Signal Processing 99 (2014) 136–146

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

The geometry of fusion inspired channel design

Yuan Wang [a,*], Haonan Wang [a], Louis L. Scharf [a,b]

[a] Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA
[b] Department of Mathematics, Colorado State University, Fort Collins, CO 80523, USA

ARTICLE INFO

Article history:
Received 22 June 2013
Received in revised form
6 October 2013
Accepted 14 December 2013
Available online 24 December 2013

Keywords:
Embedded submanifold
Information fusion
MIMO channel design
Mutual information
Two-channel system

ABSTRACT

This paper is motivated by the problem of integrating multiple sources of measurements. We consider two multiple-input–multiple-output (MIMO) channels, a primary channel and a secondary channel, with dependent input signals. The primary channel carries the signal of interest, and the secondary channel carries a signal that shares a joint distribution with the primary signal. The problem of particular interest is designing the secondary channel matrix, when the primary channel matrix is fixed. We formulate the problem as an optimization problem, in which the optimal secondary channel matrix maximizes an information-based criterion. An analytical solution is provided in a special case. Two fast-to-compute algorithms, one extrinsic and the other intrinsic, are proposed to approximate the optimal solutions in general cases. In particular, the intrinsic algorithm exploits the geometry of the unit sphere, a manifold embedded in Euclidean space. The performances of the proposed algorithms are examined through a simulation study. A discussion of the choice of dimension for the secondary channel is given.

1. Introduction

Consider the following two-channel system, as illustrated in Fig. 1,

$$\boldsymbol{x} = \boldsymbol{F}\boldsymbol{\theta} + \boldsymbol{u}$$
$$\boldsymbol{y} = \boldsymbol{G}\boldsymbol{\phi} + \boldsymbol{v}. \quad (1)$$

The first channel is the primary channel that carries the signal of interest $\boldsymbol{\theta}$. The secondary channel carries a signal $\boldsymbol{\phi}$ that shares a joint distribution with $\boldsymbol{\theta}$. The measurements $\boldsymbol{x}$ and $\boldsymbol{y}$ are linear transformations of the input signals with measurement noises $\boldsymbol{u}$ and $\boldsymbol{v}$, respectively. For example, the elements of the primary signal $\boldsymbol{\theta}$ may be the complex scattering coefficients of several radar-scattering targets and the elements of the secondary signal $\boldsymbol{\phi}$ may be intensities in an optical map of these same optical-scattering targets. The measurement $\boldsymbol{x}$ is then a range-doppler map and the measurement $\boldsymbol{y}$ is an optical image. Under a certain power constraint, we consider the design of the channel matrix $\boldsymbol{G}$ (or the precoder matrix assuming identity channel), with the primary channel fixed. This problem is an abstraction of the quite general problem of adding a new sensor suite (or communication channel) to an existing suite. The question is whether the added performance gain warrants the expense. The only constraint in this abstraction is that both channels are linear. Examples include cooperative radar, bistatic radar, radar plus optical, and so on. The application of this abstract model to a physical problem requires only the determination of the design problems of this paper. Notice that if both channels were to be designed, then a Gauss–Seidel iteration could be used, as follows:

Step 0: Initialize;
Step 1: Fix the secondary channel and update the primary channel. Then go to Step 2;
Step 2: Fix the primary channel, update the secondary channel. Return to Step 1.

The solution for step 1 can be found in many classical precoder design problems in the literature. The solution for step 2 would be based on the results in this paper.

* Corresponding author. Tel.: +1 970 491 3778.
E-mail addresses: wangy@stat.colostate.edu (Y. Wang), wanghn@stat.colostate.edu (H. Wang), scharf@engr.colostate.edu (L.L. Scharf).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.12.015

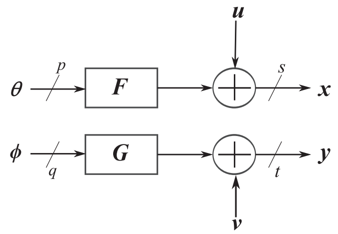

Fig. 1. A two-channel system with two linear channels.

This iterative procedure would guarantee a non-decreasing information rate and the results would converge to a local optimal solution. In the scope of this paper, we focus only on step 2, as channel one is fixed.

Various criteria have been studied for MIMO channel design problems. For example, signal-to-noise ratio (SNR) and signal-to-interference-noise ratio (SINR) [3,4], mutual information [5–8], and (weighted) minimum mean square error (MMSE) [9–11]. The criterion we use in this paper is the mutual information between the primary source signal and the measurements x and y. In the multivariate normal case, this is maximized by minimizing the log-determinant of the error covariance for the MMSE estimator. Without the assumption of multivariate normality, the determinant remains the volume of the error concentration ellipsoid. This volume is a scalar measure of how good this ellipsoid is. When the volume is small, estimates may be expected to lie near the true value of the source vector. Moreover, this information-based criterion is connected with estimation theory in a channel with Gaussian noise and arbitrary source distribution by linking the mutual information with MMSE, as in [12,13].

The linear precoding design for single or multiple MIMO channels has been studied in the literature (see [1–11] and references therein). While most of the work considered single or mutually independent sources, Fang et al. [1], Tang and Hua [2] have considered the joint precoder design with correlated sources signals. In the work of [1,2], both sources are of interest and the precoders are designed jointly to recover both source signals. In our problem, the secondary channel contains an auxiliary source and our interest focusses on the primary source. This makes our work different from prior work. That is, we have a one-channel design in a two-channel system, rather than a one-channel design or a two-channel design. The resulting precoder for the auxiliary channel does not maximize the differential rate at which it brings information about the auxiliary source, but it does maximize the differential rate at which the two channels bring information about the primary source.

Zhang et al. [28–29] have derived a suboptimal precoding matrix by a sequence of vector-variate optimization problems that maximizes the information gain brought by each row of the precoding matrix. Our matrix-variate problem, however, is more complex and in general, is not a convex problem. Moreover, this problem cannot be formulated as an SVD problem [21], in contrast to the one-channel system design. Analytical solutions are derived for some special cases. For general cases, we propose two gradient-based algorithms, one extrinsic and the other intrinsic, to approximate the optimal channel matrix. The extrinsic algorithm is a gradient-ascent algorithm with projection to the constrained space [15]. The intrinsic algorithm, a gradient-ascent algorithm of manifold, exploits the geometry that codes for the power constraint by vectorizing the channel matrix, [16–20]. Extrinsic and intrinsic algorithms are not entirely new. Our contribution is to implement a manifold optimization for our problem by exploiting the geometry of the power constraint.

The rest of the paper is organized as follows. We formulate the channelization problem in Section 2 and point out the challenges for design in a two-channel system. In Section 3, we give an analytical solution when the conditional covariance of $\boldsymbol{\phi}$ given $\boldsymbol{\theta}$ is the identity matrix. In Section 4, we propose two numerical algorithms, the extrinsic and intrinsic gradient searches, to approximate the optimal channel matrix for general cases. A simulation study is presented to illustrate the performance of the proposed algorithms in Section 4.3. In Section 5, we discuss the choice of number of measurements for the secondary channel. Section 6 concludes the paper.

Notation: The set of length m real vectors is denoted by $\mathbb{R}^m$ and the set of $m \times n$ real matrices is denoted by $\mathbb{R}^{m\times n}$. Bold upper case letters denote matrices, bold lower case letters denote column vectors, and italics denote scalars. The scalar x_i denotes the ith element of vector $\boldsymbol{x}$, and $\boldsymbol{X}_{i,j}$ denotes the element of $\boldsymbol{X}$ at row i and column j. The diagonal matrix with diagonal elements $\boldsymbol{x}$ is denoted as $\text{Diag}(\boldsymbol{x})$. The $n \times n$ identity matrix is denoted by $\boldsymbol{I}_n$. The transpose, inverse, trace and determinant of a matrix are denoted by $(\cdot)^T$, $(\cdot)^{-1}$, $\text{tr}(\cdot)$ and $\det(\cdot)$, respectively.

A covariance matrix is denoted by bold upper case $\boldsymbol{Q}$ with specified subscripts: $\boldsymbol{Q}_{zz}$ denotes the covariance matrix of a random vector $\boldsymbol{z}$; $\boldsymbol{Q}_{z_1 z_2}$ is the cross-covariance matrix between $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$; $\boldsymbol{Q}_{z_1 z_1|z_2}$ is the conditional covariance matrix of $\boldsymbol{z}_1$ given $\boldsymbol{z}_2$.

2. Overview

2.1. Problem statement

The two channels of the system described in (1) have input signals $\boldsymbol{\theta} \in \mathbb{R}^p$ and $\boldsymbol{\phi} \in \mathbb{R}^q$, respectively. The signal $\boldsymbol{\theta}$ is of key interest and $\boldsymbol{\phi}$ is a secondary signal that is jointly distributed with $\boldsymbol{\theta}$. The first channel $\boldsymbol{x} \in \mathbb{R}^s$ is a direct measurement of $\boldsymbol{\theta}$, while the secondary channel $\boldsymbol{y} \in \mathbb{R}^t$ is an indirect measurement of $\boldsymbol{\theta}$ through $\boldsymbol{\phi}$. Both $\boldsymbol{x}$ and $\boldsymbol{y}$ contain information about $\boldsymbol{\theta}$, and one can expect that fusing measurements from both channels would provide a better estimate than using a single measurement. Our interest is to design the channel matrix $\boldsymbol{G}$, with the first channel fixed, such that the rate at which $\boldsymbol{x}$ and $\boldsymbol{y}$ bring information about $\boldsymbol{\theta}$ is maximized.

We make the following assumptions:

(a1) The signals $\boldsymbol{\theta} \in \mathbb{R}^p$ and $\boldsymbol{\phi} \in \mathbb{R}^q$ are jointly Gaussian distributed as

$$\begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\phi} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\mu}_\theta \\ \boldsymbol{\mu}_\phi \end{pmatrix}, \begin{pmatrix} \boldsymbol{Q}_{\theta\theta} & \boldsymbol{Q}_{\theta\phi} \\ \boldsymbol{Q}_{\phi\theta} & \boldsymbol{Q}_{\phi\phi} \end{pmatrix} \right).$$

Here $\boldsymbol{Q}_{\theta\theta}$, $\boldsymbol{Q}_{\phi\phi}$ are the positive definite covariance for $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, $\boldsymbol{Q}_{\theta\phi}$ is the cross covariance between $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$.