Face Classification via Sparse Approximation

Elena Battini Sönmez\textsuperscript{1}, Bülent Sankur\textsuperscript{2}, and Songul Albayrak\textsuperscript{3}

\textsuperscript{1} Computer Science Department, Bilgi University, Dolapdere, Istanbul, TR
\textsuperscript{2} Electric and Electronic Engineering Department, Boğaziçi University, Istanbul, TR
\textsuperscript{3} Computer Engineering Department, Yıldız Teknik University, Istanbul, TR

Abstract. We address the problem of 2D face classification under adverse conditions. Faces are difficult to recognize since they are highly variable due to such factors as illumination, expression, pose, occlusion and resolution. We investigate the potential of a method where the face recognition problem is cast as a sparse approximation. The sparse approximation provides a significant amount of robustness beneficial in mitigating various adverse effects. The study is conducted experimentally using the Extended Yale Face B database and the results are compared against the Fisher classifier benchmark.

Keywords: Face classification, sparse approximation, Fisher classifier.

1 Introduction

Automatic identification and verification of humans using facial information has been one of the most active research areas in computer vision. The interest on face recognition is fueled by the identification requirements for access control and for surveillance tasks, whether as a means to increase work efficiency and/or for security reasons. Face recognition is also seen as an important part of next-generation smart environments, \cite{1, 2}.

Face recognition algorithms under controlled conditions have achieved reasonably high levels of accuracy. However, under non-ideal, uncontrolled conditions, as often occur in real life, their performance becomes poor. Their main handicaps are the changes in face appearances caused by such factors as occlusion, illumination, expression, pose, make-ups and aging. In fact the intra-individual face differences due to some of these factors can easily be larger than the inter-individual variability \cite{3, 6}.

We briefly point out below some of the main roadblocks to wide scale deployment of reliable face biometry technology.

Effects of Illumination: Illumination changes can vary the overall magnitude of light intensity reflected back from an object and modify the pattern of shading and shadows visible in an image, \cite{6}. It was shown that varying illumination is most detrimental to both human and machine accuracies in recognizing faces. It suffices to quote simply the fact in FRGC face recognition, the 17 algorithms competing in the controlled illumination track have achieved a median verification rate of 0.91 while in contrast, the seven algorithms competing in the
uncontrolled illumination experiment have achieved a median verification rate of only 0.42 (both figures at a false acceptance rate of .001). The difficulties posed by variable illumination conditions, therefore, remain still one of the main roadblocks to reliable face recognition systems.

**Effects of Expression:** Facial expression is known to affect the face recognition accuracy though in the current literature, a full-fledged analysis of the deterioration caused by expressions has not been documented. Instead, most studies either focus on expression recognition alone or on face identification alone. It is quite interesting that this dichotomy is also encountered in biological vision. There is strong evidence that facial identity and expression might be processed by separate systems in the brain, or at best they are loosely linked, [7].

**Effects of Pose:** Facial pose or viewing angle is a major impediment to machine-based face recognition. As the camera pose changes, the appearance of the face changes due to projective deformation (causing stretching and foreshortening of different part of face); also self-occlusions and/or uncovering of face parts can arise. The resulting effect is that image-level differences between two views of the same face are much larger than those between two different faces viewed at the same angle. While machines fail badly in face recognition under viewing angle changes, that is, when trained on a gallery of a given pose and tested with a probe of set of a different viewing angle, [8], humans have no difficulty in recognizing faces at arbitrary poses. It has been reported in [8] that the performance of PCA-based method decreases dramatically beyond 32 degree yaw and those for LDA beyond 17 degree rotation.

**Effects of Occlusion:** The face may be occluded by facial accessories such as sunglasses, a snow cap, a scarf, by facial hair or other paraphernalia. Furthermore the subjects in an effort to eschew being identified can purposefully cover parts of their face. Although it is very difficult to systematically experiment with all sorts of natural or intentional occlusions, results reported in [8] shows that methods like PCA and LDA fail quite badly (e.g., sunglasses and scarf scenes in AR database). Recent work on recognition by parts shows that methods that rely on local information, can perform fairly well under occlusion, [4], [5], [9].

**Effects of Low Resolution:** The performance loss of face recognition with decreasing resolution is well known and documented, [10]. For example, 30 percentage point drops are reported in [10] as the resolution changes from 65 × 65 to 32 × 32 faces. The robustness to varying resolution becomes relevant especially in uncontrolled environments where the face may be captured by a security camera at various distances within its field of view.

In this paper we consider the effects of the resolution, the ones of degree of over-completeness, as well as the illumination compensation capability, the robustness to noise and to planar rotation of a recently introduced sparse approximation based classification algorithm. That is, we investigate the robustness of a non-linear face recognition method, called Sparse Representation-based Classifier (SRC), [9], *vis-à-vis* the well-known Fisher linear discriminant (FLDA) Classifier. The rationale of this approach is to use an over-complete dictionary whose base elements consist of training samples themselves, and to search for a