Evolutionary Learning of Feature Fusion Hierarchies

Fabien Scalzo FSCALZO@MEDNET.UCLA.EDU
Division of Neurosurgery, Geffen School of Medicine, University of California, Los Angeles, CA, USA

George Bebis BEBIS@CSE.UNR.EDU
Mircea Nicolescu MIRCEA@CSE.UNR.EDU
Leandro Loss LOSS@CSE.UNR.EDU
Computer Vision Lab. Department of Computer Science, University of Nevada, Reno, NV, USA

Abstract

We present a hierarchical feature fusion model for image classification that is constructed by an evolutionary learning algorithm. The model has the ability to combine local patches whose location, width and height are automatically determined during learning. The representational framework takes the form of a two-level hierarchy which combines feature fusion and decision fusion into a unified model. The structure of the hierarchy itself is constructed automatically during learning to produce optimal local feature combinations. A comparative evaluation of different classifiers is provided on a challenging gender classification image database. It demonstrates the effectiveness of these Feature Fusion Hierarchies (FFH).

1. Introduction

The generalization of new image acquisition devices and the development of new feature extractors have recently increased the interest of combining complementary modalities or features to perform automatic image classification. Hierarchical approaches (Singh et al., 2008; Podolak, 2008; Kim & Oh, 2008) to image classification are particularly interesting to solve complex problems because they are capable to decompose them into tasks that are often easier to tackle. However, these approaches often tend to manually define the structure of their hierarchy depending on the features involved (Tan & Triggs, 2007), and can only exploit a limited number of features. The current paper addresses these problems by presenting a framework that performs gender classification based on a large set of features extracted from facial images. The structure of the model as well as its parameters are estimated by a genetic learning algorithm that explores the space of possible hierarchies.

2. Feature Fusion Hierarchies

Feature Fusion Hierarchies (FFH) address the problem of fusing high-dimensional registered feature sets for image classification. The representational framework takes the form of a two-level hierarchy which combines local feature fusion and decision fusion into a unified model (Figure 1).

Given a feature set $I(x, y, f)$, where (x, y) denotes a position in the image, and f is a feature, the feature fusion level is defined as a set of compound features $\mathcal{C}$. Each compound feature $\mathcal{C}_i$ combines a subset of features $f_{\mathcal{C}_i} \in f$ over a local window $\theta_{\mathcal{C}_i}$. This fusion is done using a dimensionality reduction technique, denoted $\mathcal{R}_i(I_{f_{\mathcal{C}_i},\theta_{\mathcal{C}_i}})$, and learned in a supervised way (*e.g.* LDA). A key property of this function $\mathcal{R}_i$ is to operate locally in the sense that it exploits local adaptive windows (Scalzo & Piater, 2007) whose parameters $\theta_{\mathcal{C}_i} = \{x, y, Sx, Sy\}$ are automatically adjusted during learning (position in the image (x, y), width Sx and height Sy). The output of the function $S_i = \mathcal{R}_i(I_{f_{\mathcal{C}_i},\theta_{\mathcal{C}_i}})$ corresponds to a lower dimensionality response vector. An additional classifier is learned on the top of the first level to form the second level $\mathcal{D}$ corresponding to the decision fusion. Its input data correspond to the compound feature output $\{S_1, S_2, \ldots, S_n\}$ merged into a single vector S.

3. Learning of Fusion Hierarchies

A canonical genetic algorithm is used to explore the space of possible hierarchies (both the *structure* and the *parameters* are estimated). The optimal solution is the one that offers the best classification rate on the validation data and minimizes the number of features used as well as the size of the patches.

3.1. Genome Representation

Each evolving genome in the population is represented as a binary vector encoding the structure and the pa-

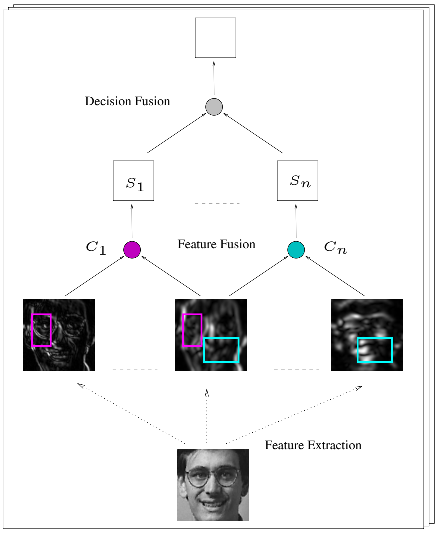

Figure 1. Overview of a Feature Fusion Hierarchy (FFH).

rameters of a specific hierarchy (Fig. 2). A genome defines the hierarchy as a set of N_c combinations $\mathcal{C}_i$. The *structure* of $\mathcal{C}_i$ corresponds to the subset of features that are combined $f_{\mathcal{C}_i} = \{f_1, \ldots, f_n\}$ whereas its *parameters* define the location (x, y) and size (Sx, Sy) of the local window in the image on which the fusion is performed.

Given n features at the first level, the structural part is represented as a n-length binary vector encoding the presence of the features in the combination. For the parameter part, variables $\{x, y, Sx, Sy\}$ are each represented as b bits vector.

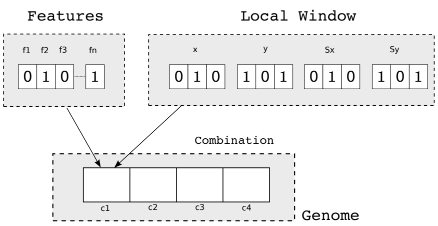

Figure 2. A Feature Fusion Hierarchy made of four compound features $(c1, c2, c3, c4)$ is encoded into a genome. The structural part and the parameters are embedded into a single binary vector.

3.2. Fitness, Crossover and Mutation

The fitness function $\text{fit}(h)$ is used to evaluate each individual in the population. It is set proportional to the classification rate r of the genome encoded hierarchy g, $\text{fit}(g) = r(g) + \alpha_1 n + \alpha_2 s^{-1}$, where n is the number of zeros in the structure part of the genome g and s is the total area covered by the patches. Parameters α_1 and α_2 are used respectively to support combinations that have a fewer number of features and are defined over a smaller window. A bi-parental random crossover and a single point mutation operator are used in our algorithm to produce new individuals.

4. Experiments

The effectiveness of the proposed framework is evaluated on a gender classification problem. Given a set of 400 facial images (Sun et al., 2002) captured under various conditions, the task is to correctly identify the gender of the subject present in the image. Each image is convolved with 35 Gabor filters and 5 Laplacian filters to produce the initial feature set on which our Feature Fusion Hierarchies (FFH) are constructed. The classification results after a three-fold cross-validation are reported in Table 1 for LDA, SVM and KSR classifiers. It can be observed that the use of the Feature Fusion Hierarchies (FFH) reduces significantly the classification error of a PCA-based framework and outperforms the results obtained by PCA-GA approach (Sun et al., 2002). This can be explained by the fact that our FFH approach exploits local features whereas PCA-GA computes the projections on the entire image.

	LDA	SVM	KSR
PCA	14.2%	8.9%	-%
GA-PCA(Sun et al., 2002)	9%	4.7%	-%
FFH (this paper)	**7.2%**	-%	**3.8%**

Table 1. Results for three different classifiers are reported for PCA, GA-PCA (Sun et al., 2002) and the Feature Fusion Hierarchies (FFH).

References

Kim, Y., & Oh, I. (2008). Classifier ensemble selection using hybrid genetic algorithms. *Pattern Recogn. Lett.*, *29*, 796–802.

Podolak, I. T. (2008). Hierarchical classifier with overlapping class groups. *Expert Syst. Appl.*, *34*, 673–682.

Scalzo, F., & Piater, J. (2007). Adaptive patch features for object class recognition with learned hierarchical models. *Beyond Patches*.

Singh, R., Vatsa, M., & Noore, A. (2008). Hierarchical fusion of multi-spectral face images for improved recognition performance. *Information Fusion*, *9*, 200–210.

Sun, Z., Bebis, G., Yuan, X., & Louis, S. J. (2002). Genetic Feature Subset Selection for Gender Classication: A Comparison Study. *IEEE International Conference on Image Processing*.

Tan, X., & Triggs, B. (2007). Fusing gabor and lbp feature sets for kernel-based face recognition. *Analysis and Modeling of Faces and Gestures* (pp. 235–249).