerable Pyramid [13] with 2 scales and 6 orientations and 2 parameters (α and β) per sub-band

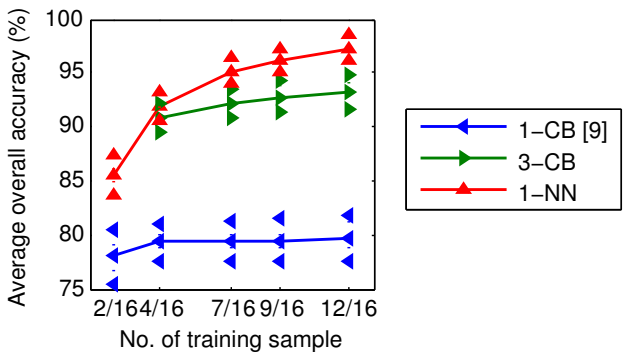


Fig. 4. Evolution of the average overall accuracy as a function of the number of training samples on the VisTex database.

Fig. 3 (b) shows some examples of texture image samples associated to both clusters. After performing a K-means classifier, each clusters are separated according to their principal orientation. The two centroids are representative of the diversity inside this texture class.

4. RESULTS

4.1. Context

To evaluate the performance of the supervised classification algorithms, the database is split into a training database and a disjoint testing database. Practically, N_{Tr} training samples are randomly selected for each texture class, the remaining sample are used as testing samples. Two databases are considered here: VisTex [14] with 40 classes and $N_{sa} = 16$ images per class (128×128 pixels), and Brodatz [15] with 13 classes and $N_{sa} = 112$ images per class (128×128 pixels). The VisTex database contains some texture images with different illumination conditions, while the Brodatz one exhibits a higher intra-class variability due to various viewing conditions such as rotated image. In the following, 100 Monte Carlo runs have been used to evaluate the performance of the different classifiers (overall accuracy and error bars).

4.2. Results and discussion

In this experiment, the stationary wavelet decomposition (with 2 scales) with Daubechies' filter db4 is considered. Fig. 4 draws the evolution of the average overall accuracy as a function of the number of training samples on the VisTex database for the nearest neighbour (1-NN in red), the one centroid [9] (1-CB in blue) and the proposed K-CB classifier with 3 centroids (3-CB in green). A gain of more than 10 points is observed when 3 centroids are considered instead of only 1. Hence, the proposed 3-CB classifier allows a better characterization of the intra-class diversity. Note also that performances of the 3-CB classifier are close to those of the 1-NN. However, the computation complexity is significantly lower with the K-CB classifier[2], since only K computations of the similarity measure between the query and the centroids are necessary while N_{Tr} are required for the 1-NN classifier.

[2]The computation of the centroids being done off-line for the K-CB classifier.

N_{Tr}	1-NN K	1-CB [9] $N_{Sa}/2$	K-CB $N_{Sa}/2$	1-NN $N_{Sa}/2$
VisTex	79 % ±2	73 % ±2	89 % ±2	94 % ±1
Brodatz	70 % ±3	70 % ±1	97 % ±2	99 % ±1

Table 1. Average overall accuracy for the different supervised classifiers on the VisTex and Brodatz databases.

Since computational considerations play a key role for any practical application, the 1-NN classifier cannot be used for large database. As a consequence, the K-CB classifier is a good trade-off between classification accuracy and computation complexity.

Table 1 displays the average overall accuracy for the different classifiers (1-NN, 1-CB and K-CB) on the VisTex and Brodatz databases (K equals respectively 3 and 10 for the two databases). Note that two 1-NN classifiers have been considered, one with K training samples per class (the same complexity as the K-CB) and one where half of the database is used for training. In this experiment, the steerable pyramid with 2 scales and 8 orientations has been used for the decomposition. As observed in Table 1, for the same computational complexity, the proposed K-CB classification exhibits a gain of about 10 and 27 points compared to the nearest neighbour classifier (1-NN) with $N_{Tr} = K$. Hence, an adapted selection of centroids with the K-means allows an increase of the overall accuracy. Note also, that the performance of the K-CB are very close to the 1-NN classifier with the same number of training data. In addition, the proposed K-CB classifier based on a parametric point of view outperforms a texton-based approach such as the one proposed by Varma and Zisserman in [2] which has an overall accuracy of about 82% on the VisTex database with 20 centroids.

5. CONCLUSION

In this paper, a K-centroids-based (K-CB) supervised classification has been introduced to handle the natural intra-class diversity of texture images. The proposed K-CB provides a good compromise between the k-NN classifier and 1-Centroid method proposed by Choy and Tong [9] in the framework of the stochastic parametric modeling. Moreover, we show also the superiority of our approach against the non-parametric one proposed by Varma and Zisserman [2] Further works will concern the development of a supervised classification algorithm with an adapted and automatic number of centroids per texture class (which will depend on the intra-class diversity).

6. REFERENCES

[1] A. Kibriya and E. Frank, "An empirical comparison of exact nearest neighbour algorithms," *Knowledge Discovery in Databases: PKDD 2007*, pp. 140–151, 2007.

[2] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International*

Journal of Computer Vision, vol. 62, pp. 61–81, 2005, 10.1023/B:VISI.0000046589.39864.ee.

[3] J. Puzicha, T. Hofmann, and J.M. Buhmann, "Non-parametric similarity measures for unsupervised texture segmentation and image retrieval," in *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, Jun 1997, pp. 267 –272.

[4] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *International Journal of Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001.

[5] O.G. Cula and K.J. Dana, "3D texture recognition using bidirectional feature histograms," *International Journal of Computer Vision*, vol. 59, no. 1, pp. 33–60, 2004.

[6] O.G. Cula, K.J. Dana, F.P. Murphy, and B.K. Rao, "Bidirectional imaging and modeling of skin texture," *Biomedical Engineering, IEEE Transactions on*, vol. 51, no. 12, pp. 2148–2159, 2004.

[7] S.G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 7, pp. 674–693, 1989.

[8] M.N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *Image Processing, IEEE Transactions on*, vol. 11, no. 2, pp. 146 –158, Feb 2002.

[9] S.-K. Choy and C.-S. Tong, "Supervised texture classification using characteristic generalized gaussian density," *Journal of Mathematical Imaging and Vision*, vol. 29, pp. 35–47, Aug. 31 2007.

[10] S. Amari and S.C. Douglas, "Why natural gradient?," in *Acoustics, Speech and Signal Processing, IEEE International Conference on*, May 1998, vol. 2, pp. 1213–1216.

[11] A. Schutz, Y. Berthoumieu, Turcu. F., C. Nafornita, and A. Isar, "Barycentric distribution estimation for texture clustering based on information-geometry tools," in *Electronics and Telecommunications (ISETC), 10th International Symposium on*, Nov. 2012.

[12] M. Novey, T. Adali, and A. Roy, "A complex generalized Gaussian distribution – characterization, generation, and estimation," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1427–1433, 2010.

[13] E.P. Simoncelli, W. T. Freeman, E.H. Aldelson, and D.J. Heeger, "Shiftable multiscale transforms," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 587–607, Mar. 1992.

[14] "MIT Vision and Modeling Group. Vision Texture.," Available: http://vismod.media.mit.edu/pub/VisTex.

[15] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover, New York, 1966.