Advances in Texture Analysis for Emphysema Classification

Rodrigo Nava[1], J. Victor Marcos[2], Boris Escalante-Ramírez[1], Gabriel Cristóbal[2], Laurent U. Perrinet[3], and Raúl San José Estépar[4]

[1] Posgrado en Ciencia e Ingeniería de la Computación, Universidad Nacional Autónoma de México, Mexico City, Mexico
[2] Instituto de Óptica, Spanish National Research Council (CSIC), Serrano 121, Madrid 28006, Spain
[3] INCM, UMR6193, CNRS & Aix-Marseille University, 31 ch. Aiguier, 13402 Marseille Cedex 20, France
[4] Brigham and Women's Hospital, Harvard Medical School, Boston MA, United States

Abstract. In recent years, with the advent of High-resolution Computed Tomography (HRCT), there has been an increased interest for diagnosing Chronic Obstructive Pulmonary Disease (COPD), which is commonly presented as emphysema. Since low-attenuation areas in HRCT images describe different emphysema patterns, the discrimination problem should focus on the characterization of both local intensities and global spatial variations. We propose a novel texture-based classification framework using complex Gabor filters and local binary patterns. We also analyzed a set of global and local texture descriptors to characterize emphysema morphology. The results have shown the effectiveness of our proposal and that the combination of descriptors provides robust features that lead to an improvement in the classification rate.

Keywords: Co-occurrence matrices, Emphysema, Gabor filters, LBP, Sparsity, Tchebichef, Texture analysis.

1 Introduction

COPD is a progressive and irreversible lung condition, which is characterized by tissue damage. It hinders air from passing through airpaths and causes that alveolar sacs lose their elastic quality, increasing the risk of death. COPD can manifest as either emphysema, bronchitis or both; the former is the most common manifestation that destroys lung parenchyma [1].

Literature recognizes three types of emphysema: *i*) **Paraseptal** (PS) or distal acinar emphysema, which is characterized by destruction of distal airway structures, alveolar ducts, and alveolar sacs. The process is localized around the pleura; *ii*) **Panlobular** (PL) or panacinar emphysema destroys uniformly the entire alveolus, it is predominant in the lower half of the lungs; and *iii*) **Centrilobular** (CL) or centriacinar emphysema, which is the most common type

J. Ruiz-Shulcloper and G. Sanniti di Baja (Eds.): CIARP 2013, Part II, LNCS 8259, pp. 214–221, 2013.

of emphysema. It begins in the respiratory bronchioli and spreads peripherally, most damage is usually contained to the upper half of the lungs.

Spirometry is the gold standard criterion to establish a diagnosis of emphysema. It measures the volume of air that a patient is able to expel from lungs after a maximal inspiration. Nevertheless, this method does not allow to discriminate pathological subphenotypes of emphysema. On the other hand, HRCT is a minimally invasive imaging technique capable of providing both high-contrast and high-resolution details of lungs and airways; it has shown its potential for identifying changes in lung parenchyma and abnormalities associated with emphysema.

Hayhurst et al. [2] showed that Hounsfield Unit (HU) frequency distributions in patients who had CL differed significantly from patients with Normal Tissue (NT). Low-attenuation areas in HRCT images have been found to represent macroscopic and microscopic changes due to emphysema. Such areas are determined using the density mask method, which measures the amount of emphysematous lung by calculating the percentage of voxels lesser than a threshold; commonly, the threshold lies somewhere between -910 and -980 HU.

Texture-based classification of lung HRCT images may provide new insights towards the construction of a reliable computer-aided diagnosis system. New methods include features extracted using local binary patterns [3]. A simpler alternative based on kernel density estimation of local histograms has been proposed in [4]. A different approach was proposed in [5] where the authors used meta-data to label lung samples, whereas in [6], the Riesz transform was used to obtain textural features in interstitial lung abnormalities but it has not been tested in analysis of emphysema subtypes. However, researchers have analyzed texture in HRCT images using simple descriptors. In this work, we claim that the combination of both global and local descriptors will provide robust features because global characteristics and local information are encode simultaneously. Thus, an improvement in the classification rate can be attained.

This paper is organized as follows: the datasets are described in Section 2. In Section 3 we defined a set of global and local descriptors used in the present study and provided their mathematical foundations. The results are presented in Section 4. Finally, our work is summarized in Section 5.

2 Material

We used two datasets labeled by experienced pulmonologists: the **Bruijne and Sørensen dataset** (BS) was provided by Prof. Dr. Bruijne and Dr. Sørensen [3]. It consists of 168 non-overlapping annotated ROIs of size 61×61 pixels and belong to three types of patterns: NT=59, CL=50, and PS=59; and **Brigham and Women's Hospital dataset** (BWH). This dataset was provided by researchers from the Brigham and Women's Hospital using a subset of the COPDGene study [4]. 1337 ROIs from 267 CT scans were randomly selected; the distribution per pattern is: NT=370, PS=184, PL=148. BWH includes three subtypes of CL patterns (mild, moderate, and severe): CL1=170, CL2=287, and CL3=178

respectively. The size of the samples was chosen to fit the physical extent of emphysema within the secondary lobule corresponding to 31×31 pixels.

3 Methods

We propose the combination of Complex Gabor Filters (CGF) and Local Binary Patters (LBP) for a better characterization of emphysema; the former are global descriptors, whereas the latter are local descriptors. Additionally, a wide set of texture descriptors have been analyzed. To assign a given patch to one of several emphysema patterns, we used a methodology composed of three stages: *i*) feature extraction with global and local descriptors; *ii*) dimensionality reduction using Kernel-Fisher Discriminant Analysis (KFDA); and *iii*) classification with k-Nearest Neighbors (kNN). In the following paragraphs we summarize the main characteristics of the descriptors used in the current study.

Complex Gabor Filters [7] are defined as the product of Gaussian functions and complex sinusoids. They are band-pass filters that constitute a complete but non-orthogonal basis set and their shape match with psychophysical properties of receptive fields [8]. They can be divided into two parts: $g_e(x,y) = K\exp\{-\frac{1}{2}(\frac{\tilde{x}^2+\gamma^2\tilde{y}^2}{\alpha^2})\}\cos(2\pi u_0\tilde{x})$, which is an even filter, whereas $g_o(x,y) = K\exp\{-\frac{1}{2}(\frac{\tilde{x}^2+\gamma^2\tilde{y}^2}{\alpha^2})\}\sin(2\pi u_0\tilde{x})$ is an odd filter. K represents a normalizing constant, u_0 is the central frequency, (α,γ) are the constants of the Gaussian envelope along x and y-axes respectively. $\tilde{x} = x\cos\theta - y\sin\theta$, $\tilde{y} = x\sin\theta + y\cos\theta$, and θ denotes the orientation. Further filtering parameters were tuned by following the design constraints recommended in [9].

We used a bank made of 24 filters distributed in 4 scales (s) and 6 orientations; for each of them, we computed $E_{(s,\theta)} = I(x,y) \star g_{e(s,\theta)}(x,y)$ and $O_{(s,\theta)} = I(x,y) \star g_{o(s,\theta)}(x,y)$ where $I(x,y)$ is the given patch and the $\star$ indicates convolution. Then, we extracted the magnitude coefficients $(M_{(s,\theta)}(x,y))$ as:

$$M_{(s,\theta)}(x,y) = \sqrt{E^2_{(s,\theta)}(x,y) + O^2_{(s,\theta)}(x,y)} \tag{1}$$

Since $M_{(s,\theta)}(x,y)$ are considered as random variables, we extracted the mean (μ), the standard deviation (σ), the skewness (Υ), and the kurtosis (Ψ) from them to characterize the response of any image and build a feature vector, $\overline{f_{CGF}}$, as follows:

$$\begin{aligned}\overline{f_{CGF}} = \big[&\mu_{(0,0)},\, \sigma_{(0,0)},\, \Upsilon_{(0,0)},\, \Psi_{(0,0)},\, \ldots,\\ &\mu_{(s-1,\theta-1)},\, \sigma_{(s-1,\theta-1)},\, \Upsilon_{(s-1,\theta-1)},\, \Psi_{(s-1,\theta-1)}\big]\end{aligned} \tag{2}$$

Log-Gabor Filters (LGF) [10] are defined in frequency domain as Gaussian functions shifted from the origin; they have a null DC component and can be split into radial and angular filters: $\hat{G}(\rho,\theta) = \exp\{-\frac{1}{2}[\frac{\log(\frac{\rho}{u_0})}{\log(\frac{\alpha_\rho}{u_0})}]^2\}\exp\{-\frac{1}{2}[\frac{(\theta-\theta_0)}{\alpha_\theta}]^2\}$, where (ρ,θ) represent the polar coordinates, u_0 is the central frequency, θ_0 is the orientation angle, α_ρ and α_θ determine the scale and the angular bandwidth respectively. We applied setting recommendations that appear in [9] and computed the feature vectors, $\overline{f_{LGF}}$, by convolving a bank of 24 log-Gabor filters

distributed in 4 scales and 6 orientations with the input images and then we followed the procedure presented in Eq. (2).

Sparse Gabor Coding (SGC). Gabor filters provide redundant representations, which may hamper classification tasks. As proposed first by [11], this problem may be solved using a greedy algorithm. This approach corresponds to first choosing a single filter, Φ_i, that best fits the image, $I(x, y)$, along with a suitable coefficient a_i, such that the *single* source $a_i\Phi_i$ is a good match to the image: $i = \arg\max_j(\langle \frac{I(x,y)}{\|I(x,y)\|}, \frac{\Phi_j}{\|\Phi_j\|} \rangle)$, where $\langle \cdot, \cdot \rangle$ represents the inner product.

The associated coefficient is the scalar projection: $a_i = \langle I(x, y), \frac{\Phi_i}{\|\Phi_i\|^2} \rangle$. Knowing this choice, the image can be decomposed as: $I(x, y) = a_i\Phi_i + \mathbf{R}$ where $\mathbf{R}$ is the residual image. We repeat this 2-step process on the residual until some stopping criterion is met. This procedure is known as the matching pursuit algorithm, which has proven to be a good approximation for natural images [12]. Measuring the ratio of extracted energy in the images, 256 edges were on average enough to extract 90% of the energy of whitened images on all datasets. We thus used this set of sparse coefficients as the input vector for the classification framework.

Gray-level Co-occurrence Matrices (GLCM) were proposed by Haralick [13]. They evaluate spatial relationship among gray levels. Each pixel in an image $I(x, y)$ is assigned to one of N_g gray levels. The GLCM matrix consist of a set of $\{P_{ij}|i, j = 1, \ldots, N_g\}$ values. P_{ij} represents the number of occurrences of two pixels with gray levels i and j separated by a distance d in the direction of the angle θ. The GLCM's elements are normalized, providing the relative frequency of occurrence for a pair of gray levels.

The element $p\,(i, j)$ denotes the probability of finding the pair of levels (i, j) in the image, which is obtained as: $p(i, j) = P_{ij}(\sum_{i,j}^{N_g} P_{ij})^{-1}$. 10 features were chosen to capture texture properties: energy, contrast, correlation, homogeneity, entropy, autocorrelation, dissimilarity, cluster shade, cluster prominence, and maximum probability. In our study, N_g was set to 8 according previous works focused on texture analysis [14]. d was set to 1 while four different angle values were assessed: $0, 45, 90$, and 135 degrees. Thus, a total of 40 descriptors (10 statistical features for each of the four orientations) were obtained for each texture.

Discrete Tchebichef Moments (DTM) [15] are computed by projecting the image $I(x, y)$ onto the set of Tchebichef polynomial kernels. DTM provides a unique representation of the image in the spanned Tchebichef space. The moment $T_{pq}\,(p, q = 0,\, 1,\, \ldots,\, N - 1)$ of order $s = p + q$ is defined as:

$$T_{pq} = \frac{1}{\tilde{\rho}(p, N)\tilde{\rho}(p, N)} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \tilde{t}_p(x)\tilde{t}_q(y)I(x, y) \tag{3}$$

where $\tilde{t}_p(x)$ and $\tilde{t}_q(x)$ are scaled Tchebichef polynomials and $\rho(n, N)$ is its squared norm. T_{pq} quantifies the correlation between the image, $I(x, y)$, and the kernel $\tilde{t}_p(x)\tilde{t}_q(y)$. Hence, this magnitude will be higher for images characterized by repetitive patterns occurring at a similar rate to the kernel. The following feature evaluates the similarity between the image and the varying

patterns implemented by s-order Tchebichef kernels: $T(s) = \sum_{p+q=s} |T_{pq}|$, ($s = 0, 1, \ldots, 2N-2$). The analysis based on DTM yields a feature vector of length $2N-1$ to describe texture attributes.

Local Binary Patterns [16] are based on the idea that textural properties within homogeneous regions can be mapped into patterns, which represent micro-features. LBP uses a 3×3 square mask called "texture spectrum". The values in the square mask are compared with the central pixel, those ones lesser are labeled with "0" otherwise they are labeled with "1". The labeled pixels are multiplied by a fixed weighting function according with their positions to form a chain. Afterward, the values of the eight pixels are summed to obtain a label: $LBP_{P,R}(g_c) = \sum_{p=0}^{P-1} s(g_p - g_c)2^p$, where $\{g_p | p = 0, \ldots, P-1\}$ are the values of the neighbors. The comparison function $s(x)$ is defined as a Heaviside function: $s(x) = \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases}$

Uniform Local Binary Patterns ($LBP^{uni}_{P,R}$) [17]. Over 90% of LBP patterns can be described with few spatial transitions, which are the changes (0/1) in a pattern chain. Ojala introduced the measure $U(LBP_{P,R}(g_c)) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)|$, which corresponds to the number of spatial transitions. So that, the uniform LBP ($LBP^{uni}_{P,R}$) can be obtained as:

$$LBP^{uni}_{P,R}(g_c) = \begin{cases} \sum_{p=0}^{P-1} s(g_p - g_c) \text{ if } U(LBP_{P,R}(g_c)) \leq 2 \\ P+1 \qquad\qquad\qquad\qquad \text{otherwise} \end{cases} \tag{4}$$

after the process is completed; a labeled image, $L(x, y)$, is generated and the pixel-wise information is encoded as a histogram, H_i.

3.1 Multi-class Kernel Fisher Discriminant Analysis

It must be considered that the size of a training set should be exponentially increased with the dimensionality of the input space. Since the previous methods generate high-dimensional feature vectors and a limited dataset is available, we used KFDA [18], which maps original data into a new feature space preventing overfitting. Let $\boldsymbol{X}^1 = \{x_1^1, x_2^1, \ldots, x_{l_1}^1\}, \ldots, \boldsymbol{X}^C = \{x_1^C, x_2^C, \ldots, x_{l_C}^C\}$ be feature vectors from C classes and let $\boldsymbol{K}(m, n)$ be the kernel matrix defined as $\boldsymbol{K}(m, n) = k(X_m, X_n)$ where $\boldsymbol{X} = \bigcup_{i=1}^{C} \boldsymbol{X}^i$. We used the Gaussian kernel, $k(x, y) = \exp\{-\frac{1}{2}\frac{\|x-y\|^2}{a^2}\}$, $a = 333$.

The "between scatter matrix" is defined by $\boldsymbol{P} = \sum_{j=1}^{C} l_j(\mu_j - \mu)(\mu_j - \mu)^T$ with $\mu_j = \frac{1}{l_j}\sum_{\forall n \in X^j} \boldsymbol{K}(m, n)$ and $\mu = \frac{1}{l}\sum_{\forall n} \boldsymbol{K}(m, n)$. The "within class scatter matrix" is defined by $\boldsymbol{Q} = \boldsymbol{K}\boldsymbol{K}^T - \sum_{j=1}^{C} l_j \mu_j \mu_j^T$; since $\boldsymbol{Q}$ must be a positive definite matrix, we used $\boldsymbol{Q} = \boldsymbol{Q} + r\boldsymbol{I}$ to guarantee that $\boldsymbol{Q}$ is positive definite. Finally, α^* is built with the $C-1$ largest eigenvalues of $\boldsymbol{Q}^{-1}\boldsymbol{P}$ and the projection can be computed as: $y = \boldsymbol{K}\alpha^*$.

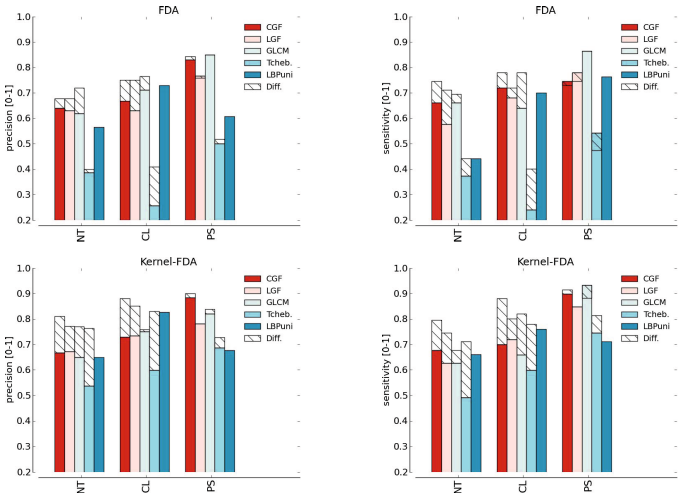

Fig. 1. BS classification rates (Three classes). The first row shows the results using FDA whereas in the second row, the results using KFDA are shown. Note that in almost all the cases, the extended methods ($Diff$), which are built by concatenating a single descriptor and its corresponding $LBP^{uni}_{P,R}$ histogram, achieved higher rates.

4 Experiments and Results

Parameter selection is a fundamental step in any classification problem; we used 10-fold cross-validation to estimate global parameters resulting in $k = 20$ neighbors in the kNN classifier as the best case. Then, we applied leave-one-out cross-validation to measure Specificity (Sp), Sensitivity (S), and Precision (P). We carried out a comparison of each method using both Fisher Discriminant Analysis (FDA) and KFDA, (see Fig. 1 for BS and Fig. 2 for BWH). Since KFDA generates non-linear boundaries among classes, the classification rates are better than those achieved with FDA. Furthermore, we compared each method with its extended version, which is built by concatenating a single descriptor and its corresponding $LBP^{uni}_{P,R}$ histogram into a single sequence to represent a mixture descriptor.

We computed the F_1-Score $= 2 * \frac{P*S}{P+S}$ for each algorithm and measured the accuracy of the tests. For BS dataset, our proposal, CGF + $LBP^{uni}_{P,R}$, achieved the highest F_1-Score with 0.8637. A straightforward comparison with the work of Bruijne and Sørensen [3] is not possible because they reported a classification rate using patches of 31×31 pixels as the best case. Here, we used larger patches, which implies the risk of including different lobes that might have different emphysema and might decrease the classification performance. Using the BWH dataset, our proposal also achieved the highest F_1-Score with 0.6899. Mendoza et al. [4] reported a F_1-Score of 0.6440 using the kernel density estimation approach.

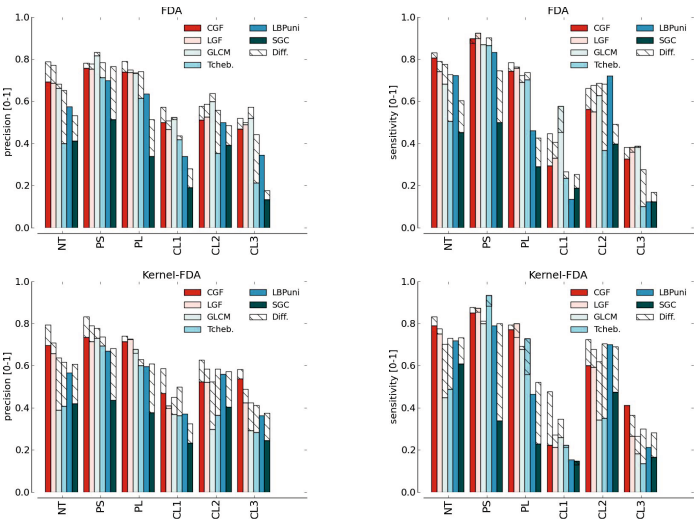

Fig. 2. BWH classification rates (Six classes). The first row shows the results using FDA while in the second row the classification rates using KFDA are shown. The extended methods, $Diff$, achieved higher rates than single texture approaches.

5 Conclusions

We proposed a novel approach to quantify emphysema patterns based on global and local descriptors to form a single sequence that represent any given texture patch. This approach simultaneously encodes global characteristics with local information that leads to better classification rates. Additionally, we analyzed six texture descriptors and compared their performance. Since the size of extended descriptors increases exponentially, we applied KFDA via the kernel trick to avoid computing a mapping function. This procedure resulted in an improvement of the classification rates.

Acknowledgments. This work has been sponsored by the grant UNAM PAPIIT IN113611. R. Nava gives a special thank to CONACYT for the doctoral scholarship 167161; V. Marcos thanks the programme "Juan de la Cierva" from the Spanish Ministry of Economy and Competitiveness; and R. San José was supported by NHLBI grants K25 HL104085-03 and 1R01HL116931-01.

References

1. Galban, C., Han, M., Boes, J., Chughtai, K., Meyer, C., Johnson, T., Galban, S., Rehemtulla, A., Kazerooni, E., Martínez, F., Ross, B.: Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression. Nat. Med. 18(11), 1711–1715 (2012)

2. Hayhurst, M., Flenley, D., Mclean, A., Wightman, A., Macnee, W., Wright, D., Lamb, D., Best, J.: Diagnosis of pulmonary emphysema by computerised tomography. The Lancet 324, 320–322 (1984)
3. Sørensen, L., Shaker, S., de Bruijne, M.: Quantitative analysis of pulmonary emphysema using Local Binary Patterns. IEEE Trans. Med. Imag. 29(2), 559–569 (2010)
4. Mendoza, C., Washko, G., Ross, J., Diaz, A., Lynch, D., Crapo, J., Silverman, E., Acha, B., Serrano, C., Estepar, R.: Emphysema quantification in a multi-scanner HRCT cohort using local intensity distributions. In: 9th IEEE International Symposium on Biomedical Imaging (ISBI), pp. 474–477 (2012)
5. Sørensen, L., Nielsen, M., Lo, P., Ashraf, H., Pedersen, J., de Bruijne, M.: Texture-based analysis of COPD: A data-driven approach. IEEE Trans. Med. Imag. 31(1), 70–78 (2012)
6. Depeursinge, A., Foncubierta–Rodriguez, A., Van de Ville, D., Müller, H.: Multiscale lung texture signature learning using the riesz transform. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 517–524. Springer, Heidelberg (2012)
7. Gabor, D.: Theory of communication. J. Inst. Elec. Eng (London) 93III, 429–457 (1946)
8. Daugman, J.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A 2, 1160–1169 (1985)
9. Nava, R., Escalante-Ramírez, B., Cristóbal, G.: Texture image retrieval based on log-gabor features. In: Alvarez, L., Mejail, M., Gomez, L., Jacobo, J. (eds.) CIARP 2012. LNCS, vol. 7441, pp. 414–421. Springer, Heidelberg (2012)
10. Field, D.: Relations between the statistics of natural images and the response properties of cortical cells. J. Opt. Soc. Am. A 4(12), 2379–2394 (1987)
11. Perrinet, L.U., Samuelides, M., Thorpe, S.J.: Sparse spike coding in an asynchronous feed-forward multi-layer neural network using matching pursuit. Neurocomputing 57C (2002)
12. Perrinet, L.U.: Role of homeostasis in learning sparse representations. Neural Computation 22(7), 1812–1836 (2010)
13. Haralick, R., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst., Man, Cybern., Syst. SMC-3(6), 610–621 (1973)
14. Randen, T., Husøy, J.H.: Filtering for texture classification: A comparative study. IEEE Trans. Pattern Anal. Mach. Intell. 21, 291–310 (1999)
15. Marcos, V., Cristóbal, G.: Texture classification using Tchebichef moments. J. Opt. Soc. Am. A 30(8), 1580–1591 (2013)
16. Ojala, T., Pietikainen, M., Harwood, D.: Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In: 12th International Conference on Pattern Recognition - Conference A: Computer Vision Image Processing (IAPR), vol. 1, pp. 582–585 (1994)
17. Ojala, T., Pietikäinen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with Local Binary Patterns. IEEE Trans. Pattern Anal. Mach. Intell. 24(7), 971–987 (2002)
18. Mika, S., Ratsch, G., Weston, J., Scholkopf, B., Mullers, K.: Fisher discriminant analysis with kernels. In: IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing IX, pp. 41–48 (1999)