Digital Signal Processing 50 (2016) 24–33

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

On the fusion of non-independent detectors

Luis Vergara *, Antonio Soriano, Gonzalo Safont, Addisson Salazar

Universitat Politècnica de València, Valencia, Spain

ARTICLE INFO

Article history:
Available online 17 December 2015

Keywords:
Decision theory
Decision fusion
Dependence factors
Optimum fusion

ABSTRACT

Independence between detectors is normally assumed in order to simplify the algorithms and techniques used in decision fusion. In this paper, we derive the optimum fusion rule of N non-independent detectors in terms of the individual probabilities of detection and false alarm and defined dependence factors. This has interest for the implementation of the optimum detector, the incorporation of specific dependence models and for gaining insights into the implications of dependence. This later is illustrated with a detailed analysis of the two equally-operated non-independent detectors case. We show, for example, that not any dependence model is compatible with an arbitrary point of operation of the detectors, and that optimality of the counting rule is preserved in presence of dependence if the individual detectors are "good enough". We have derived also the expressions of the probability of detection and false alarm after fusion of dependent detectors. Theoretical results are verified in a real data experiment with acoustic signals.

1. Introduction

Fusion of detectors is a well-established issue, appearing in such related areas as sensor data fusion [1], multimodal fusion [2], mixture of experts [3] and classifier combiners [4]. Although frequently employing different terminologies and experiencing distinct implementation constraints, all of these areas share similar problems when considering the optimum design of fusion methods. Three different levels of fusion can be carried out: observation (in classification, "feature" is preferred) fusion, score fusion and decision fusion. In principle, observation fusion should be privileged, as it captures all the underlying statistical information about the problem. However, all the original observation components might not be accessible at the fusion center. This occurs in sensor networks, where transmission bandwidth conservation and distributed processing lead to only individual decisions being transmitted to the fusion center. Moreover, even given simultaneous accessibility to all the observation components, the problem of estimating the multidimensional probability densities (MPD) required for optimum, likelihood-ratio-based observation fusion, remains a complex one. This is especially the case when dealing with heterogeneous observations and/or statistical dependence between the observation vector components. Score fusion alleviates the problem of heterogeneity, as the scores afforded by each individual detector include some type of normalization: the scores are generally estimates of the *a posteriori* probability of every hypothesis derived from an observation. But in areas such as biometrics, scoring normalization is required prior to fusion [5]. An additional advantage garnered from score fusion is that the number of components to be fused is limited to the number of detectors. In any case, scores, like observations, are continuous variables that lend a significant degree of complexity to the task of estimating the underlying MPD, and might not be available and could not be accessible in distributed detection architectures. Consequently, decision fusion is the ideal choice when the observations or the scores are unavailable at the fusion center, and/or when MPD estimation is to be avoided or replaced by the (simpler) estimation of multidimensional probability masses (MPM).

A great deal of effort has been dedicated over the past few decades to finding optimum fusion methods at the three different levels. Some of this research, mostly recent, has assumed the presence of statistical dependence, so factoring the corresponding and underlying MPD or MPM in unidimensional marginals is not possible. In [6], a copula-based approach is therefore presented to fuse heterogeneous observations. Copulas are useful in managing heterogeneity when dependence is present: the MPD is factorized in the marginals, thereby capturing the heterogeneity, and a multidimensional copula (MPD of uniformly distributed variables derived from the original observation components) captures the dependence. Copulas are also useful in defining a variety of dependence models. Reference [7] is representative of score fusion, where the authors use non-parametric estimates of the score MPD. Optimum fusion at the decision level requires knowing the MPM of deci-

* Corresponding author. Fax: +34963877309.
E-mail address: lvergara@dcom.upv.es (L. Vergara).

http://dx.doi.org/10.1016/j.dsp.2015.11.009
1051-2004/

sions made by the individual detectors. This can be obtained by multidimensional integration of the MPD in limits defined by the individual thresholds. Thus in [8], a parametric copula model (with known or unknown parameters) is assumed for the MPD, then the corresponding MPM are derived. In [9], the work of [8] is extended to the case of "mis-specified" copulas and multibit quantizers. Optimum design of the individual detector rules, in conjunction with the optimum definition of the decision fusion rule when dependence is present, have also been analyzed by several authors. In [10], for example, the two-sensor case is considered for the binary-quantizer Gaussian shift-in-mean problem. Determining optimal individual rules requires an iterative algorithm, where only one rule is modified at a time while the others are assumed to be fixed. Conclusions are then derived about the convergence properties of the algorithm for possible two-sensor fusion rules (AND, OR, XOR). The authors in [11] find the general expression for an optimum individual rule, given the other rules; to avoid the iterative, one-by-one search for optimal individual rules, the authors propose transforming the original observations to undo the statistical dependence and thereby render individual likelihood ratio tests optimal. Though ambitious in their quest for optimality, these works do present various drawbacks:

- It is generally assumed that the input of the individual detector is only one of the components from the entire observation vector. While this may be a reasonable assumption in distributed detection, it is not usually the case in the aforementioned related areas (see [12] for an example of fusion in the financial area), where highly dimensional patterns appear at the input of every individual detector.
- Different amounts of available information about the MPD of the observations are required or must be estimated. In [8] and [9], for example, the marginals are assumed to be known, while the copulas are known or must be estimated. In [10], Gaussianity is assumed, while in [11], the proposed transformation depends on the specific statistical model considered and may not be feasible for arbitrary scenarios.
- The algorithms are complex and mostly iterative. Hence, applying these techniques is only justified if the observations are not simultaneous available at the fusion center—not as a means of avoiding use of the MPD.
- Thus, these techniques do not generally offer an intuitive and simple understanding of the implications of dependence for the optimality of the fusion center rules.

Simpler approaches are based on working directly with the MPM, or with equivalent representations of both the marginals and the dependence of the decisions. The optimum decision fusion rule employing the MPM is well-known (see Section IV.A of [8] and [13]). Optimum fusion rules are also proposed in [14], employing correlation coefficients of decisions and probabilities of false alarm and detection for some specific correlation structures. Based on the results from [14,15] derives the optimal conditions for simple counting rules in the case of identical detectors, where every hypothesis is characterized by a constant correlation coefficient. Practical application of these methods requires estimating the MPM or equivalent representations. This estimation can be made from sample training records of synchronized decisions under every hypothesis.

The work presented in this paper belongs to this class of approaches. Each individual detector will be characterized by its point of operation (probability of detection PD and probability of false alarm PFA). Dependence between decisions in every hypothesis will be captured by a new defined parameters: dependence factors, DFs. These parameters are factors linking the marginal masses with the MPM in much the same way that the copula function links the marginals with the MPD, although by no means the DFs exhibit properties similar to those of copula functions [16]). We will obtain the optimum fusion rule in terms of the individuals PD and PFA and on the DFs. This facilitate the implementation of the optimum fused detector as the nature and complexity of every detector is thus of no concern, only its PD and PFA, and the DFs can be estimated from training records. Moreover, we will see that the optimum fusion rule in presence of dependence is an obvious extension of the optimum fusion rule of independent detectors. This makes possible a straightforward consideration of particular dependence models. It also allows gaining insights into the implications of dependence in different aspects such as optimality of the counting rule, detector performance or compatibility between the dependence model and the points of operation of the individual detectors.

The remainder of this paper is organized as follows. In Section 2, it is derived the optimum fusion rule in terms of the individual PD and PFA, and the DFs. Then, Section 3 focus in the two equally-operated non-independent case; this particularization is a convenient assumption for mathematical tractability, and it does not inhibit deriving interesting conclusions about the influence of the statistical dependence in the decision fusion problem. Section 4 provides three representative examples to illustrate the theory and a real data application to verify the theoretical predictions, with conclusions discussed in Section 5.

2. Optimum fusion rule in terms of dependence factors

Let us consider the case of N detectors indexed by $n = 1 \ldots N$. Every detector generates a binary decision $u_n = 1, 0$ and is characterized by some particular probability of detection P_{dn} and probability of false alarm P_{fn}. We define the vector of decisions $\mathbf{u} = [u_1 \ldots u_N]^T$. The optimum fusion rule is given by [17]

$$\Lambda(\boldsymbol{u}) = \frac{P_{H_1}(\mathbf{u})}{P_{H_0}(\mathbf{u})} \underset{H_0}{\overset{H_1}{\gtrless}} t. \tag{1}$$

Where $P_{H_j}(\mathbf{u})$ is the MPM corresponding to hypothesis H_j and t is a threshold which determines the PD and PFA. $\Lambda(\boldsymbol{u})$ is the likelihood ratio. In the following we are going to express the optimum fusion rule in terms of the individual PD and PFA of every detector and of some properly defined dependent factors.

Let us define also the truncated vector of decisions $\mathbf{u}_n = [u_n \ldots u_N]^T$, $n = 1 \ldots N$, hence $\mathbf{u}_1 \equiv \mathbf{u}$ and $\mathbf{u}_N \equiv u_N$. Using the probability chain rule, the MPM corresponding to the hypothesis H_j may be expressed in the form

$$\begin{aligned} P_{H_j}(\mathbf{u}) &= P_{H_j}(u_1 \mid \mathbf{u}_2) P_{H_j}(u_2 \mid \mathbf{u}_3) \ldots P_{H_j}(u_{N-1} \mid \mathbf{u_N}) P_{H_j}(u_N) \\ &= \frac{P_{H_j}(u_1 \mid \mathbf{u}_2)}{P_{H_j}(u_1)} \frac{P_{H_j}(u_2 \mid \mathbf{u}_3)}{P_{H_j}(u_2)} \ldots \frac{P_{H_j}(u_{N-1} \mid \mathbf{u_N})}{P_{H_j}(u_{N-1})} \\ &\quad \times \prod_{n=1}^{N} P_{H_j}(u_N) = \alpha_j(u_1 \ldots u_N) \prod_{n=1}^{N} P_{H_j}(u_N). \end{aligned} \tag{2}$$

In this way, the MPM is factorized in the marginals and in a dependence function $\alpha_j(u_1 \ldots u_N)$ in a manner similar to that performed with a copula function [16]. Now, let the DFs under H_j be defined by

$$\beta_{jn}(\mathbf{u}_{n+1}) = \frac{P_{H_j}(u_n = 1 \mid \mathbf{u}_{n+1})}{P_{H_j}(u_n = 1)}, \quad n = 1 \ldots N-1. \tag{3}$$

Notice that

$$\frac{P_{H_j}(u_n = 0 \mid \mathbf{u}_{n+1})}{P_{H_j}(u_n = 0)} = \frac{1 - P_{H_j}(u_n = 1 \mid \mathbf{u}_{n+1})}{1 - P_{H_j}(u_n = 1)}$$