Digital Signal Processing **10**, 237–248 (2000)
doi:10.1006/dspr.1999.0358, available online at http://www.idealibrary.com on IDEAL®

Applying Logistic Regression to the Fusion of the NIST'99 1-Speaker Submissions

Stéphane Pigeon, Pascal Druyts, and Patrick Verlinde

SIC Laboratory, Royal Military Academy, Av. Renaissance 30,
1000 Brussels, Belgium

E-mail: stephane.pigeon@elec.rma.ac.be; pascal.druyts@elec.rma.ac.be; and patrick.verlinde@tele.rma.ac.be

Pigeon, Stéphane, Druyts, Pascal, and Verlinde, Patrick, Applying Logistic Regression to the Fusion of the NIST'99 1-Speaker Submissions, *Digital Signal Processing* **10** (2000), 237–248.

This contribution formulates the decision fusion problem encountered in the design of a multiexpert identity verification system. For this purpose, a Bayesian classifier is used and the influence of the a priori class probabilities is investigated. Logistic regression is introduced as a particular case of the naive Bayesian classifier and is applied to the fusion of all submitted data of the 1-speaker recognition task, part of the NIST'99 campaign. © 2000 Academic Press

Key Words: decision fusion; Bayes; logistic regression.

0. INTRODUCTION

The automatic verification of a person's identity is becoming an important task in several applications, especially in the field of automatic access to restricted physical or virtual environments. Passwords, personal magnetic cards, and PINs are already widely used in this context. Although they are quite convenient to use, they can be forgotten, lost, or stolen. Therefore, a new kind of method is emerging, based on so-called *biometric* measures such as vocal (speech), visual (face/profile), fingerprint, or any other information that physically characterizes the person to be identified. A system based on a particular set of biometric features is referred to as an *expert* for that biometric feature set. To gain robustness, multimodal systems tend to combine several experts together, such as the frontal, facial, and speech experts used in [10]. In our contribution, all experts used are taken from among the NIST'99 evaluation campaign submissions and thus refer to the same speech modality. Although we cannot expect as much improvement by fusing speech-related experts as compared to the fusion independent biometrics, it is interesting to see how

1051-2004/00 $35.00

fusion improves the overall system performance, even in such a particular case as such the one we are working with in this paper.

The first section introduces the decision fusion framework adopted in this work. Section 2 reminds the reader about the general Bayesian theory. Section 3 introduces the independence hypothesis that will be used to derive the logistic regression model presented in Section 4. Section 5 illustrates the influence of the a priori probabilities that are inherent to the Bayesian approach. Section 6 presents the results of fusing all NIST'99 speech experts and, finally, Section 7 concludes this work.

1. DECISION FUSION IN AN IDENTITY VERIFICATION SYSTEM

The purpose of an identity verification system is to decide whether someone claiming the identity of a registered user is indeed that client or an impostor. In a mono-modal system, this is done by comparing the score obtained for that person with a decision threshold. Such a system can make two types of errors: (i) reject a client (i.e., *false rejection*—FR—or *miss*) and (ii) accept an impostor (i.e., *false acceptance*—FA—or *false alarm*). The performance of a speaker verification system is usually given in terms of global error rates computed during tests, namely the *false rejection rate* (FRR—the number of FR divided by the number of client claims) and the *false acceptance rate* (FAR—the number of FA divided by the number of impostor claims) [1]. The *equal error rate* (EER) stands for the operating point at which the FAR and FRR are the same.

One possible and straightforward way of building a multi-modal verification system from n such mono-modal system is to input all n scores provided in parallel into a fusion module which has to take the decision to *accept* or *reject* the claim. This is a typical *decision fusion* approach, in which the fusion module receives as input the *decisions* issued by the several individual experts, and typically has access to neither the input feature vectors of these experts (*feature fusion*) nor the original raw data streams (*data fusion*) [3]. Once the choice of a particular fusion scheme has been made, two main alternatives still remain for the fusion module: a global (i.e., the same for all persons) or a personal (i.e., tailored to the specific characteristics of each authorized person) approach. For the sake of simplicity and since the personal approach requires much more training data, we opted for a global fusion module. As, in a verification system dealing with n modalities, the fusion module has to realize a mapping from $\mathcal{R}^n$ into the binary set {*reject, accept*}, this can be seen as a multi-dimensional classification problem, splitting an n-dimensional space into two classes. Bayesian classifiers will be introduced in the next section.

2. BAYESIAN FRAMEWORK FOR DECISION FUSION

In a number of references, such as [4], a general overview of Bayesian decision theory is presented in the case of the classification problem. We give here only