On Applicability of Principal Component Analysis to Concept Learning from Images

Boris Strandjev
Musala Soft Ltd
Sofia, Bulgaria

Gennady Agre
Institute of Information and Communication Technologies – BAS, Sofia, Bulgaria

Abstract — The paper presents some experiments investigating the applicability of the Principal Component Analysis method for solving several concept learning tasks defined on images of faces. The results have shown that, in most cases, the applied transformation improves the classification accuracy of used concept learning algorithms. In addition, the experiments have confirmed a possible relation between the quality of the obtained improvements and the complexity of the concepts to be learnt. This relation has the potential to be an objective measure of “concept complexity”.

Keywords: Principal component analysis, Concept learning, Eigenfaces, Concept complexity.

I. INTRODUCTION

One of the hardest tasks the modern computer science tries to solve is making intelligent agents more aware of the environment around them. This usually involves accounting for different properties and detecting different features. An important subtask in this huge domain is helping the intelligent agents to assign images they perceive into some predefined classes so that they can take corresponding predefined actions.

Image processing is a resource intensive task by default. It comprises of analyzing millions of picture pixels and detecting features that they form. Thus, the tasks for classifying images as well as learning different concepts from them are naturally very complex too. It has been more than a century since Pearson first introduced the Principal Component Analysis (PCA) as a method of reducing the dimensions of a large space without losing important information [1]. However, it took almost ninety years until the advantages of using this approach for image processing were discovered. Sirovich and Kirby [2] used PCA to represent images of faces in a low dimensional space. They also introduced the term “eigenface” which is an eigenvector in the special case of PCA applied to the space of face images. Turk and Pentland used this representation for solving face recognition tasks [3]. Following their work PCA has been also successfully applied for image reconstruction of partially obstructed faces [4], 3D face image reconstruction [5], analysis of whole body images and postures [6] and many others.

The relative novelty of combining PCA with image processing could explain why only very few works examine the problem of applicability of PCA to the task of concept learning from images. The only such attempt we are aware of is [7]; however it does not present any concrete results. Aiming to fill this gap, this paper presents the results of several experiments on the applicability of PCA for solving concept learning problems defined over face images. The remaining sections of the paper are organized as follows: Section 2 presents our experiment setup. Section 3 describes the results of the conducted experiments. Section 4 is devoted to discussion and proposes an approach to measure the complexity of a concept to be learnt. Section 5 presents the conclusions and our intentions for possible future enhancements of the proposed approach.

II. EXPERIMENT SETUP

A. Data preparation

For the purposes of the experiments 102 frontal images of volunteers were collected. The people on the images are diverse in age, gender and face characteristics such as hair length, presence of moustaches and glasses, etc. (Fig. 1). The images were subjected to an algorithm [8] extracting the face bounding boxes with the main face characteristics (eyes, mouth and ears) aligned for each image. Such an alignment is crucial for improving the results of PCA application.

B. Principal component analysis transformation

PCA defines a transformation on a set of vectors producing the eigenvectors that are consequently used to define a lower-dimension space, in which the initial vector set should be projected. PCA guarantees that the axes of the new space are chosen in a way that maximizes the dispersion, so it can be considered as “trying to achieve the largest possible differentiation”. After the transformation PCA advocates that only the first few dimensions can be kept, dropping the others, whilst the most significant properties of the analyzed set will be preserved.

![Fig. 1. Examples of images used in the experiments after aligning their main characteristics](image1.png)
When the input is a set of face images the eigenvectors that PCA finds can be interpreted as grayscale images if they are normalized to have all their coordinates in the range 0...255. If the vectors are structured in a matrix of the same dimensions as the input images and then visualized as grayscale images they start resembling human faces as shown on Fig. 2, this is why the eigenvectors are called eigenfaces in the case of PCA applied to images of faces.

The output of the face bounding algorithm is a set of 110x150 pixels grayscale face images, so each image consists of 16500 pixels. Direct application of PCA transformation on this set will require computational complexity of the order of the cube of the number of pixels in an image. For such big images this will require too many operations and memory. That is why Turk and Pentland [3] have proposed an optimization based on the fact that the analyzed set comprises only a small number of images. As a result they have executed an auxiliary PCA that allows finding the eigenfaces of the image set in complexity of the order of the cube of the number of images in the set. The same optimization has been used for the experiments described in this paper.

Fig. 2. Visualization of the first few eigenfaces found when applying PCA to the input set.

Choosing the number of eigenfaces to be kept is always a challenging task and that is why the results for different number of preserved eigenfaces will be shown. In the next section we will discuss the effect of changing this parameter in more details.

C. Concept learning tasks used in the experiments

In order to evaluate the effect of PCA application the following concept learning tasks have been formulated:

- **Face recognition in presence of noise** – in this task the algorithm receives as an input one of the images from the training set with some random noise applied to it and must determine the initial image. The amount of noise is used as a task parameter. This task is one of the targets in the work of Turk and Pentland [3].
- **Learning the concept of gender** – the algorithm should determine the gender of the person whose face is presented on the analyzed photo.
- **Learning the concept of beauty** – the algorithm should determine whether the person depicted on the analyzed photo is beautiful or not. Since this is a subjective criteria, two ‘experts’ (A and B) were asked to make preliminary classification of the beauty of every image in the training set. Thus two classifications ‘beautiful according to A’ and ‘beautiful according to B’ were defined.
- **Learning a random concept** – in this task every image from the training set was randomly assigned to one of two classes. Because there is no relation of an image and its classification it is expected that all classifiers will behave randomly on this task.

D. Measuring the effect of PCA

The difference in the classification accuracy achieved by classifiers before and after application of PCA to the training set of images was used as a measure for effectiveness of the transformation for solving the mentioned above concept learning tasks. In all experiments a weighted k nearest neighbors classifier [9] was used as a concept learning algorithm. The parameter k was chosen in a way that maximizes the classification accuracy on the training set. Each of the k nearest neighbors got assigned weight of one over the square root of the Euclidean distance to it from the object to be classified.

The classification accuracies were evaluated by t-resampling test executed with 30 iterations. At each iteration 70% of examples were used as a training set and the rest 30% - as a testing set. The t-test outputs the difference $d$ in classification accuracy of compared algorithms and statistic $t$ that evaluates the confidence level of the achieved results. Even though Dietterich has argued that this is an imprecise procedure [11], we have followed Mitchell, who has claimed that this is the best approach if the amount of samples is too small [10].

III. EXPERIMENT RESULTS

A. Face recognition in presence of noise

In this experiment each image from the training set was used as a test example. The selected image was randomly corrupted according to the predefined level of noise and then
matched against the uncorrupted training set. Fig. 4 shows an example of such test image with different levels of noise applied to it. The nearest neighbor algorithm was used for classifying the corrupted image. The percent of correctly recognized faces was used as a measure of classification accuracy of the algorithm. The classification accuracy of the algorithm applied to the PCA-transformed set of images as function of the noise level and the number of preserved eigenfaces is shown in Table 1.

![Image](image.png) An example of a test image with different levels of noise: A) 0 B) 0.1 C) 0.3 D) 0.5

The comparison of the behavior of the classification algorithms with and without application of PCA has shown an interesting fact – the results were the same for any level of noise, when all possible eigenfaces were used. Moreover in that case the classifications were coincided for each test image.

TABLE I. ACCURACY OF THE FACE RECOGNITION TASK AT DIFFERENT LEVELS OF NOISE AND NUMBER OF PRESERVED EIGENFACES

<table>
<thead>
<tr>
<th>Noise / Eigenfaces</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>100%</td>
<td>73.53%</td>
<td>42.16%</td>
<td>23.53%</td>
<td>12.75%</td>
</tr>
<tr>
<td>10</td>
<td>100%</td>
<td>99.02%</td>
<td>77.45%</td>
<td>48.04%</td>
<td>28.43%</td>
</tr>
<tr>
<td>15</td>
<td>100%</td>
<td>100%</td>
<td>86.27%</td>
<td>56.86%</td>
<td>33.33%</td>
</tr>
<tr>
<td>20</td>
<td>100%</td>
<td>100%</td>
<td>89.22%</td>
<td>62.75%</td>
<td>36.27%</td>
</tr>
<tr>
<td>102</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>94.12%</td>
<td>63.73%</td>
</tr>
</tbody>
</table>

These results have shown that PCA can be successfully applied for solving the face recognition task since it does not decrease the classification accuracy while significantly increases the computational effectiveness of the nearest neighbor classifier. In our case the PCA-based representation of each image contained only 102 real coefficients instead of the initial 16500 pixels.

B. Learning a person gender from facial photo

In these experiments the distribution of examples in the two classes is almost equal – about 56% of the examples are photos of men and about 44% - of women (see Table 2). The classification accuracy of the weighted k-NN algorithm achieved after application of PCA transformation as function of the number of preserved eigenfaces is shown in Table 3. It can be seen that the results are significantly higher in comparison with those of the default classifier (which classifies each test example to the predominant class). Furthermore, the precision is not linearly dependent on the number of eigenfaces preserved, but the results are relatively close.

Application of the t-resampling test has shown that the average improvement of the weighted k-NN algorithm after applying PCA over the same algorithm without using PCA transformation is $d = 0.5$ and $t = 1.874$ which corresponds to about 94% confidence level.

TABLE II. ACUURACY OF ALGORITHMS THAT DO NOT USE PCA

<table>
<thead>
<tr>
<th>Concept \ Property</th>
<th>Male</th>
<th>Female</th>
<th>Default classifier accuracy</th>
<th>Wk-NN accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>57</td>
<td>45</td>
<td>55.88%</td>
<td>84.33%</td>
</tr>
</tbody>
</table>

TABLE III. ACCURACY OF GENDER CLASSIFICATION AT DIFFERENT NUMBER OF EIGENFACES

<table>
<thead>
<tr>
<th>Eigenfaces</th>
<th>4</th>
<th>10</th>
<th>20</th>
<th>41</th>
<th>60</th>
<th>102</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wk-NN (%)</td>
<td>83.67</td>
<td>83.25</td>
<td>83.67</td>
<td>84.83</td>
<td>84.25</td>
<td>83.67</td>
</tr>
</tbody>
</table>

These results have proved that application of PCA is beneficial for solving the task of learning the person gender from photos. Moreover, the transformation limits the analysis to a rather smaller space, which makes the computations significantly faster.

C. Learning the concept of beauty

As it has already been discussed the concept of beauty is very subjective and its definition considerably differs from one person to another. That is why the experiments have been conducted based on classifications given by two ‘experts’ who have defined the concepts ‘Beautiful person according to A’ and ‘Beautiful person according to B’. As it can be seen from Table 4 the experts have very different understanding of beauty with B being a lot more conservative than A. Expert B has found only 23% of the presented people as beautiful, while this percentage for A is 43%. This fact has proved again that the task for learning the concept of beauty is rather hard to define.

TABLE IV. DISTRIBUTION OF THE CLASSES IN THE BEAUTY LEARNING TASK ACCORDING TO A AND B AND THE ACCURACY OF ALGORITHMS THAT DO NOT USE PCA

<table>
<thead>
<tr>
<th>Classification</th>
<th>Beautiful</th>
<th>Not beautiful</th>
<th>Default classifier accuracy</th>
<th>Wk-NN accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>44</td>
<td>58</td>
<td>56.86%</td>
<td>53.67%</td>
</tr>
<tr>
<td>B</td>
<td>23</td>
<td>79</td>
<td>77.45%</td>
<td>82.44%</td>
</tr>
</tbody>
</table>

Table 5 shows the classification accuracy of the weighted k-NN algorithm on the two tasks formulated above as function of the number of preserved eigenfaces. It can be seen that for the case A the best result has been obtained when preserving only the first eigenface. This means that this eigenface carries most of the important information that ‘expert’ A used to distinguish beautiful people. All other eigenfaces can be interpreted as irrelevant features for this task that only decrease the performance of the classifier. In case B the number of eigenfaces leading to the best result is a bit larger - 15, but still remains relatively small. That has shown that the concept of beauty formulated by ‘expert’ B is based on inspecting more features than that of A since it requires more eigenfaces for solving the ‘beauty definition’ task.
The comparison of the classification algorithms with and without PCA applied gives the following results:

- For case A the classification accuracy improved with 6.83 after applying PCA. \( t \) has been found to be 2.738 which corresponds to confidence level over 99%. This has shown that PCA transformation is quite suitable for solving the learning task defined by ‘expert’ A.
- For the case B the classification accuracy has been improved by 1.44 and \( t \) has been found to be 1.631, which corresponds to confidence level of 88%. Such a confidence is too low and should be considered as insufficient from statistical point of view. However, there are no indications that application of PCA leads to decrease of the classification accuracy and having in mind that PCA immensely increases the algorithm performance, it still seems worth to apply this transformation for the task defined by ‘expert’ B too.

Another thing that should be mentioned in the context of the beauty task is that the improvement compared to the naive approach (i.e. default classification) is far from what we have got in the gender task. For both beauty definition cases the improvement is about 5%. A possible explanation is that the concept of beauty is much more complex and could be harder to identify.

D. Learning a random concept

A random concept is supposed to be impossible to be learnt since the concept classification does not depend on any features of the analyzed objects. Thus, despite of the number of training examples, the classification results cannot be improved. In order to simulate a random concept we have randomly distributed all training examples in two classes. Such a distribution is shown in Table 6.

As it can be seen class 1 has a little bit more representatives than class 2. It should be mentioned that this random class distribution has coincided with the relative distribution of classes in the gender task. This fortunate fact has allowed us to make more precise comparison on the way the algorithm behaves on different concept learning tasks.

Table 7 shows how the PCA-based classifier has performed on the random concept learning task. As it can be seen the classification accuracy does not vary more than in the previous tasks, but its value is rather lower than what the default classifier would have achieved.

When comparing the classifications with and without PCA we obtain improvement of 1.89 and confidence coefficient \( t = 1.32 \). This coefficient is really low and could be easily explained by the total unpredictability of the analyzed set.

Table 8 summarizes the results of all described above experiments by showing the percentage of improvement in the classification error when using PCA transformation before classifying. This table confirms the overall conclusion that the PCA transformation positively influences the results of the concept learning tasks.
features from a face image. People cannot even verbally express the parameters they consider when evaluating beauty. That is why we consider the task for learning the concept of beauty as more complex than learning the concept of gender.

- The task of learning a random concept is the most complex among considered ones since the training examples, in practice, do not contain any relevant features that can be effectively used for solving it.

This hypothetical ordering of the complexity of the analyzed concepts is confirmed by the results shown in Table 9. It presents the five concept learning tasks considered in the paper. For each task we have shown the classification accuracy of the default classifier and the accuracy achieved by the PCA-based classifier, choosing the number of used eigenfaces to be the one that produces the highest accuracy. For the face recognition task the noise level is equal to 0.3.

We have calculated the improvement in classification accuracy in comparison with that of the default classifier. As it can be seen the percentage of improvement greatly decreases with the increase of the complexity of the concept to be learnt. The improvement in the face recognition task is increased greatly. The gender learning task faces 50% improvement while the beauty learning tasks – about 8%. The accuracy of the PCA-based classifier is actually decreased in the random concept learning task.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Task</th>
<th>Class 1</th>
<th>Class 2</th>
<th>Default classifier</th>
<th>The best PCA-based classifier</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image recognition</td>
<td>n/a</td>
<td>n/a</td>
<td>1%</td>
<td>100%</td>
<td>10 000%</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>57</td>
<td>45</td>
<td>55.88%</td>
<td>84.83%</td>
<td>52%</td>
<td></td>
</tr>
<tr>
<td>Beauty A</td>
<td>44</td>
<td>58</td>
<td>56.86%</td>
<td>60.83%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>Beauty B</td>
<td>23</td>
<td>79</td>
<td>77.45%</td>
<td>84.17%</td>
<td>9%</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>45</td>
<td>57</td>
<td>55.88%</td>
<td>51.17%</td>
<td>-8%</td>
<td></td>
</tr>
</tbody>
</table>

There are other works in the area of evaluation of concept complexity. For example, Esposito [12] also has made some empirical tests proving that the complexity of a concept is in inverse proportion to the accuracy that any classification algorithm could achieve on the task. The results presented in this paper have confirmed his theory. On the other hand this would not be true if we have attempted to analyze the results achieved by the classifier without applying PCA. For example, the classification accuracy of the weighted k-NN algorithm without PCA achieved on solving “Beauty case A” task is lower than that of the default classifier.

In this section we briefly present a hypothesis on how the proposed approach might be used to measure the complexity of a concept to be learnt. Further investigations in the area are required to determine if our observations are just random coincidence or can be generalized to a formal approach to such estimation.

V. CONCLUSIONS AND FURTHER WORK

In this paper we have discussed the applicability of the Principle Component Analysis to the tasks related to learning different abstract concepts from images of faces. The conducted experiments have shown the potential of this transformation. Additionally a possible approach for measuring the complexity of the concept to be learnt has been briefly discussed.

There are many improvements that can be added to the current work. For example, more precise and reliable results could be expected if more training examples are used in experiments. Another possible direction of experiments is to apply the same approach to images of objects different from faces. For example, such images could be whole body figures. It will be interesting to see if the presented results will be repeated in such experiments.

The briefly mentioned idea for measuring the concept complexity will also require more thorough research. Defining new tasks for learning different concepts from images and investigating the results might prove or disprove the initial theory.

ACKNOWLEDGMENT

The work on this paper was partially supported by EC FP7-REGPOT-2012-2013-1 project “AComIn: Advanced Computing for Innovation”.

REFERENCES

[8] B. Strandjev, Investigation of the applicability of the principal component analysis method for image classification tasks, Ms Thesis, Faculty of Mathematics and Informatics, Sofia University, 2012 (in Bulgarian).