

Automatic Scene Recognition for Low-Resource Devices using Evolving Classifiers

Javier Andreu, Rashmi Dutta Baruah, Plamen Angelov

School of Computing and Communications

Infolab21, Lancaster University

Lancaster, LA1 4WA, United Kingdom

{j.andreu,r.duttabaruah,p.angelov}@lancaster.ac.uk

Abstract– In this paper an original approach is proposed which makes possible autonomous scenes recognition performed on-line by an evolving self-learning classifier. Existing approaches for scene recognition are off-line and used in intelligent albums for picture categorization/selection. The emergence of powerful mobile platforms with camera on board and sensor-based autonomous (robotic) systems is pushing forward the requirement for efficient self-learning and adaptive/evolving algorithms. Fast real-time and online algorithms for categorisation of the real world environment based on live video stream are essential for understanding and situation awareness as well as for localization and context awareness. In scene analysis the critical problem is feature extraction mechanism for a quick description of the scene. In this paper we apply a well known technique called spatial envelop or GIST. Visual scenes can be quite different but very often they can be grouped in similar types/categories. For example, pictures from different cities across the Globe, e.g. Tokyo, Vancouver, New York, Moscow, Dusseldorf, etc. bear the similar pattern of an urban scene – high rise buildings, despite the differences in the architectural style. Same applies for the beaches of Miami, Maldives, Varna, Costa del Sol, etc. One assumption based on which such automatic video classifiers can be build is to pre-train them using a large number of such images from different groups. Variety of possible scenes suggests the limitations of such an approach. Therefore, we use in this paper the recently propose evolving fuzzy rule-based classifier, *simpl_eClass*, which is self-learning and thus updates its rules and categories descriptions with each new image. In addition, it is fully recursive, computationally efficient and yet linguistically transparent.

Key words - scene recognition, evolving fuzzy rule based classifier, GIST, image processing

I. INTRODUCTION

The rapid evolution of MEMS (Micro-Electro-Mechanical Systems) has brought to our lives hundreds of sensors which can fit in the palm of our hand and also be embedded in personal devices such as PCs, PDAs, tablets and mobile devices. One of the sources of information (a kind of a sensor) which is integrated within such devices is the video camera. Thanks to these built-in cameras it is possible to obtain video input from nearly all portable devices nowadays.

There is a body of research which is focused on the event detection such as landmarks [1]-[2] and on scene recognition such as [3]-[4]. In the field of scene recognition, there are works that developed intelligent photo albums such as [5] or

photo retrieval and classification systems as [6]. However, only in [1] the algorithm that has been used is recursive, on-line, adaptive and can be embedded on a chip. Most of the other published approaches are off-line, computationally expensive and not suitable for embedded applications.

A very important by-product of having computationally lean algorithms is the energy efficiency. In terms of energy consumption one can differentiate between two types of digital capturing images; i) one is CCD (charge coupled devices), and the other one is ii) CMOS (complementary metal oxide semiconductor). Both of them convert images into electric charges. In terms of functioning, the former determines a charge for every pixel which is transferred into another chip as an analogue signal; the latter defines a voltage charge for each pixel conversion and performs digitalization by itself. Although, both offer an excellent imaging performance, CMOS is widely used because it is more energy efficient compare to CCD [7]. It also offers more functionality such as discretization and signal digitalization integrated on the chip than CCD. However, the picture quality of CMOS is not as good as the ones that can be obtained from a CCD camera.

Consequently, in order to maximize the performance and energy consumption of these devices, quick and cheap algorithms in terms of resources will have an important role.

For this purpose, we have developed and studied new algorithms based on recursive and computationally efficient evolving classifiers [8]-[9]. We used GIST descriptor [10]-[11] to compute the features from images. The idea of this descriptor is quite simple; it aims to develop a low dimensional representation of the scene but without partitioning. Basically, GIST feature instances are defined by perceptual dimension such as roughness, openness, naturalness, expansion. All these characteristics represent the spatial structure of a scene. They are easily computed using spectral and superficial localized information. In addition, the use of GIST in scene recognition has been widely experimented [12]-[13].

The evolving fuzzy rule-based classifiers that we will apply in this study [8]-[9] represent a new paradigm of online self-learning which combines the classifier design and classification (prediction) and can also be used as unsupervised tool for clustering (the labeling of clusters can be left for the user). Similar techniques of evolving classifiers have been recently reported in human activity recognition [14] and intelligent robotics applications [15].

The remainder of the paper is organized as follows; in section II related works are presented; in section III feature

extraction method is given; in section IV image processing methods are described, in section V the evolving classifier is portrayed, a summary of the entire approach is given in section VI; section VII analyses experimental results and finally section VIII concludes the paper.

II. RELATED WORK

Nowadays, the topic of natural scene classification has been widely studied. Fundamentally, most of the publications focus on modeling the scene using statistical information which is fed later to classification. They differ in the selected method to compute the features. For example [16] which is one of the first works in the field, proposed a feature framework which differentiate between indoor and outdoor scenes. It uses low level features such as; texture, color, and frequency information. Some standards, such as MPEG-7 offer a detailed description of these features. Other, more complex studies, try to relate graphical model with object recognition [17]. Further approaches also try to split the scene semantics concepts [18]. For this purpose, grids of sub-regions of the image are created and later semantics concepts (depicted in semantic nets) are related to these regions. One outstanding contribution to this field was the work [10] which used GIST [11] to build efficient classifiers. However, their work is in off-line mode.

Many other studies used GIST-based scene descriptors. For example [19] used GIST combined with clustering methods to recognize landmarks in the scene. In [20] the recognition of facial expression is enhanced by using context information of the scene using GIST descriptors. In [21] a context-based recognition approach was presented for mobile robots capturing the GIST-based description of the scene.

Evolving classifiers has also been used for recognizing landmarks [22] and for autonomous self-location of mobile robots [9]. In this paper we propose an innovative approach to automatic scene analysis using on-line and evolving self-learning fuzzy rule-based (FRB) classifier that is *potential free*. This results in an algorithm which is computationally and energy efficient and very precise.

III. GIST DESCRIPTION OF A SCENE

Usually, it is said that humans are able to recognize a scene from a single glance. This is because our brain captures a spatial representation of the scene, recognizing some meaningful objects or salient information of the scene. However, our rational capacities cannot process all information of a complex scene in just a single glance. What our brain does is to just construct a representative picture of perceptual and conceptual low-level (color, contours); intermediate (shapes, texture regions) and high-level information (semantic knowledge). Based on behavioral studies, Friedman [12] observed that regardless of the visual complexity of the scene, humans are able to recognize either categories or special layout of the scene.

More recently, based on Friedman's fundamental studies, a team at MIT studied the application of the GIST to images [10],[22]-[23] to analyze characteristics and properties of images, in order to create a structured summary of a real-world scene. It is stated that descriptive factors exist that can

be instrumental for the diagnostic structure of the scene. Torralba et al. came up with the example that urban zones are structured vertically while the forests are textured [11].

Unlike the other approaches, which partition the scene in bins, the standard GIST-based approach considers the scene as a whole, but analyses it spectrally. Consequently, there is no need to detect individual landmarks and objects in the scene. Instead, a scene can be represented by a vector of spectral values. These values are spatial envelope properties such as expansion, openness, roughness.

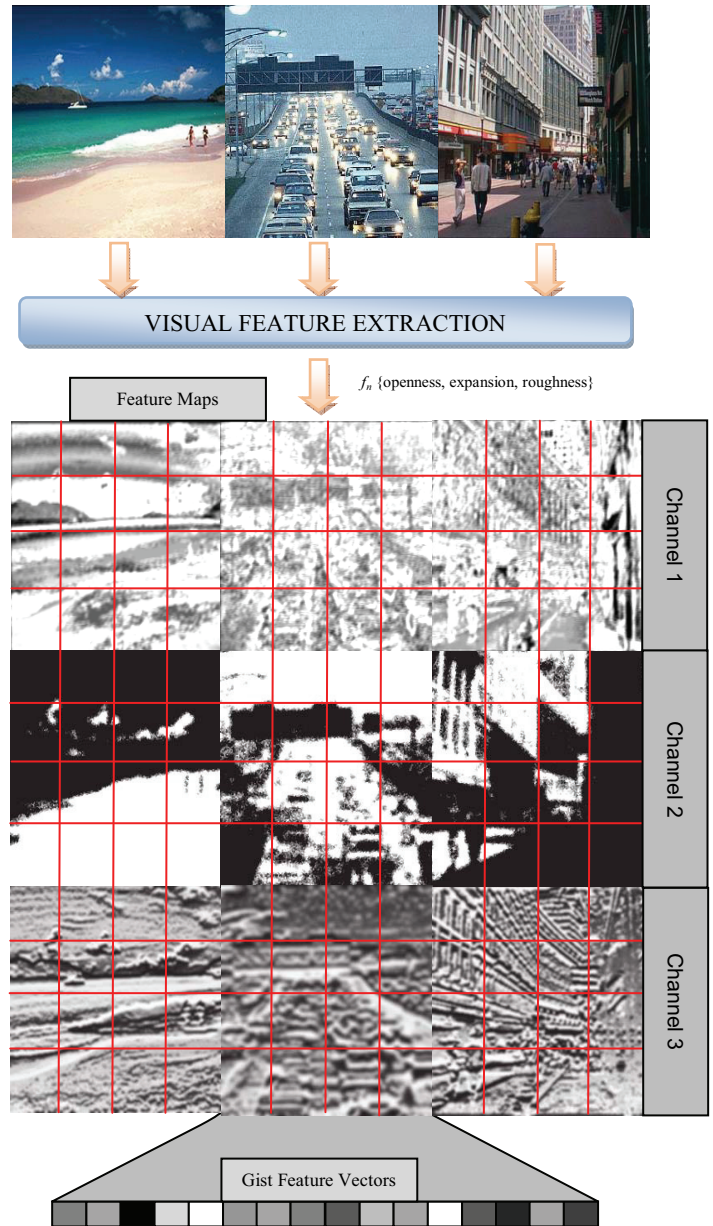


Fig 1. Feature vector of spatial envelop properties

Each property is equivalent to a dimension of the spatial envelope space. These features are meaningful enough to describe the spatial structures of the scene. To estimate these properties spectral and coarsely localized information is used.

To have a rough idea of how GIST-based analysis works let's consider the fundament of the approach.

Most widely used image-based representation for the analysis of the spectral dynamic properties is provided by the discrete Fourier Transform (FT). FT filters represent the spatial domain of the input figure in a frequency domain of sine and cosine components. The discrete form of the FT is a sampled version of it. In other words, it is a set of most representative samples of frequencies. FT-based image representations have shown to have a very good performance in scene recognition problems.

In the fundament of GIST approach are Gabor filters which bear some similarity to Fourier filters. However, the former ones are defined by harmonic functions modulated by Gaussian distributions. As a result of adjustment into Gaussians, these types of filters are limited by frequency bands. With a judicious selection of frequencies (e.g. by successive factors of 2, by octaves), a succession of Gabor filters can be assimilated to a wavelet transform. It results in a very good method of image compression and discretization.

In Figure 1 several Gabor filters with selected channels are computed on a grid of the image (4x4) and indexed into an array. This array is called GIST of the scene. This array represents the spatial layout of the scenes, i.e two images with similar array values (layout) are often similar scenes and vice versa.

IV. ONLINE PREPROCESSING OF FEATURES.

For online recognition of scenes, a fast computation of features is necessary. This can be achieved by means of a fast/recursive computation of parameters. On this sense, in the literature we may find several research works which already have considered this issue. For example, [25] proposes an efficient and recursive computation of Gabor filters to cope with the multi-resolution/orientation of images (which can be a computational intense process). Other works like [26] defined a fast computation of Gabor functions using a bi-orthogonal condition to compute the dual frames. A very precise solution to increase speed on computation of Gabor features is to compute the Gaussian envelope recursively. This way, [27] proposes a recursive way to compute the Gaussian kernel of the Gabor filter general formula [28].

In most of the cases, number of Gist features from the resultant Gabor filters applied, is usually big. In our case, the resultant 4x4 grid partition leads to a feature map of 544 gist features (from 34 Gabor features per region). Consequently, a dimensionality reduction algorithm should be also considered. Although, the recursive computation of these methods is not as extended as the ones of Gabor filters, there has been some progress in this line.

Most widely, used approaches to attain this dimensional reduction are by means of methods such as Principal Component Analysis (PCA) or Fast Independent Component Analysis (FastICA). The latter is already formulated as a computationally efficient method; the former by definition is not. However, there are some works that have recently

achieved to formulate the computation of PCA values online [29].

Based on the trend of some works in biology and biotechnology, PCA/ICA projections are reinforced with a Linear Discriminant Analysis (LDA) to increase the description power of features regarding to the class.

1) Recursive Principal Component Analysis (PCA)

It is one of the most commonly used approaches for feature extraction/reduction. Principal Component Analysis (PCA) addresses the issue of correlation of raw input data by orthogonalisation. PCA mainly aims to transform the raw data that are, usually, correlated into a smaller set of uncorrelated variables. Summarizing, we may say that the transformation of new variables are performed by linear combinations of the original (measured) variables. In this connection, the overall number of variables that are required to be measured is not being reduced; what is reduced is the number of inputs to the classifier (manipulated variables by the classifier) by preserving the majority of the information in terms of data variance which is useful for correct classification.

It is mainly based on producing new variables maximising the variance of the data. PCA is very helpful in revealing the hidden dynamics that is often 'hidden' in a raw signal. This method is completely unsupervised (that makes easy to compute it online) and does not take into account any category or label. Its main target is to find efficient orthogonal projections explaining the variance of the data most. In the standard (offline) version of PCA, projections are evaluated by the value of their associated eigen-values. The number of desired orthogonal directions must be defined previously. This is because each projection has to be orthogonal to its previous, independently of the first computed one that will be update according to the coming data.

As it has been discussed before, in this work we aim to perform in an online/real time basis. In order to get a recursive/fast implementation of PCA we can follow the steps suggested by [29].

PCA projections are represented by the eigen-values, v . They can be calculated as follows:

$$v_k = \frac{1}{k} \sum_{i=1}^k u_i \cdot u_i^T \cdot \frac{v_{i-1}}{\|v_{i-1}\|} \quad (1)$$

Note u denotes an orthogonal projection and k is the current time instant/sample.

This expression is based on the definition of v as $\gamma \cdot x = C \cdot x$ and $v = \gamma \cdot x$ where γ is the magnitude of the eigen-value, $\|v\|$, and C the covariance matrix.

Expression (1) can be expressed recursively by weighting its current and last estimates by $\frac{k-1}{k}$ and $\frac{1}{k}$ respectively;

then we get:

$$v_k = \frac{k-1}{k} \cdot v_{k-1} + \frac{1}{k} \cdot u_k \cdot u_k^T \cdot \frac{v_{k-1}}{v_{k-1}} \quad (2)$$

As a result, it is possible to estimate recursively orthonormalized values to guarantee maximum independence of directions. Thus, parting from the first order eigenvector a new one is estimated by subtracting its orthogonal projection. The formula which describes this projection u_j is as follows:

$$u_k = u_{k-1} - u_{k-1}^T \cdot \frac{v_{k-1}}{\|v_{k-1}\|} \cdot \frac{v_k}{\|v_k\|} \quad (3)$$

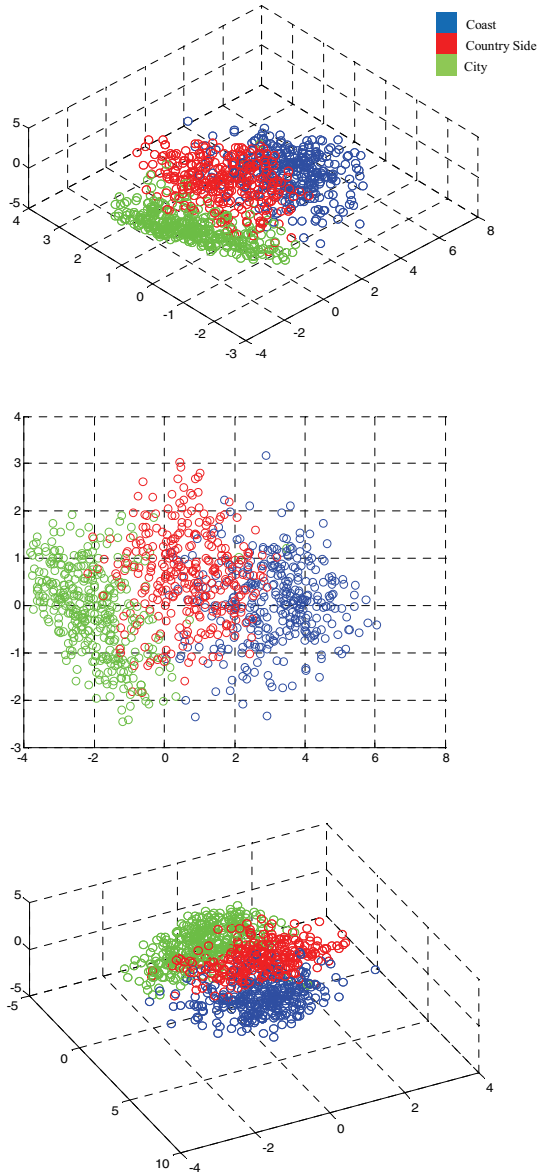


Fig 2. rPCA-LDA applied to data with three categories.

2) Recursive Linear Discriminant Analysis (LDA)

Methods explained above are very useful for reducing correlation and number of inputs to the classifier. However, they do not take into account what category of scene data belongs to. They just get the best representation of data based on its initial values (inputs) without considering the output.

We want to classify and recognize categories. Hence, bearing in mind that scenes are categorized, LDA can be applied, it is considered to use LDA to improve the description power of features, this time regarding to the class label (outputs in terms of the classifier).

While PCA/ICA finds the best projections of data in a way that maximises their separability, LDA helps finding these directions which will be best to distinguish between different classes (scenes categories). A combination of both, PCA and LDA has proven to be an effective pre-processing technique in spectroscopy. LDA method is based on the sample variance (σ^2) and sample variance per class (σ_c^2). They are used to form scatter matrices (S_b and S_w). S_w denotes the within-class scatter matrix. Likewise, S_b represents a singular matrix denoted as between-classes scatter matrix. The separation between those matrices is defined as a measure of the signal-to-noise ratio for class labelling. Eigenvectors of S_w^{-1} and S_b are consider the best coordinates to discriminate classes better. Thus, new projections of data enhance a separation between categories. In [29] a recursive method is presented to compute both LDA and PCA (fig. 2).

V. EVOLVING FUZZY RULE-BASED CLASSIFIERS

In this paper we use evolving fuzzy-rule based (FRB) classifiers of the eClass family [8], namely *eClass1* [8] and a simplified version called *simpl_eClass1* [30]. The evolving classifiers are characterized by self-learning of both structure and parameters and use evolving Takagi-Sugeno FRB models [31]. The structure of both the classifiers is based on first-order multi-input multi-output (MIMO) type FRB model of the following form.

$$\text{Rule}^i : IF(x_1 \text{ is } \sim x_1^*) AND \dots (x_n \text{ is } \sim x_n^*) THEN(y^i = x_e \theta^i) \quad (4)$$

where $y^i = [y_1^i \ y_2^i \ \dots \ y_m^i]$ is the m-dimensional output of i^{th} rule, $x_e = [1 \ x_1 \ \dots \ x_n]$ is the extended input vector, θ^i is the matrix of local sub-system parameters given as:

$$\theta^i = \begin{bmatrix} a_{01}^i & a_{02}^i & \dots & a_{0m}^i \\ a_{11}^i & a_{12}^i & \dots & a_{1m}^i \\ \vdots & \vdots & \dots & \vdots \\ a_{n1}^i & a_{n2}^i & \dots & a_{nm}^i \end{bmatrix}$$

The rules of the classifiers are formed around focal points (cluster centres), where a focal point is a representative of all the data samples within a cluster. The consequent parameters of the rules are estimated using weighted recursive least squares (wRLS) approach [31]. *Simpl_eClass1* differs from *eClass1* mainly due to the learning mechanism which is potential/density free [30]. In order to determine whether an incoming sample is a new focal point, *eClass1* requires calculation of a *potential/density* value [8],[31] for the incoming sample and update of potentials/densities of all the existing focal points. Thus, the computational complexity of *eClass1* (similarly to *eTS*) is $O(N)$, where N is the number of

existing focal points (respectively fuzzy rules). On the other hand, the determination of focal points in *simpl_eClass1* requires only mean value calculation, and it can be calculated recursively as each sample arrives (same as in *simpl_eTS+* [32]). Thus, in *simpl_eClass1* the computational complexity is reduced from $O(N)$ to $O(1)$ – by a factor of N . A detailed description of *simpl_eClass1* is presented in [30].

In both the classifiers, before the rule learning phase, the actual label of a given sample is converted into numeric form in order to enable the consequent part of a rule to be represented in the form of (4). For example, for a two class problem, *class 1* can be represented as $y = [1 \ 0]$, and *class 2* as $y = [0 \ 1]$. Similarly, for a three class problem, *class 1* is $y = [1 \ 0 \ 0]$, *class 2* is $y = [0 \ 1 \ 0]$, and *class 3* is $y = [0 \ 0 \ 1]$. As can be observed from the representation, the dimension (m) of y is same as the number of classes. For the classification task, the classifiers estimate the overall output ($\mathbf{est_Y}$) by weighted sum of outputs of all the rules using (5)-(6). The index corresponding to the highest value element of $\mathbf{est_Y}$ is assigned as the class label for the sample as given in (7) – so called winner takes all principle which is usually used in classification. For example, suppose for a given sample the value of $\mathbf{est_Y}$ is determined to be $\mathbf{est_Y} = [0.45 \ 0.23 \ 0.78]$ in a three class problem, then the sample is assigned class label 3.

$$\mathbf{est_Y} = \sum_{i=1}^N \lambda^i \mathbf{y}^i; \quad \lambda^i = \frac{\tau^i}{\sum_{j=1}^N \tau^j} \quad (5)$$

$$\tau^i = \prod_{j=1}^n \mu_j^i(x_j), \quad i = [1, N], \quad \mu_j^i = e^{-\frac{\|x - x^i\|^2}{2(r_j^i)^2}} \quad (6)$$

where $\mathbf{estY}$ is the overall estimated output, τ^i is the firing strength of the i^{th} rule, $(r_j^i), i = [1, N], j = [1, n]$ is the spread of the membership function.

$$\text{Class} = \text{Class}^{i^*}, \quad i^* = \arg \max_{i=1}^m (\mathbf{est_Y}) \quad (7)$$

VI. PROPOSED EVOLVING ONLINE METHOD.

The procedures explained above might be concatenated in a single process instance by instance. It does mean that just from a standalone data source for example a frame from a camera. The proposed method must be able to recognize pre-defined categories and also any other that has not appeared before, as a result of the evolving nature of the algorithm.

In figure 3, the entire process is represented. A single data sample or frame for a video camera is obtained and formatted to a representative numeric model. The transformation model used in this work is based on a RGB color map, however other geometrical representations such as HSL or HSV could be considered. This image representation is handed to another module where the resultant color map is partitioned and over those partitions GIST is applied. Applying 32 different feature maps and partitioning the data on 4×4 regions per map, a total of 512 features are obtained per image. For a learning model this number of features implies a very high number of dimensions in most cases. A data compression technique is

highly recommended, hence GIST features are handed to another module where PCA (for compression) in addition to LDA (enhance feature categorization) is used. A reduced data set of 22 features is obtained after this process, and then an evolving fuzzy rule base model is build up from the scratch. In first iteration, first sample will be considered as a rule and next samples update the FRB model adding, removing or updating new rules. New rules may be associated to one of existing categories and improve interpretation of that category. On other hand, by means of a small supervision or a simple rules monitoring, a new rule can be considered as a start point of a new category.

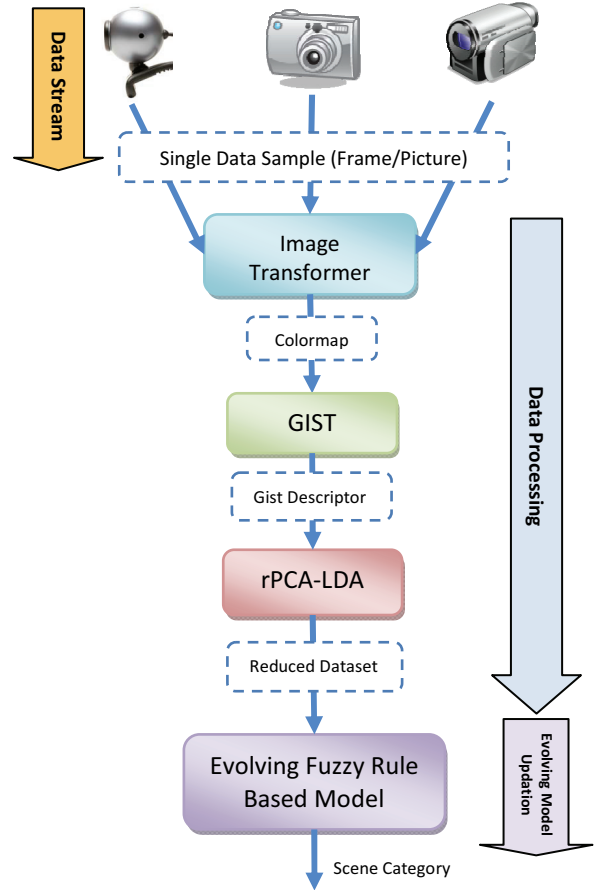


Fig. 3. Processing order of different modules (approaches) to work in a fully online (sample by sample) basis.

VII. EXPERIMENTAL RESULTS AND EVALUATION

In this section the proposed classification scheme for scene recognition (*eClass_Scenes*) which includes recursive PCA-LDA as well as two versions of evolving classifier: a FRB (*eClass*) and a potential free version (*simpl_eClass*). These approaches are compared with several well known classifiers. These classifiers are organized in two groups: i) static, off-line classifiers, and ii) on-line and evolving classifiers. The first ones, generally, provide higher classification rate, but impose heavy requirements on the memory and computational resources. Some of them can process the data incrementally but they keep most of the data in the memory and process all

of them at once. On the other hand, the latter, process the data sample by sample (on-line) always and may be able to adapt their structure (evolve) as a reaction to the variance in the data pattern. A less complex model is obtained from their adaptation and evolution. It is principally due to the on-line processing mode they operate. Significantly less computational resources are required. For this reason, they use to have a better performance in terms of memory consumption and CPU usage which compensate a slight lower accuracy.

The selected data set [33] to test the previous mentioned methods contains 8 outdoor scene categories: street, open country, inside city, forest, coast, mountain, tall buildings and highways. In total, they are 2600 colour images of 256x256 pixels. A confusion matrix of experiments is presented in Figure 4. In this experiment features were reduced to 22 out of 512 initial inputs using rPCA-LDA. Classifications rates vary between 79% and 81%. Although, in some cases categories were independently recognized with more than 90% of accuracy, for instance Inside City, Open Country, Mountain and Forest. By contrast, one category was recognized with less than a 90%. This is the case of highways. This category could wrongly be identified as coast. It is understandable that mirages of sunny highways could resemble sea water channels.

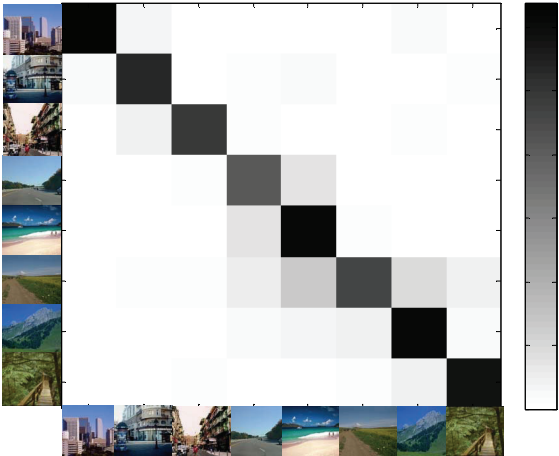


Figure 4. Confusion matrix for classification of scenes.

TABLE II. EVALUATION TIMES FOR THE WHOLE APPROACH

Method	Mean Time (seconds)
RGB Format + GIST	310.73 ms
rPCA-LDA	0.104 ms
Evolving Fuzzy Classifier	0.6/0.2 ms

TABLE I. EVALUATION PARAMETERS USING RPCA-LDA

Method	Evaluation Parameters		
	Accuracy Rate	Time, ms	Number of Rules-SV-K
Evolving			
eClass1	79.1578%	0.648	31

Method	Evaluation Parameters		
	Accuracy Rate	Time, ms	Number of Rules-SV-K
Simple_eClass1	80.580357%	0.240	29
Non-Evolving			
NaiveBayes	79.6646%	10	-
Nearest Neighbour	74.3187 %	0.31	8
SVM Classifier	80.2935%	2	140

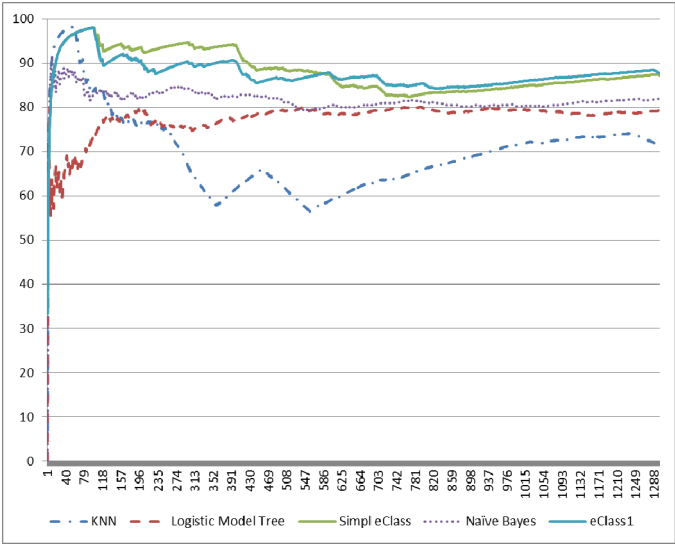


Fig 5. Evolution of rate accuracy sample by sample.

VIII. CONCLUSION

The foundation of a new approach for a fast automatic categorization of visual scenes is presented in the paper. It combines GIST feature selection and PCA-LDA methods with a self-learning evolving fuzzy rule-based classifier with very light computational load (*simpl_eClass*). *Simpl_eClass* can work in real-time, on-line and is linguistically transparent. The proposed approach is compared thoroughly with possible alternative methods including off-line and incremental approaches and the results demonstrate that it outperforms them computationally, in terms of complexity and transparency (Table I). The approach has a latency of approximately 0.3 seconds in total (including spatial descriptors, pre-processing, model updating and estimation) which makes it perfectly usable in real-time application. Moreover, this is calculated based on a Matlab application, an application using C or on-chip VHDL will further reduce this latency. It is important to stress that the computational time of model update is significantly smaller than that of the other approaches. Figure 5 clearly demonstrates that eClass1 and simpl_eClass outperform other methods in terms of accuracy. To conclude we may say that scene classification can be important for scene analysis and a better understanding of the differences in scenes etc. The results are promising as Evolving Fuzzy Classifiers can compete with renowned SVM approach despite incrementally trained where SVM sees all

the samples at once in a batch setting; furthermore, the number of structural components (rules) is significantly lower than in case of SVM (support vectors).

IX. REFERENCES

- [1] Zhou X. and Angelov P., "Real-time joint Landmark Recognition and Classifier Generation by an Evolving Fuzzy System" in *Proceedings of IEEE International Conference on Fuzzy Systems*, 2006 pp.1205-1212.
- [2] Bandera A. and Marfil R., "Incremental Hybrid Approach for Unsupervised Classification: Applications to Visual Landmarks Recognition" in *Lecture Notes in Computer Science*, 2010, vol. 6111, pp.137-146.
- [3] Li J. and Qian Y., "Automatic scene recognition for digital camera by semantic features" in *Proceedings of International Conference on Wavelet Analysis and Pattern Recognition*, 2008, vol.1, pp.327-332.
- [4] Matsushita N., Hihara D., Ushiro T., Yoshimura S., Rekimoto J., Yamamoto Y., "ID CAM: a smart camera for scene capturing and ID recognition" in *Proceedings of The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2003, pp. 227-236.
- [5] Chai Y., Xia T., Zhu J. and Li, "Intelligent Digital Photo Management System Using Ontology and SWRL" in *Proceedings of International Conference on Computational Intelligence and Security*, 2010, pp.18-22.
- [6] La Cascia M., Morana M. and Sorce S., "Mobile Interface for Content-Based Image Management," in *Proceedings of Complex, Intelligent and Software Intensive Systems*, 2010, pp. 718-723.
- [7] Rizkalla M.E. and Gundrum H.C., "Design of analog CMOS integrated circuits input stage for the operation at zero temperature coefficient using PSPICE" in *Proceedings of the 35th Midwest Symposium on Circuits and Systems*, 1992, vol.1, pp.16-19.
- [8] Angelov P. and Zhou X., "Evolving Fuzzy-Rule-based Classifiers from Data Streams in *Proceedings of IEEE Trans. on Fuzzy Systems*, 2008, vol. 16(6) *Evolving Fuzzy Systems*, pp.1462-1475.
- [9] Zhou X. and Angelov P., "An Approach to Autonomous Self-localization of a Mobile Robot in Completely Unknown Environment using Evolving Fuzzy Rule-based Classifier" in *Proceedings of IEEE Intern. Conf. on Comp. Intel. App. for Defense & Security*, 2007, pp.131-138.
- [10] Oliva A. and Torralba A., "Modeling the shape of the scene: a holistic representation of the spatial envelope" in *International Journal in Computer Vision*, 2001, vol. 42, pp. 145-175.
- [11] Oliva, A. and Torralba, A., Building the Gist of a Scene: The Role of Global Image Features in Recognition in *Progress in Brain Research: Visual perception*, 2006, vol. 155, pp. 23-36.
- [12] A. Friedman, "Framing pictures: The role of knowledge in automatized encoding and memory for gist", in *J. Exp. Psychol. Gen.*, 1979 vol. 108 pp. 316-355.
- [13] Murillo A.C., Kosecka J., "Experiments in place recognition using gist panoramas" in *Proceedings of IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*, 2009, pp. 2196-2203.
- [14] Andreu J. and Angelov P., "Real-time human activity recognition from wireless sensors using evolving fuzzy systems" in *Proceedings of IEEE International Conference on Fuzzy Systems*, 2010, pp. 1-8.
- [15] Sadeghi-Tehran P., Andreu J., Angelov P. and Zhou X., "Intelligent Leader-Follower Behaviour for Unmanned Ground-based Vehicles" *Journal of Automation, Mobile Robotics and Intelligent Systems (JAMRIS)*, vol. 5 (1), pp.1-11, 2011
- [16] Szummer M. and Picard R. "Indoor-outdoor image classification" In *Proceedings of International Workshop on Content-based Access of Image and Video Databases*, 1998, pp. 42-55.
- [17] Murphy K., Torralba A. and Freeman, W.: Using the forest to see the trees: exploiting context for visual object detection and localization. In *Communications of the ACM*. 2004 vol. 16 (3).
- [18] Ankenbrandt C.A, Buckles B.P and Petty F.E, "Scene recognition using genetic algorithms with semantic nets" in *Pattern Recognition Letters*, 1990, Vol.11 (4), pp. 285-293.
- [19] Li X., Wu C., Zach C., Lazebnik S. and Frahm J. "Modeling and Recognition of Landmark Image Collections Using Iconic Scene Graphs" in *Proceedings of the 10th European Conference on Computer Vision*, 2008, pp. 427-440.
- [20] Ruthger R., De Gelder B., "Recognition of facial expression is influenced by emotional scene gist" in *Cognitive, Affective and Behavioural Neuroscience*, 2008, vol 8 (3), pp. 264-272.
- [21] Siagian C., Itti L., "Gist: A Mobile Robotics Application of Context-Based Vision in Outdoor Environment" in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, pp. 285-293.
- [22] Zhou X. and Angelov P., "Real-Time joint Landmark Recognition and Classifier Generation by an Evolving Fuzzy System" in *IEEE World Congress on Computational Intelligence*, 2006, pp.6314-6321.
- [23] Lazebnik S., Schmid C. and Ponce J., "Beyond Bags of Features: Spatial Pyramid Matching for Recognizing Natural Scene Categories" in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2006, pp. 2169-2178.
- [24] Oliva A., and Schyns P.G., "Diagnostic colors mediate scene recognition" in *Cognitive Psychology*, 2006, vol. 41, pp. 176-210.
- [25] Kameswara R.N., Mehrotra R., Ranganathan N., "Efficient computation of gabor filter based multiresolution responses" in *Pattern Recognition* vol 27(7), pp. 925-938
- [26] Youhong L., Morris J.M., "Fast Computation of Gabor Functions" in *IEEE Signal Processing Letters*, 1996, vol 3(3), pp. 75-78.
- [27] Young I.T, Van Vliet L.J. and Van Ginkel M., "Recursive Gabor Filtering" in *IEEE Transactions on Signal Processing*, vol. 50(11), pp. 2798-2805.
- [28] Gabor D., "Theory of communications," in *Proc. Inst. Elect. Eng.*, 1946, vol. 93, pp. 429-459.
- [29] Dagher I.; "Incremental PCA-LDA algorithm" in *Computational Intelligence for Measurement Systems and Applications (CIMSA)*, 2010 pp. 97-101
- [30] Dutta-Baruah R., Angelov P. and Andreu J., "Simpl_eClass: Simplified Potencial-free Evolving Fuzzy Rule-Based Classifiers" in *IEEE International Conference on Systems, Man, and Cybernetics*, Oct. 9-12, 2011. Submitted.
- [31] Angelov P. and Filev D., "An approach to online identification of Takagi-Sugeno fuzzy models" in *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, 2004, vol. 34, pp. 484-498.
- [32] Angelov P., "Fuzzily connected multi-model systems evolving autonomously from data streams" *IEEE Transactions on Systems Man and Cybernetics*. In Press.
- [33] Spatial Envelope [online]. Available: <http://people.csail.mit.edu/>