GENERALIZED MUTUAL INTERDEPENDENCE ANALYSIS

Heiko Claussen*, *Justinian Rosca*

Siemens Corporate Research Inc.
755 College Road East
Princeton, NJ 08540

Robert Damper

School of Electronics and Computer Science
University of Southampton
Southampton SO17 1BJ, UK

ABSTRACT

The mean of a data set is one trivial representation of data from one class. Recently, mutual interdependence analysis (MIA) has been successfully used to extract more involved representations, or "mutual features", accounting for samples in the class. For example a mutual feature is a speaker signature under varying channel conditions or a face signature under varying illumination conditions. A mutual representation is a linear regression that is equally correlated with all samples of the input class. We present the MIA optimization criterion from the perspectives of regression, canonical correlation analysis and Bayesian estimation. This allows us to state and solve the above criterion concisely, to contrast the MIA solution to the sample mean, and to infer other properties of its closed form, unique solution under various statistical assumptions. We define a generalized MIA solution (GMIA) and apply MIA and GMIA in a text-independent speaker verification task using the NTIMIT database. Both methods show competitive performance with equal-error-rates of 7.5 % and 6.5 % respectively over 630 speakers.

Index Terms— Algorithms, Signal Processing, Pattern Classification, Signal Analysis, Speaker Recognition.

1. INTRODUCTION

Statistical signal processing methods such as Fisher's linear discriminant analysis (FLDA) [1], canonical correlation analysis (CCA) [2] or ridge regression [3] aim to model or extract the essence of a dataset. The goal is to find a simplified data representation that retains the information that is necessary for subsequent tasks such as classification or prediction. Each of the methods uses a different viewpoint and criteria to find this "optimal" representation. For instance FLDA is used to reduce the dimensionality of a dataset by projecting data points on a space that maximizes the quotient of the between- and within-class scatter of the training data. In this way, FLDA aims to find a simplified data representation that retains the discriminant characteristics for classification. On the other hand, CCA assumes one common source in two datasets. The dimensionality of the data is reduced by retaining the space that is spanned by pairs of projecting directions in which the datasets are maximally correlated. In contrast to this, ridge regression finds a linear combination of the inputs that best fits a known optimal response.

In this paper, we are interested in alternative criteria to extract an "optimal" dataset representation. We aim to extract an invariant representation of high dimensional instances of a single class. An approach that has been designed for this purpose is mutual interdependence analysis (MIA) ([4], [5]). We introduce MIA in section 2. In section 3 we relate MIA to CCA and FLDA to derive a more elegant MIA criterion. Section 4 is used to interpret MIA from a Bayesian perspective, to illustrate similarities and differences from ridge regression, and to find a generalized MIA solution (GMIA). In section 5 we apply MIA and GMIA to a text independent speaker verification problem. We achieve equal-error-rates (EERs) of 7.5 % and 6.5 % respectively for the full NTIMIT database [6] of 630 speakers. The paper concludes with a summary of our approaches and directions for future work.

2. MUTUAL INTERDEPENDENCE ANALYSIS

Throughout this paper, $\mathbf{x}_i^{(p)} \in \mathbb{R}^D$ denotes the ith input vector, $i = 1 \ldots N^{(p)}$ in class p. Furthermore, we use $\mathbf{X}^{(p)} \subseteq \mathbf{X}$ to represent a matrix with columns $\mathbf{x}_i^{(p)}$ and $\mathbf{X}$ to denote the matrix with columns $\mathbf{x}_i$ of all classes K. Moreover, $\mu = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i$, $\underline{\mathbf{1}}$ is a vector of ones, $\mathbf{Y} = \mathbf{X} - \mu \cdot \underline{\mathbf{1}}^T$ and $\mathbf{I}$ represents the identity matrix. The remaining notation will be clear from the context.

Assume that we wish to find a class representation $\mathbf{w}^{(p)}$ of high dimensional data vectors $\mathbf{x}_i^{(p)}$ $\left(D \geq N^{(p)}\right)$. A common first step is to select features and reduce the dimensionality of the data. However, because of possible loss of information, this preprocessing is not always desirable. Therefore, we aim to find a class representation of similar or same dimensionality as the inputs.

---

*Further affiliation of the author: School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK, hc05r@ecs.soton.ac.uk

The quality of such a representation can be evaluated by its correlation with the class instances. Our intuition is that a superior class representation is highly correlated and also has small variance of the correlations over all instances in the class. The former condition ensures that most of the energy in the samples is captured. The latter condition is indicative of membership in a single class. Note that only vectors in the span of the class vectors contribute to the cross-correlation value. Therefore, in the absence of prior knowledge, it is reasonable to constrain the search for a class representation $\mathbf{w}$ to the span of the training vectors $\mathbf{w} = \mathbf{X}^{(p)} \cdot \mathbf{c}$, where $\mathbf{c} \in \mathbb{R}^N$. This problem definition is the motivation for the MIA criterion proposed in [4].

The MIA result is defined as a direction $\mathbf{w}_{\mathrm{MIA}}^{(p)}$ that minimizes the projection scatter of the class p inputs:

$$\mathbf{w}_{\mathrm{MIA}}^{(p)} = \underset{\mathbf{w}, \mathbf{w}=\mathbf{X}^{(p)} \cdot \mathbf{c}}{\arg\min} \left(\mathbf{w}^T \cdot \mathbf{Y}^{(p)} \cdot \mathbf{Y}^{(p)^T} \cdot \mathbf{w} \right) \quad (1)$$

Note that the original space of the inputs spans the mean subtracted space plus possibly one additional dimension. By searching in the span of the original rather than mean subtracted inputs, a closed form solution of (1) can be found [4]:

$$\mathbf{w}_{\mathrm{MIA}}^{(p)} = \zeta \mathbf{X}^{(p)} \cdot \left(\mathbf{X}^{(p)^T} \cdot \mathbf{X}^{(p)} \right)^{-1} \cdot \underline{\mathbf{1}} \quad (2)$$

where ζ is a constant. The mathematical structure of this MIA solution has striking similarity with linear regression. Section 4 shows which assumptions distinguish the two problems.

3. ALTERNATIVE MIA CRITERION

In this section we will motivate the design of an alternative to (1), an unconstrained MIA criterion.

3.1. CANONICAL CORRELATION ANALYSIS

If two datasets $\mathbf{X} \in \mathbb{R}^{D \times N}$ and $\mathbf{Z} \in \mathbb{R}^{K \times N}$, of possibly different dimensionality, are influenced by a common source $\mathbf{s} \in \mathbb{R}^N$, canonical correlation analysis [2] can be used to extract this inherent similarity. The goal of CCA is to find two vectors to project the datasets such that their projection lengths are maximally correlated. Let $\mathbf{C_{XZ}}$ denote the cross covariance matrix between the datasets $\mathbf{X}$ and $\mathbf{Z}$. The CCA problem is given by maximization of the objective function:

$$J(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a}^T \cdot \mathbf{C_{XZ}} \cdot \mathbf{b}}{\sqrt{\mathbf{a}^T \cdot \mathbf{C_{XX}} \cdot \mathbf{a}} \cdot \sqrt{\mathbf{b}^T \cdot \mathbf{C_{ZZ}} \cdot \mathbf{b}}} \quad (3)$$

over the vectors $\mathbf{a}$ and $\mathbf{b}$. The intuition is that the maximally correlated projections $\mathbf{X}^T \cdot \mathbf{a}$ and $\mathbf{Z}^T \cdot \mathbf{b}$ represent an estimate of the common source.

Canonical correlation analysis can be used to extract classification relevant information from a set of inputs. Indeed, let

$\mathbf{X}$ be the union of all data points, $\mathbf{Z}$ the table of corresponding class memberships, $k = 1, \ldots, K$ and $i = 1, \ldots, N$:

$$\mathbf{Z}_{ki} = \begin{cases} 1, & \text{if } \mathbf{x}_i \in \mathbf{X}^{(k)} \\ 0, & \text{otherwise.} \end{cases}$$

The intuition is that all classification relevant information is represented by the classification table. Therefore, this information is retained in those input components of $\mathbf{X}$ that originate from a common virtual source with the classification table. As shown in [7], [8], [9] and [10], this constrained CCA problem is equivalent to Fisher's linear discriminant analysis.

3.2. A NEW VIEW ON MIA

The CCA problem can be modified to extract an invariant signal from inputs of a single class. One interpretation of CCA is from the point of view of the cosine angle between the (non mean subtracted) vectors $\mathbf{a}^T \cdot \mathbf{X}$ and $\mathbf{Z}^T \cdot \mathbf{b}$. We will use a modified CCA (MCCA) criterion as follows: First, consider the original inputs rather than the mean subtracted covariance matrices; Second, $\mathbf{Z}^T \cdot \mathbf{b} = \underline{\mathbf{1}}$, the class membership table for data from a single class. It can be seen that criterion (3) becomes independent of $\mathbf{b}$. It represents a solution $\hat{\mathbf{a}}_{\mathrm{MCCA}}$:

$$\hat{\mathbf{a}}_{\mathrm{MCCA}} = \underset{\mathbf{a}, \mathbf{a}=\mathbf{X}^{(p)} \cdot \mathbf{c}}{\arg\max} \frac{\mathbf{a}^T \cdot \mathbf{X}^{(p)} \cdot \underline{\mathbf{1}}}{\sqrt{\mathbf{a}^T \cdot \mathbf{X}^{(p)} \cdot \mathbf{X}^{(p)^T} \cdot \mathbf{a}}} \quad (4)$$

This criterion is maximized when the correlation of $\mathbf{a}$ with all inputs $\mathbf{x}_i^{(p)}$ is as uniform as possible. The solution to this problem can be shown to be:

$$\mathbf{a} = \alpha \mathbf{X}^{(p)} \cdot \left(\mathbf{X}^{(p)^T} \cdot \mathbf{X}^{(p)} \right)^{-1} \cdot \underline{\mathbf{1}} \quad (5)$$

with $\alpha = \frac{\mathbf{a}^T \cdot \mathbf{X}^{(p)} \cdot \mathbf{X}^{(p)^T} \cdot \mathbf{a}}{\mathbf{a}^T \cdot \mathbf{X}^{(p)} \cdot \underline{\mathbf{1}}}$. Note that α is a scalar that results in scale independent solutions. As can easily be seen, the solution (5) of the modified CCA problem in (4) is identical to the MIA solution in (2). Thus, one can easily argue for the equivalence of the MCCA and MIA criteria.

The new formulation (4) of MIA highlights its properties:

Corollary 3.1. *The MIA problem has no solution if the inputs are zero mean i.e., if $\mathbf{X}^{(p)} \cdot \underline{\mathbf{1}} = \underline{\mathbf{0}}$.*

This is obvious from (4).

Corollary 3.2. *Any combination $\hat{\mathbf{a}}_{\mathrm{MCCA}} + \mathbf{b}$ with $\mathbf{b}$ in the nullspace of $\mathbf{X}^{(p)}$ is also a solution to (4).*

This means that only the component of $\mathbf{a}$ that is in the span of $\mathbf{X}^{(p)}$ contributes to the criterion in (4).

Corollary 3.3. *The solution of (4) is not unique if the N inputs $\mathbf{X}^{(p)}$ do not span the D-dimensional space.*

This follows from corollary 3.2. A unique solution can be found by further constraining (4). One such constraint is that $\mathbf{a}$ be a linear combination of the inputs $\mathbf{X}^{(p)}$.

Corollary 3.4. *The MIA solution reduces to the mean of the inputs in the special case when the covariance of the data $\mathbf{C_{XX}}$ has equal eigenvalues.*

Indeed, (4) can be rewritten as:

$$\hat{\mathbf{a}}_{\mathrm{MCCA}} = \underset{\mathbf{a}, \mathbf{a} = \mathbf{X}^{(p)} \cdot \mathbf{c}}{\arg \max} \frac{\mathbf{a}^T \cdot \mu^{(p)}}{\sqrt{\mathbf{a}^T \cdot \mathbf{C}_{\mathbf{XX}}^{(p)} \cdot \mathbf{a} + \left(\mathbf{a}^T \cdot \mu^{(p)}\right)^2}} \quad (6)$$

After normalizing $\mathbf{a} = \frac{\mathbf{X}^{(p)} \cdot \mathbf{c}}{\|\mathbf{X}^{(p)} \cdot \mathbf{c}\|}$ and using the spectral decomposition theorem [11], it can be shown that $\mathbf{a}^T \cdot \mathbf{C}_{\mathbf{XX}}^{(p)} \cdot \mathbf{a}$ is invariant to $\mathbf{a}$. The function in (6) is monotonically increasing in $\mathbf{a}^T \cdot \mu^{(p)}$. Therefore, the optimum of (6) is obtained when $\frac{\mathbf{a}^T \cdot \mu^{(p)}}{\|\mathbf{a}\|}$ is maximum. This means $\hat{\mathbf{a}}_{\mathrm{MCCA}} = \mu^{(p)}$.

4. A BAYESIAN MIA FRAMEWORK

In this section we motivate and analyze MIA from a Bayesian point of view. This allows us to find a generalized MIA formulation that can incorporate uncertainties and other prior knowledge. Furthermore, we show which assumptions distinguish MIA from linear regression.

The general linear regression model is $\mathbf{y} = \mathbf{X} \cdot \beta + \mathbf{n}$. As discussed in [12], the expected value $E\{\beta | \mathbf{y}\}$ from the conditional probability $p(\beta | \mathbf{y})$ can be introduced as a biased estimator of β. Let $\mathbf{n} \sim N(\mathbf{0}, \mathbf{C_n})$ and $\beta \sim N(\mu_\beta, \mathbf{C}_\beta)$ be independent Gaussian variables. Also, let $p(\mathbf{y}) = N(\mu_\mathbf{y}, \mathbf{C_y})$ and $p(\mathbf{y}, \beta) = N\left(\begin{bmatrix} \mu_\mathbf{y} \\ \mu_\beta \end{bmatrix}, \begin{bmatrix} \mathbf{C_y} & \mathbf{C}_{\mathbf{y}\beta} \\ \mathbf{C}_{\beta\mathbf{y}} & \mathbf{C}_\beta \end{bmatrix}\right)$. After a few mathematical transformations (see [12]), the posterior expectation of β given $\mathbf{y}$ is found to become:

$$\begin{aligned} E\{\beta | \mathbf{y}\} &= \\ &= \mu_\beta + \left(\mathbf{X}^T \cdot \mathbf{C_n}^{-1} \cdot \mathbf{X} + \mathbf{C}_\beta^{-1}\right)^{-1} \cdot \mathbf{X}^T \cdot \mathbf{C_n}^{-1} \cdot (\mathbf{y} - \mathbf{X} \cdot \mu_\beta) \end{aligned} \quad (7)$$

Ridge regression follows from this result by further assuming $\mu_\beta = \mathbf{0}$, $\mathbf{C}_\beta = \sigma_\beta^2 \mathbf{I}$ and $\mathbf{C_n} = \sigma_\mathbf{n}^2 \mathbf{I}$:

$$\beta_{\mathrm{RIDGE}} = \left(\mathbf{X}^T \cdot \mathbf{X} + \frac{\sigma_\mathbf{n}^2}{\sigma_\beta^2} \mathbf{I}\right)^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y}$$

Ridge regression is a generalization of the least squares solution to the regression problem and helps when $\mathbf{X}^T \cdot \mathbf{X}$ is not full rank or where we have numerical instability. Ridge regression assumes availability of the desired output $\mathbf{y}$ to aid the estimation of a weighting vector β. Thereafter, β is used to predict future outcomes of $\mathbf{y}$. In contrast to this, MIA aims to extract $\mathbf{y} \equiv \mathbf{w}$ such that the expectations of $\mathbf{y}$'s correlation with the high dimensional inputs are equal $E\{\mathbf{X}^T \cdot \mathbf{y}\} = \zeta \mathbf{1}$. By left-multiplying the linear model with $\mathbf{X}^T$ and substituting $\mathbf{X}^T \cdot \mathbf{y}$ with $\mathbf{g}$, $\mathbf{X} \cdot \beta$ with $\mathbf{w}$ and $\mathbf{X}^T \cdot \mathbf{n}$ with $\mathbf{f}$, the model becomes:

$$\mathbf{g} = \mathbf{X}^T \cdot \mathbf{w} + \mathbf{f} \quad (8)$$

In the following, let us assume that $\mathbf{w} \sim N(\mu_w, \mathbf{C_w})$ and $\mathbf{f} \sim N(\mathbf{0}, \mathbf{C_f})$ are independent Gaussian variables and $\mathbf{g} \sim N(\zeta \mathbf{1}, \mathbf{C_g})$. Note, the independence of $\mathbf{w}$ and $\mathbf{f}$ is implied by the previous assumptions on β and $\mathbf{n}$. We can find a generalized MIA criterion (GMIA) by using an equivalent derivation to (7) and replacing $\mathbf{g}$ with its expectation:

$$\begin{aligned} \mathbf{w}_{\mathrm{GMIA}} &= \\ &= \mu_\mathbf{w} + \mathbf{C_w} \cdot \mathbf{X} \cdot \left(\mathbf{X}^T \cdot \mathbf{C_w} \cdot \mathbf{X} + \mathbf{C_f}\right)^{-1} \cdot \left(\zeta \mathbf{1} - \mathbf{X}^T \cdot \mu_\mathbf{w}\right) \\ &= \mu_\mathbf{w} + \left(\mathbf{X} \cdot \mathbf{C_f}^{-1} \cdot \mathbf{X}^T + \mathbf{C_w}^{-1}\right)^{-1} \cdot \mathbf{X} \cdot \mathbf{C_f}^{-1} \cdot \left(\zeta \mathbf{1} - \mathbf{X}^T \cdot \mu_\mathbf{w}\right) \end{aligned} \quad (9)$$

Equation (9) suggest various properties of MIA. Note that the first line becomes identical to (2) if $\mathbf{C_w} = \mathbf{I}$, $\mu_\mathbf{w} = \mathbf{0}$ and $\mathbf{C_f} = \mathbf{0}$. In general it is desirable that the MIA representation be robust to small variations in $\mathbf{X}$ (e.g., due to noise). Equation (9) indicates when small variations in $\mathbf{X}$ do not have a large effect on the GMIA result. Furthermore, (9) allows us to integrate additional prior knowledge such as smoothness of $\mathbf{w}_{\mathrm{GMIA}}$ through the prior $\mathbf{C_w}$, correlation of consecutive instances $\mathbf{x}_i$ through the prior $\mathbf{C_f}$ etc. Moreover, we can use the GMIA formulation to define an iterative procedure that tackles datasets with large N and D. In this case it might be unfeasible to compute the matrix inverse. By using subsets of the input data, one can iteratively compute $\mu_\mathbf{w}$ as a MIA representation of the whole dataset from smaller subsets.

5. APPLICATIONS OF MIA

Possible MIA applications include novelty detection, classification, dimensionality reduction and feature extraction. MIA can be used when it is desirable to extract a single representation from a set of high dimensional data vectors. Such high dimensional data is common in the fields of audio and image processing, bioinformatics, spectroscopy etc. The usefulness of MIA was already shown on challenging real world applications such as illumination robust face recognition and text independent speaker verification [5], where performance is competitive with well studied approaches such as Gaussian mixture models, kernel PCA and FLDA.

The key to an effective application of MIA is the transfer of the input data to a space in which the assumed linear model in (8) holds. For example, as discussed in [5], in speaker verification this is the case for voiced speech in the logarithmic Fourier domain. Next, in contrast to previous work, we use different input feature and preprocessing. It can be argued that the new feature $\mathbf{x}_i^{(p)} = \log(\mathrm{abs}(\mathrm{IFFT}(\mathrm{DCF}(\mathrm{FFT}(\mathbf{s}_i^{(p)})))))$ also fits the MIA model for voiced speech $\mathbf{s}^{(p)}$. The NTIMIT database consists of 10 utterances per person that are classified in three different text types. In contrast to [5], the voiced segments are extracted for each utterance individually. We use every other utterance to learn a MIA/GMIA signature
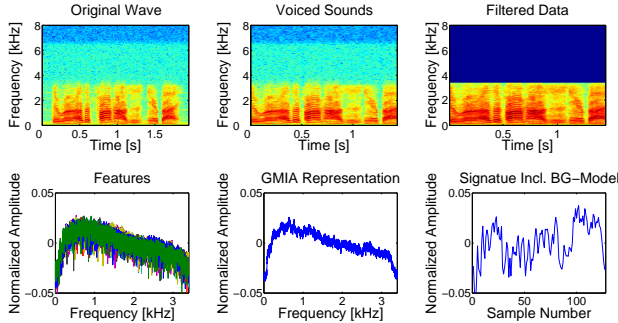
Fig. 1: Processing chain for text-independent speaker verification using GMIA

Table 1: MIA and GMIA performance comparison using various NTIMIT database segments.

Method	EER [%]	Identification [%]	Number of Speakers	Comments
GMIA	6.0	52	168	Only test section used
	6.9	39	438	Only male speakers used
	6.5	37	630	Full database used
MIA	6.9	48	168	Only test section used
	8.4	35	438	Only male speakers used
	7.5	32	630	Full database used
MIA [5]	6.8	56	168	50-50 partitioning for training and testing
GMM [13]	7.2	69	168	Similar/dissimilar speakers excluded
GMM [14]	9.6	N/A	168	Only test section used
GMM [15]	12.4	N/A	168	Only test section used
	8.8		630	Full database used
Phoneme GMM [16]	15.7	N/A	438	Only male speakers used

of each speaker in order to vary the text type composition. To achieve a high spectral resolution and for improved comparability to previous results, we choose a window size of one second. This induces high dimensional data vectors of $D = 6800$. Additionally, we use a background speaker model to improve our classification performance. For GMIA we assume $\mu_{\mathbf{w}} = \underline{0}$, $\mathbf{C_w} = \mathbf{I}$ and $\mathbf{C_f} = \Delta^T \cdot \Delta$ with $\Delta \in \mathbb{R}^{(N-2) \times N}$ being a tridiagonal Toeplitz matrix with backward shift elements $[1, -2, 1]$. The procedure that has been used to extract GMIA speaker signatures is illustrated in Fig. 1.

The new speaker verification EER results of MIA and GMIA are presented in Table 1 lines 1 and 2, column 2. These are competitive with other results from the literature including state-of-the-art Gaussian mixture model (GMM) approaches on challenging noisy data (see lines 3-7 in Table 1). The overall EER drops lower if the database includes both male and female speakers. For comparability with previous results in the literature we also include the identification rates of the algorithms in column 3. The identification rate of our MIA and GMIA implementation is below the state-of-the-art GMM approaches ([13] reported 60.7 % identification rate on the full 630 speakers of the NTIMIT database). We expect that this gap between the identification performances can be reduced by an improvement of the currently used background model. Moreover, the performance can be further optimized as a function of the window size, feature length, design etc.

6. CONCLUSION

The goal of MIA is to compute a unique invariant or characteristic of a dataset that can be used in class recognition problems. By definition, the MIA invariant is a linear combination of class examples that has equal correlation with all the training samples in the class. Another equivalent view is to find a direction to project the dataset such that projection lengths are maximally correlated. Both goals have MIA as unique solution. This paper discusses similarities of MIA with regression, Bayesian estimation and a modified CCA approach. Each brings insights into the value and properties of MIA. Furthermore, a general MIA solution (GMIA) is found. The performance of both MIA and GMIA is shown to be competitive to other modern algorithms. Future work will investigate statistical properties of MIA for a large number of inputs and computational tractability in large dimensions.

7. REFERENCES

[1] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[2] H. Hotelling, "Relation between two sets of variates," *Biometrika*, vol. 28, pp. 322–377, 1936.

[3] A. Tikhonov, "On the stability of inverse problems," *Doklady Akademii Nauk SSSR*, vol. 39, no. 5, pp. 195–198, 1943.

[4] H. Claussen, J. Rosca, and R. Damper, "Mutual interdependence analysis," in *Independent Component Analysis and Blind Signal Separation*, Heidelberg, Germany, 2007, pp. 446–453, Springer-Verlag.

[5] H. Claussen, J. Rosca, and R. Damper, "Mutual features for robust identification and verification," in *International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, NV, 2008, pp. 1849–1852.

[6] W. M. Fisher, G. R. Doddington, K. M. Goudie-Marshall, C. Jankowski, A. Kalyanswamy, S. Basson, and J. Spitz, "NTIMIT," CDROM, 1993.

[7] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*, Academic Press, Padstow, Cornwall, UK, 1979.

[8] M. S. Bartlett, "Further aspects of the theory of multiple regression," *Proceedings of the Cambridge Philosophical Society*, vol. 34, pp. 33–40, 1938.

[9] T. Hastie, R. Tibshirani, and A. Buja, "Flexible discriminant analysis by optimal scoring," *Journal of the American Statistical Association*, vol. 89, no. 428, pp. 1255–1270, 1994.

[10] F. R. Bach and M. I. Jordan, "A probabilistic interpretation of canonical correlation analysis," Tech. Rep. 688, Department of Statistics, University of California, Berkeley, 2005.

[11] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice-Hall, Upper Saddle River, NJ, 2000.

[12] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*, Prentice-Hall, Inc., Upper Saddle River, NJ, 1993.

[13] D. A. Reynolds, "Speaker identification and verification using Gaussian mixture speaker models," *Speech Communication*, vol. 17, no. 1-2, pp. 91–108, 1995.

[14] C. Sanderson, "Speech processing & text-independent automatic person verification," Tech. Rep. 08, IDIAP, Martigny, Switzerland, 2002.

[15] B. Wildermoth and K.K. Paliwal, "GMM based speaker recognition on readily available databases," in *Microelectronic Engineering Research Conference*, Brisbane, Australia, 2003, pagination unknown.

[16] D. Gutman and Y. Bistritz, "Speaker verification using phoneme-adapted Gaussian mixture models," in *European Signal Processing Conference*, Toulouse, France, 2002, vol. 3, pp. 85–88.