

La matrice de co-occurrence des matrices de covariance : un nouveau modèle de codage pour la classification de textures

Ioana Ilea, Lionel Bombrun, Salem Said, Yannick Berthoumieu

▶ **To cite this version:**

Ioana Ilea, Lionel Bombrun, Salem Said, Yannick Berthoumieu. La matrice de co-occurrence des matrices de covariance : un nouveau modèle de codage pour la classification de textures. GRETSI, Sep 2017, Juan les Pins, France. hal-01629143

HAL Id: hal-01629143
https://hal.science/hal-01629143

Submitted on 6 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La matrice de co-occurrence des matrices de covariance : un nouveau modèle de codage pour la classification de textures

Ioana ILEA[1,2], Lionel BOMBRUN[1], Salem SAID[1], Yannick BERTHOUMIEU[1]

[1]Université de Bordeaux, Bordeaux INP, Laboratoire IMS, Groupe Signal et Image.

[2]Université Technique de Cluj-Napoca.

{ioana.ilea, lionel.bombrun, salem.said, yannick.berthoumieu}@ims-bordeaux.fr

Résumé – Cet article présente un nouveau modèle de codage pour la classification à partir d'un ensemble de matrices de covariance : la matrice de co-occurrence associée à un dictionnaire de matrices de covariance. Contrairement aux modèles de codage de l'état de l'art (SDMR, R-VLAD and VRF), un tel modèle local exploite la répartition spatiale des patchs. À partir du modèle de mélange génératif de distributions gaussiennes Riemanniennes, nous présentons ce modèle local. Une expérience sur la classification d'images de texture est ensuite effectuée sur les bases de textures VisTex et Outex_TC000_13 afin de évaluer le potentiel de la méthode proposée.

Abstract – This paper introduces a novel coding model for the classification from covariance matrice set: the co-occurrence matrix associated to a dictionary of covariance matrices. Contrary to state-of-the-art coding models (BoRW, R-VLAD and RFV), such local model exploits the spatial distribution of the patches. Starting from the generative mixture model of Riemannian Gaussian distributions, we introduce this local model. An experiment on texture image classification is then conducted on the VisTex and Outex_TC000_13 databases to evaluate its potential.

1 Introduction

La tâche de classification à partir du contenu "texture" des images de type matériel vise à attribuer une ou plusieurs étiquettes d'appartenance à une classe. Atteindre un tel objectif est l'un des problèmes les plus fondamentaux pour un large panel d'applications telles que l'inspection industrielle [1], l'indexation d'image [2], l'imagerie médicale [3, 4], la télédétection [5, 6], la reconnaissance d'objet ou de visage [7, 8]. Dans le cadre général de la classification supervisée, les techniques de codage des descripteurs exploitées dans la famille d'approches de type "sac-de-mots" ont largement prouvé leur efficacité comme le démontrent les travaux de la communauté. Ainsi, à partir d'un espace de descripteur donné, les techniques de type sac-de-mots consistent à générer dans un premier temps un dictionnaire composé par un ensemble fini de mots, suivi d'une étape de codage qui associe à chaque image une carte d'activation [9].

Dans le contexte de l'analyse de texture, des travaux récents [10, 11, 12, 13, 14] ont proposé des représentations compactes et sélectives à partir de descripteurs structurés et localisés que sont les matrices de covariance. La structure tient aux propriétés de symétrie et de définie positivité (SDP) des matrices de covariance locale (MCL). Considérant les propriétés intrinsèques à la géométrie Riemannienne propre à l'espace des matrices SDP, le contenu du présent article vise à l'étude des différentes techniques de codage fondées sur l'utilisation d'un dictionnaire de mots de type MCL dans l'objectif de la classification d'images texturées.

L'article est structuré comme suit. La section 2 introduit l'architecture fonctionnelle générale de la méthode de classification. Par la suite, les sections 3 et 4 sont dédiées à la présentation des deux principales étapes du procédé de génération du codage des contenus visuels, l'apprentissage du dictionnaire et le codage à proprement dite. Finalement, la section 5 présente les résultats obtenus dans le cadre d'une expérimentation menée sur plusieurs bases de données d'images texturées pour évaluer le potentiel du modèle de codage proposé qui étend le principe de co-occurrence dans l'espace des matrices de covariances.

2 Architecture fonctionnelle générale

La figure 1 présente l'architecture fonctionnelle de la méthode de classification fondée sur l'analyse statistique des descripteurs locaux.

1. Au cours de la première étape (appelée extraction des descripteurs locaux), un jeu de caractéristiques de bas niveau est calculé sur chaque image de la base de données. Ces descripteurs sont généralement calculés à partir d'un partitionnement spatial, régional et régulier de l'image par patchs. Ainsi, un ensemble de matrices caractéristiques (signatures) est obtenu pour chaque image de la base. Il s'agit par exemple des matrices de covariance calculées à partir de l'estimation des statistiques d'ordre deux sur les composantes de couleur et/ou de certaines composantes de dépendances spatiales (coefficients d'ondelettes, etc.).

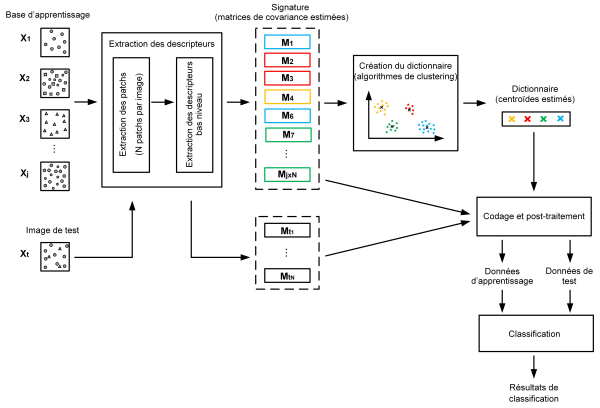


FIGURE 1 – Schéma fonctionnel de la méthode de classification fondée sur l'analyse statistique des descripteurs locaux.

2. La deuxième étape consiste à créer le dictionnaire de mots. Pour cela, un algorithme de regroupement (clustering) comme l'algorithme des k-moyennes ou de type espérance-maximisation (EM) pour les modèles de mélange est appliqué sur l'ensemble des échantillons du jeu de caractéristiques observées. En utilisant l'un de ces algorithmes, l'ensemble des observations est partitionné en un nombre prédéfini de groupes, chacun d'eux étant décrit par des paramètres tels que le centroïde du groupe, la dispersion et le poids associé. Ces paramètres estimés sont appelés mots-clés et sont regroupés dans le dictionnaire.
3. La troisième étape est la phase de codage. Au cours de cette étape, chaque ensemble de signature est projeté sur l'espace des mots-clés du dictionnaire. Pour cela, diverses approches ont été proposées dans de précédents travaux dans le cas des caractéristiques définies par un ensemble de matrices de covariance. Il s'agit du sac de mots Riemannien (SDMR) [12], des vecteurs de descripteurs Riemanniens agrégés localement (R-VLAD) [13] et des vecteurs Riemanniens de Fisher (VRF) [14]. Inspiré par le concept de matrices de co-occurrences des niveau de gris (GLCM) [15], largement exploité dans le contexte de la caractérisation de texture à partir des simples niveaux de gris, la principale contribution de notre article est de proposer une nouvelle approche de codage qui exploite l'arrangement spatial entre les matrices de covariance extraites. A la différence des SDMR, R-VLAD, VRF, clairement nous souhaitons exploiter l'arrangement spatial des images texturées.
4. Pour l'étape de classification finale, l'image de test est étiquetée dans la classe la plus similaire. En pratique, des classificateurs tels que l'algorithme des k-plus proches voisins, les machines à vecteurs de supports ou l'approche des forêts aléatoires sont généralement utilisés.

Les deux sections suivantes se concentrent respectivement sur la deuxième et la troisième étape du schéma fonctionnel présenté figure 1.

3 Apprentissage du dictionnaire

Soit l'ensemble des N observations $\mathcal{M} = \{\mathbf{M}_n\}_{n=1:N}$, avec $\mathbf{M}_n \in \mathcal{P}_m$ modélisé par un mélange de K distributions gaussiennes Riemanniennes. Sous l'hypothèse d'indépendance, la densité de probabilité (ddp) de $\mathcal{M}$ est donnée par :

$$p(\mathcal{M}|\theta) = \prod_{n=1}^{N} p(\mathbf{M}_n|\theta) = \prod_{n=1}^{N}\sum_{k=1}^{K} \varpi_k p(\mathbf{M}_n|\bar{\mathbf{M}}_k, \sigma_k), \quad (1)$$

où $p(\mathbf{M}_n|\bar{\mathbf{M}}_k, \sigma_k)$ est la ddp de la loi gaussienne Riemannienne (DGR) définie sur la varitée $\mathcal{P}_m$ des matrices réelles, symétriques et définies positives de taille $m \times m$ [16]. La ddp de la DGR, respectant l'élément de volume Riemannien, a été introduit dans [16] comme :

$$p(\mathbf{M}|\bar{\mathbf{M}}, \sigma) = \frac{1}{Z(\sigma)} \exp\Big\{ - \frac{d^2(\mathbf{M}, \bar{\mathbf{M}})}{2\sigma^2} \Big\}, \quad (2)$$

où $Z(\sigma)$ est le coefficient de normalisation indépendant du centroïde $\bar{\mathbf{M}}$ et $d(\cdot)$ est la distance Riemannienne entre deux matrices de covariance définie par $d(\mathbf{M}_1, \mathbf{M}_2) = \left[\sum_i (\ln \lambda_i)^2\right]^{\frac{1}{2}}$, avec λ_i, $i = 1, \ldots, m$ les valeurs propres de $\mathbf{M}_1^{-1}\mathbf{M}_2$.

Le dictionnaire est donc composé des K mots-clés qui sont les paramètres des distributions de chaque composante dans le modèle de mélange défini dans (1). Il comprend le poids du mélange ϖ_k, le centroïde $\bar{\mathbf{M}}_k$ et le paramètre de dispersion σ_k. En pratique, les paramètres du modèle de mélange sont estimés en considérant un algorithme k-means ou EM intrinsèque. Pour plus d'informations sur la mise en œuvre de l'algorithme EM, le lecteur intéressé trouvera les informations complémentaires dans [17].

Dans la partie expérimentale, afin de s'assurer que chaque classe est représentée par un ensemble de mots-clés, une stratégie intra-classe est adoptée pour estimer le dictionnaire. Cela signifie qu'un modèle de mélange est appris pour chaque classe dans l'ensemble d'apprentissage et le dictionnaire final est obtenu en concaténant chaque mot-clés de toutes les classes.

Une fois le dictionnaire créé, une étape de codage est exploitée pour coder chaque image de la base à partir des mots-clés. Pour cela, différentes stratégies peuvent être adoptées telles que le sac de mots Riemanniens (SDMR) [12], les vecteurs de descripteurs Riemanniens agrégés localement (R-VLAD) [13], les vecteurs Riemanniens de Fisher (VRF) [14] et les co-occurrences des covariances (CoC). La section suivante décrit chacune de ces stratégies.

4 L'étape de codage

A partir de l'ensemble $\mathcal{M} = \{\mathbf{M}_n\}_{n=1:N}$, avec $\mathbf{M}_n \in \mathcal{P}_m$, l'objectif de l'étape de codage est de projeter l'ensemble $\mathcal{M}$ sur les éléments du dictionnaire.

4.1 Sac de mots Riemanniens (SDMR)

La méthode des sacs de mots (SDM) est probablement l'une des méthodes les plus utilisées pour coder une image. Cette

approche est utilisée dans un large panel d'applications en vision par ordinateur et en traitement du signal et de l'image. Cependant, lorsque les descripteurs vivent dans un espace non-euclidien, tel que celui des matrices de covariance de taille $m \times m$, l'approche doit être réadaptée afin de tenir compte de la géométrie particulière de cet espace. Pour cela, les sacs de mots Riemanniens (SDMR) [12] et les sacs de mots log-Euclidien (SDM-LE) [18] ont été introduits.

Dans ces modèles, l'espace des descripteurs est partitionné en K régions de Voronoï. Ensuite, chaque observation $\mathbf{M}_n$ est affectée au cluster k, pour $k = 1, \ldots, K$ par :

$$\arg\max_k \ \varpi_k \ p(\mathbf{M}_n|\bar{\mathbf{M}}_k, \sigma_k), \tag{3}$$

où $p(\mathbf{M}_n|\bar{\mathbf{M}}_k, \sigma_k)$ est la ddp de la DGR définie par (2). En pratique, l'hypothèse d'homoscédasticité est généralement admise (*i.e.* $\sigma_k = \sigma \ \forall k \in [1, K]$) et les mots clés sont considérés équiprobables (*i.e.* $\varpi_k = 1/K$). De plus, pour chaque image dans la base de données, la signature est déterminée en calculant l'histogramme du nombre d'occurrences de chaque mot clé du dictionnaire.

Le modèle SDMR est une méthode simple mais efficace. Néanmoins, il souffre d'un inconvénient majeur, il ne compte que le nombre de descripteurs locaux attribués à chaque région de Voronoï. Afin d'améliorer les performances de classification, certains auteurs ont proposé des modèles qui incluent des statistiques du second ordre. C'est le cas des modèles R-VLAD et VRF présentés ci-après.

4.2 Vecteurs de descripteurs Riemanniens agrégés localement (R-VLAD)

La version Riemannienne des descripteurs VLAD, appelée R-VLAD a été développée dans [13]. Pour chaque cluster c_k, $k \in [1, K]$, un vecteur contenant les différences entre le centroïde $\bar{\mathbf{M}}_k$ et chaque observation $\mathbf{M}_i$ est calculé. Ensuite, la somme des différences concernant chaque cluster c_k est déterminée par :

$$\mathbf{v}_k = \sum_{\mathbf{M}_i \in c_k} \mathrm{Log}_{\bar{\mathbf{M}}_k} \mathbf{M}_i, \tag{4}$$

où $\mathrm{Log}(\cdot)$ est l'application log Riemannienne [19]. Ce modèle considère deux hypothèses :

— chaque observation $\mathbf{M}_i$ appartient à un seul cluster c_k.

— une hypothèse d'homoscédasticité, $\sigma_k = \sigma \ , \forall k = 1, \ldots, K$

Afin de relaxer ces deux hypothèses, les vecteurs Riemannien de Fisher ont été proposés dans [14].

4.3 Vecteurs Riemannien de Fisher (VRF)

À partir du modèle génératif introduit dans (1), le modèle VRF est obtenu en calculant la dérivée de la log-vraisemblance par rapport aux paramètres de la distribution [14].

$$\frac{\partial \log p(\mathcal{M}|\theta)}{\partial \bar{\mathbf{M}}_k} = \sum_{n=1}^{N} \gamma_k(\mathbf{M}_n) \ \sigma_k^{-2} \ \mathrm{Log}_{\bar{\mathbf{M}}_k}(\mathbf{M}_n), \tag{5}$$

$$\frac{\partial \log p(\mathcal{M}|\theta)}{\partial \sigma_k} = \sum_{n=1}^{N} \gamma_k(\mathbf{M}_n) \Big\{ - \frac{Z'(\sigma_k)}{Z(\sigma_k)} + \frac{d^2(\mathbf{M}_n, \bar{\mathbf{M}}_k)}{\sigma_k^3} \Big\}, \tag{6}$$

$$\frac{\partial \log p(\mathcal{M}|\theta)}{\partial \alpha_k} = \sum_{n=1}^{N} \left[\gamma_k(\mathbf{M}_n) - \varpi_k \right], \tag{7}$$

où $Z'(\sigma)$ est la dérivée du coefficient de normalisation $Z(\sigma)$ par rapport au paramètre de dispersion σ. Le terme $\gamma_k(\mathbf{M}_n)$ correspond à la contribution de l'observation $\mathbf{M}_n$ au cluster c_k, il est défini par :

$$\gamma_k(\mathbf{M}_n) = \frac{\varpi_k \ p(\mathbf{M}_n|\bar{\mathbf{M}}_k, \sigma_k)}{\sum_{j=1}^{K} \varpi_j \ p(\mathbf{M}_n|\bar{\mathbf{M}}_j, \sigma_j)}, \tag{8}$$

Notez que la paramétrisation suivante des poids dans le modèle de mélange est utilisée afin de respecter les contraintes de positivité et de somme à un des poids :

$$\varpi_k = \frac{\exp(\alpha_k)}{\sum_{j=1}^{K} \exp(\alpha_j)}. \tag{9}$$

Comme expliqué dans [14], les descripteurs R-VLAD sont un cas particulier des descripteurs VRF. Ils sont obtenus à partir des VRF lorsque seule la dérivée par rapport à $\bar{\mathbf{M}}_k$ est considérée (5) et lorsque les deux hypothèses présentées dans la section 4.2 sont admises.

4.4 Co-occurrences des covariances (CoC)

Ces trois modèles (SDMR, R-VLAD et VRF) ont montré des résultats prometteurs, mais toutes ces méthodes n'utilisent pas une caractéristique principale : la répartition spatiale des patchs. Inspiré par le concept de GLCM pour l'analyse de texture [15, 20], nous introduisons une nouvelle approche de codage : les co-occurrences de covariances (CoC).

Pour un dictionnaire de K mots, la matrice (de taille $K \times K$) de co-occurrence des matrices de covariance d'une image I décrit les interactions spatiales entre les matrices de covariances $\mathbf{M}_n$ calculées sur des patchs séparés d'une distance (Δ_x, Δ_y). L'élément $\mathbf{C}_{\Delta_x,\Delta_y}(k, l)$ de cette matrice de co-occurrence contient le nombre de fois qu'une matrice de covariance appartenant au cluster l apparait dans le voisinage $\mathcal{N}_{\Delta_x,\Delta_y}$ d'une matrice de covariance appartenant au cluster k :

$$\mathbf{C}_{\Delta_x,\Delta_y}(k, l) = \sum_{\mathbf{M}_n \in \mathcal{M}} \sum_{\mathbf{M}_p \in \mathcal{N}_{\Delta_x,\Delta_y}(\mathbf{M}_n)} \begin{cases} 1 & \text{si } \mathbf{M}_n \in c_k \\ & \text{et } \mathbf{M}_p \in c_l \\ 0 & \text{sinon.} \end{cases} \tag{10}$$

Une fois les matrices de co-occurrence calculées, la mesure de similarité entre deux CoC $\mathbf{C}^1$ et $\mathbf{C}^2$ est obtenue par leur intersection donnée par :

$$\sum_{k=1}^{K} \sum_{l=1}^{K} min \left(\mathbf{C}^1_{\Delta_x,\Delta_y}(k, l), \ \mathbf{C}^2_{\Delta_x,\Delta_y}(k, l) \right) \tag{11}$$

5 Classification d'images texturées

Dans cette partie, nous présentons une application à la classification d'images texturées. L'objectif de cette expérimentation est d'évaluer le potentiel des quatre modèles de codage présentés dans la partie 4 : SDMR, R-VLAD, VRF et CoC.

Pour cette expérience, deux bases de données sont considérées : les bases de données VisTex [21] et Outex_TC000_13 [22]. La base VisTex est composée de 40 classes de texture. Chaque classe est représentée par un ensemble de 64 images de taille 64×64 pixels. La base de données Outex_TC000_13 contient 68 classes de texture, chaque classe étant représentée par un ensemble de 20 images de taille 128×128 pixels. Pour les deux bases de données, les étapes d'extraction et de classification des descripteurs illustrées dans la figure 1 sont similaires. Nous considérons le même protocole que celui présenté dans [14]. Tout d'abord, les matrices de covariance sont calculées sur des fenêtres glissantes de taille 15×15 pixels. Ces matrices de covariance décrivent les interactions entre les intensités de l'image $I(x, y)$ et les normes des dérivées de premier et second ordre de $I(x, y)$ dans les deux directions x et y [10]. Ensuite, une fois que les images sont codées avec l'un des quatre modèles présentés (SDMR, R-VLAD, VRF ou CoC), un classifieur SVM avec un noyau gaussien est utilisé pour l'étape de classification. En pratique, le paramètre de dispersion de ce noyau est optimisé par une procédure de validation croisée sur l'ensemble d'apprentissage.

Le tableau 1 présente les résultats de classification en terme de taux de reconnaissance moyens obtenus sur les bases VisTex et Outex TC000_13 pour les 4 modèles de codage. Pour le modèle VRF, la contribution de chaque paramètre (centroïde, dispersion, poids) est analysée. Par exemple, la ligne "VRF : ϖ" montre le taux de classification lorsque seule la dérivée par rapport aux poids est prise en compte pour extraire le VRF, ... Pour le modèle CoC, un 8-voisinage avec un déplacement de 2 pixels entre les patchs est considéré. Comme illustré dans le tableau 1, les meilleurs résultats de classification sont obtenus pour le modèle proposé CoC qui exploite l'information sur la répartition spatiale des patchs. Un gain d'environ 1% à 2% est observé sur les bases VisTex et Outex_TC000_13 par rapport aux autres modèles de codage de l'état de l'art (SDMR, R-VLAD et VRF).

6 Conclusion

Ce papier a introduit un nouveau modèle local pour la classification d'images sur l'espace des matrices de covariance. Sur la base du concept des matrices de co-occurrence, ce modèle local exploite la répartition spatiale des patchs, ce qui permet d'améliorer les performances de classification par rapport aux modèles de codage standard (SDMR, R-VLAD et VRF).

L'extension de ces travaux portera sur la définition d'un modèle de codage à partir d'une version floue des matrices de co-occurrence [23].

Méthode	VisTex	Outex_TC000_13
SDMR [12]	86.87 ± 1.56	83.86 ± 1.41
R-VLAD [13]	87.91 ± 0.74	83.13 ± 1.50
VRF : ϖ [14]	89.42 ± 0.63	84.97 ± 0.87
VRF : σ [14]	79.32 ± 1.38	76.75 ± 1.48
VRF : $\bar{\mathbf{M}}$ [14]	87.77 ± 0.84	84.20 ± 0.65
VRF : σ, ϖ [14]	82.13 ± 1.19	79.35 ± 1.39
VRF : $\bar{\mathbf{M}}, \varpi$ [14]	88.73 ± 0.89	84.57 ± 0.54
VRF : $\bar{\mathbf{M}}, \sigma$ [14]	89.43 ± 0.79	84.01 ± 0.65
VRF : $\bar{\mathbf{M}}, \sigma, \varpi$ [14]	89.80 ± 0.57	84.22 ± 0.62
CoC	$\mathbf{91.08 \pm 0.61}$	$\mathbf{85.19 \pm 0.97}$

TABLE 1 – Performances de classification sur VisTex et Outex.

Références

[1] C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz, "Exploring features in a bayesian framework for material recognition," in *IEEE CVPR*, 2010, pp. 239–246.

[2] P.S. Hiremath and Jagadeesh Pujari, "Content based image retrieval using color, texture and shape features," *ADCOM*, pp. 780–784, 2007.

[3] R. de Luis-García, C-F. Westin, and C. Alberola-López, "Gaussian mixtures on tensor fields for segmentation : applications to medical imaging," *Comput. Med. Imag. Grap.*, vol. 35, no. 1, pp. 16–30, 2011.

[4] P. Cirujeda, Y. D. Cid, H. Muller, D. Rubin, T. A. Aguilera, B. W. Loo, M. Diehn, X. Binefa, and A. Depeursinge, "A 3-D Riesz-covariance texture model for prediction of nodule recurrence in Lung CT," *IEEE Trans. Med. Imaging*, vol. 35, no. 12, pp. 2620–2630, 2016.

[5] C. Zhu and X. Yang, "Study of remote sensing image texture analysis and classification using wavelet," *Int J Remote Sens.*, vol. 19, no. 16, pp. 3197–3203, 1998.

[6] O. Regniers, L. Bombrun, V. Lafon, and C. Germain, "Supervised classification of very high resolution optical images using wavelet-based textural features," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3722–3735, June 2016.

[7] N.-G. Vu, H.M. Dee, and A. Caplier, "Face recognition using the POEM descriptor," *Pattern Recogn.*, vol. 45, no. 7, pp. 2478–2488, July 2012.

[8] T.P. Nguyen, N.-S. Vu, and A. Manzanera, "Statistical binary patterns for rotational invariant texture classification," *Neurocomputing*, vol. 173, pp. 1565–1577, 2016.

[9] F. Perronnin, J. Sánchez, and T. Mensink, *Improving the Fisher kernel for large-scale image classification*, vol. 6314 of *Lecture Notes in Computer Science*, pp. 143–156, Springer Berlin Heidelberg, 2010.

[10] O. Tuzel, F. Porikli, and P. Meer, *Region Covariance : A Fast Descriptor for Detection and Classification*, vol. 3952 of *Lecture Notes in Computer Science*, pp. 589–600, Springer Berlin Heidelberg, 2006.

[11] S. Jayasumana, R.I. Hartley, M. Salzmann, H. Li, and M.T. Harandi, "Kernel methods on the Riemannian manifold of symmetric positive definite matrices," in *IEEE CVPR*, 2013, pp. 73–80.

[12] M. Faraki, M. T. Harandi, A. Wiliem, and B. C. Lovell, "Fisher tensors for classifying human epithelial cells," *Patt. Recogn.*, vol. 47, no. 7, pp. 2348 – 2359, 2014.

[13] M. Faraki, M. T. Harandi, and F. Porikli, "More about VLAD : A leap from Euclidean to Riemannian manifolds," in *IEEE CVPR*, June 2015, pp. 4951–4960.

[14] I. Ilea, L. Bombrun, C. Germain, R. Terebes, M. Borda, and Y. Berthoumieu, "Texture image classification with Riemannian Fisher vectors," in *IEEE ICIP*, 2016, pp. 3543–3547.

[15] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[16] S. Said, L. Bombrun, Y. Berthoumieu, and J. H. Manton, "Riemannian gaussian distributions on the space of symmetric positive definite matrices," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 2153–2170, April 2017.

[17] S. Said, L. Bombrun, and Y. Berthoumieu, "Texture classification using Rao's distance on the space of covariance matrices," in *GSI*, 2015.

[18] M. Faraki, M. Palhang, and C. Sanderson, "Log-euclidean bag of words for human action recognition," *IET Computer Vision*, vol. 9, no. 3, pp. 331–339, 2015.

[19] N. J. Higham, *Functions of Matrices : Theory and Computation*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[20] G. Liu, G. S. Xia, W. Yang, and L. Zhang, "Texture analysis with shape co-occurrence patterns," in *ICPR*, Aug 2014, pp. 1627–1632.

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

[22] "Outex Texture Database," Center for Machine Vision Research of the University of Oulu. Available : http ://www.outex.oulu.fi/index.php ?page=classification.

[23] A. Ledoux, O. Losson, and L. Macaire, "Texture classification with fuzzy color co-occurrence matrices," in *IEEE ICIP*, Sept 2015, pp. 1429–1433.