EFFICIENT TWO STAGE DECODING SCHEME TO ACHIEVE CONTENT IDENTIFICATION CAPACITY

Farzad Farhadzadeh, Ke Sun and Sohrab Fredowsi

Computer Science Department, University of Geneva

ABSTRACT

We introduce a scheme to address the trade–off between the identification rate, search and memory complexities in large–scale identification systems. We use a special database organization by assigning database entries to a set of possibly overlapping clusters. The clusters are generated based on statistics of both database entries and queries. The decoding procedure is accomplished in two stages. First, a list of clusters related to the query is detected. Then, refinement checks are performed on members of the detected clusters to produce a unique index. We investigate the minimum achievable search complexity for binary symmetric sources.

Index Terms— Content identification, identification capacity, clustering.

1. INTRODUCTION

Identification by nearest neighbor search is a research problem that emerge in vast signal processing tasks including, but not limited to, human biometrics [1], multimedia security (copy detection, content identification and tracking [2]) and physical object security [2].

An identification system consists of two main phases. In the *enrollment* phase, feature vectors representing digital contents, humans, multimedia contents or physical objects are extracted and stored in a database. In the *identification* phase, a query, i.e., a noisy and degraded counterpart of an enrolled data, is presented for identification, which is accomplished by comparing to feature vectors stored in the database.

Willems et al. [1] investigated the capacity of an identification system C, which is defined as the maximum achievable exponential rate of the number of distinguishable objects in a database. They showed that approximately 2^{NR} items can be distinguished from each other, if and only if $R < C$ and N, the dimensionality of the feature space, becomes very large. They showed that the identification capacity C is equal to the mutual information between the enrollment and identification observations, see also [3].

To achieve such a capacity, an identification system can perform an exhaustive search on all entries of the database to find the best match. However, it is not feasible in modern applications where the size of a database can be billions. The conventional approach to reduce the search complexity is to use a multidimensional index structure. Space-partitioning methods like quadtree [4] divide the data space along predefined planes. Data-partitioning index trees like R-tree [5] divide the data space according to the distribution of data. Although these access methods generally work well for low-dimensional spaces, their performance degrades dramatically as the number of dimensions increases – a phenomenon known as the *curse of dimensionality* [6].

Weber et al. [7] compared indexing techniques to methods based on vector-approximations (VA). Weber et al. [7] showed that for searching in high-dimensional spaces, quantization methods like VA outperform indexing methods. In an information-theoretical context, they can be referred to as quantization methods.

This paper is a continuation of the work initiated by Willems [8] and generalized by Farhadzadeh et al. [9]. The main goal of this sequel is to introduce a search strategy based on a two–stage decoding scheme, so as to achieve the identification capacity and to reduce the search complexity. Section 2.1 presents the two–stage identification system. Section 3 investigates minimum search complexity of a binary symmetric system. Section 4 shows the simulation results to validate the theoretical findings. Concluding remarks follow in Section 5.

2. MODEL DESCRIPTION

2.1. Model Description

The usual identification setup (Fig. 1) consists of two stages: *enrollment* and *identification*. In the enrollment phase, a randomly generated sequence (vector) $\mathbf{x}(w)$ of length N is extracted from each item w and stored in a database. $\mathbf{x} = (x_1, \ldots, x_N)$ has symbols $x_n, 1 \leq n \leq N$ taking values in a discrete alphabet $\mathcal{X}$. The database C is a collection of M indexed sequences denoted by

$$C = \{\mathbf{x}(1), \ldots, \mathbf{x}(M)\}. \tag{1}$$

Let's assume the components of $\mathbf{X} = (X_1, X_2, \cdots, X_N)$ are independent and identically distributed (i.i.d.) according to $\{Q_s(x); x \in \mathcal{X}\}$. Hence, the probability that sequence $\mathbf{x}$ oc-

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

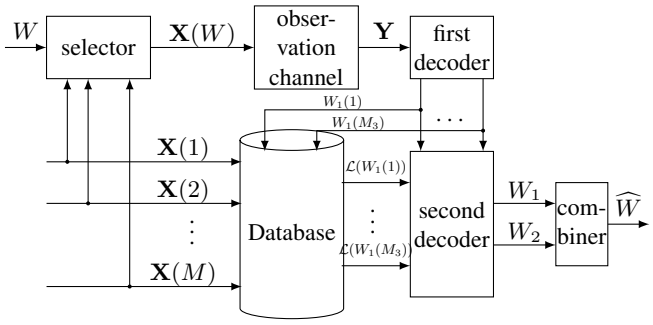

Fig. 1: A two-stage identification system.

curs for the item indexed by w is

$$\Pr\{\mathbf{X}(w) = \mathbf{x}\} = \Pi_{n=1}^{N} Q_s(x_n). \tag{2}$$

Note that this probability does not depend on the index w. All sequences in C are generated prior to the identification phase.

To organize the list of enrolled items, we use a set of cluster centroids $C_u = \{\mathbf{u}(1), \ldots, \mathbf{u}(M_1)\}$, where $\mathbf{u} = (u_1, \ldots, u_N)$ and $u_n, 1 \leq n \leq N$ take values in a discrete alphabet $\mathcal{U}$. Consider the cluster centroids are generated according to

$$P(u) = \sum_{x,y} Q_s(x) Q_c(y \,|\, x) P_{U|XY}(u \,|\, x, y) \tag{3}$$

for a given $P_{U|XY}(u \,|\, x, y)$, where $y \in \mathcal{Y}$ is the output of a memoryless observation channel $\{Q_c(y \,|\, x); x \in \mathcal{X}, y \in \mathcal{Y}\}$. Equation (3) indicates that the cluster centroids are related to both the items $\mathbf{x}$ and the observation channel output $\mathbf{y} = (y_1, \ldots, y_N)$. For each $\mathbf{u}(w_1) \in C_u$, we construct a list of enrolled items $\mathcal{L}(w_1) = \{w : d(\mathbf{x}(w), \mathbf{u}(w_1)) \leq \delta N\}$ having cardinality M_2. $d(\cdot, \cdot)$ indicates a similarity metric and $\delta \geq 0$.

In the identification phase, an enrolled item will be presented for identification. The probability of each item w to be presented for identification are all equal, and

$$\Pr\{W = w\} = 1/M, \ \forall\, w \in \{1, 2, \cdots, M\}. \tag{4}$$

When item w is presented for identification, its corresponding sequence $\mathbf{x}(w)$, which is "selected" from the database C, is observed through the observation channel. The resulting channel output sequence is $\mathbf{y}$, and

$$\Pr\{\mathbf{Y} = \mathbf{y} \,|\, \mathbf{X}(w) = \mathbf{x}\} = \Pi_{n=1}^{N} Q_c(y_n \,|\, x_n). \tag{5}$$

After observing $\mathbf{y}$, identification starts by constructing a list of cluster indices $\mathbf{w}_1$ with cardinality M_3. This index list $\mathbf{w}_1 = \big(w_1(1), \ldots, w_1(M_3)\big)$, $w_1(i) \in \{1, 2, \cdots, M_1\}, 1 \leq i \leq M_3$, is constructed by a so-called "first decoder" dec_1, a device having no knowledge of the entries $\mathbf{x}$ that were generated. Hence

$$\mathbf{w}_1 = \mathrm{dec}_1(\mathbf{y}). \tag{6}$$

Then, at the second decoding stage, a refinement decision is made, based on the list of cluster indices $\mathbf{w}_1$ and their members. This decision consisting of $w_1 \in \{1, 2, \cdots, M_1\}$ and $w_2 \in \{1, 2, \cdots, M_2\}$ is taken by a so-called "second decoder" dec_2. Hence

$$(w_1, w_2) = \mathrm{dec}_2(\mathbf{y}, \mathbf{w}_1, C). \tag{7}$$

Finally, a combiner com based on the estimated cluster index w_1 and the index w_2 forms an estimate of the index of the presented item for identification. Hence

$$\widehat{w} = \mathrm{com}(w_1, w_2). \tag{8}$$

We assume that $\widehat{w} \in \{1, 2, \cdots, M\}$.

The reliability of our identification system is measured by the error probability

$$P_{\mathcal{E}} = \Pr\left\{\widehat{W} \neq W\right\}. \tag{9}$$

2.2. Statement of Result

The rate quadruple (R_1, R_2, R_3, R) with $R \geq 0$ is called *achievable*, if for $\epsilon > 0$ and N large enough, there exist mappings $\mathrm{dec}_1(\cdot)$, $\mathrm{dec}_2(\cdot, \cdot, \cdot)$, and $c(\cdot, \cdot)$, such that

$$\begin{aligned} \log_2(M_1) &\leq N(R_1 + \epsilon), \\ \log_2(M_2) &\leq N(R_2 + \epsilon), \\ \log_2(M_3) &\leq N(R_3 + \epsilon), \\ \log_2(M) &\geq N(R - \epsilon), \\ \text{and } \Pr\{\widehat{W} \neq W\} &\leq \epsilon. \end{aligned} \tag{10}$$

We call R identification rate, R_1 cluster rate, R_2 refinement rate, and R_3 detected cluster list rate. The following theorem states a fundamental trade-off between these rates to achieve the identification capacity.

Theorem 1. The region of achievable rate quadruples $\mathcal{R}$ for the identification system introduced above is given by

$$\begin{aligned} \big\{(R_1, R_2, R_3, R) : & R_1 \geq I(X, Y; U), \\ & R_2 \geq \max(0, R - I(X; U)), \\ & R_3 \geq I(X; U|Y), \\ & 0 \leq R \leq I(X; Y), \\ & \text{for } P(x, y, u) = Q_s(x) Q_c(y \,|\, x) P(u \,|\, x, y), \\ & \text{where } |\mathcal{U}| \leq |\mathcal{Y}| \cdot |\mathcal{X}| + 2\big\}, \end{aligned} \tag{11}$$

where $I(\cdot\,;\,\cdot)$ indicates the mutual information [10]. We refer the readers to [9] for the proof of the theorem consisting of the achievability, the converse, and the cardinality bound.

Following Theorem 1, the total memory–complexity exponent of the identification setup related to storage of the cluster–centroids and their corresponding items [9] is

$$M_e = \max\Big\{I(U; X, Y) + R - I(U; X),\ I(U; X, Y)\Big\}. \tag{12}$$

The total search–complexity exponent related to the two–stage decoding scheme corresponding to finding related clusters for a given query $\mathbf{y}$ (cluster check) and checking their members (refinement check) [9] is

$$S_e = \begin{cases} I(U;X,Y) & , R \le I(U;X) + I(U;Y), \\ R + I(U;X|Y) - I(U;X), & R \ge I(U;X) + I(U;Y). \end{cases} \tag{13}$$

Remark 1. Under Markov chain condition $X \leftrightarrow Y \leftrightarrow U$, the region of achievable rate $\mathcal{R}$ coincides with the results shown in [8]. In this case $R_3 = 0$, which means that the first decoder is a unique decoder, sending a single $\mathbf{u}$ to the second decoder. Moreover, $R_1 + R_2 \ge R$, which means that clusters can overlap, i.e., each item can belong to multiple clusters.

Remark 2. Under Markov chain condition $U \leftrightarrow X \leftrightarrow Y$, $R_1 + R_2 = R$, which means that the clusters are disjoint, i.e., each item belongs to at most one cluster. This condition is exploited in an information retrieval system [11], where the authors proposed a disjoint clustering based on k-means. Under this condition, the identification system can achieve identification capacity if the first decoder detects a list of clusters.

3. MINIMUM SEARCH COMPLEXITY

Consider a common case of the above identification system. Assume binary uniform sequences, thus $Q_s(x) = 1/2$ for $x \in \mathcal{X} = \{0,1\}$. Also, assume a binary symmetric observation channel, thus $Q_c(y\,|\,x) = q$ if $y \neq x$, where $y \in \mathcal{Y} = \{0,1\}$. Let $u \in \mathcal{U} = \{0,1\}$.

Example 1. We generate $10,000$ trials of the conditional distribution $P(u\,|\,x,y)$ at random. The four relevant probabilities $p_1 \triangleq P_{U|XY}(0\,|\,0,0)$, $p_2 \triangleq P_{U|XY}(0\,|\,0,1)$, $p_3 \triangleq P_{U|XY}(0\,|\,1,0)$ and $p_4 \triangleq P_{U|XY}(0\,|\,1,1)$ are uniformly chosen over $[0,1]$. Based on (12) and (13), Fig. 2 (red o' s) shows the search–memory complexity exponents of the binary biometric system with the rate $R = 1/2$ and $q = 0.1$. Fig. 2 (blue $\times$'s) and (black $\cdot$'s) show the exponent of the search–memory complexity under Markov chain $X \leftrightarrow Y \leftrightarrow U$ and $U \leftrightarrow X \leftrightarrow Y$ conditions, respectively. Clearly, the proposed scheme under the general condition can achieve a better search-complexity compared to those Markov conditions.

We evaluate the minimum search–complexity exponent evaluated by (13) and its corresponding conditions of the binary system.

Theorem 2. Let $Q_s(x) = 1/2$ for $x \in \{0,1\}$, and let $Q_c(y\,|\,x)$ be a BSC with the cross-over probability $0 \le q < 1/2$ and $u \in \{0,1\}$. The minimum search–complexity exponent $S_e^* = (1-q)(1 - H_2(p_1^*))$ in an identification system using two–stage decoding can be achieved if $P(y\,|\,u)$ and $P(x\,|\,u)$ are BSCs with the same cross-over probability $P_b = p_1^* \star q/2 = H_2^{-1}(1 - R/2)$, where $p_1^* = \left(H_2^{-1}(1-R/2) - q/2\right)/(1-q)$, $H_2(a) = -a\log_2(a) - (1-a)\log_2(1-a)$ for $0 \le a \le 1/2$ denotes the binary entropy function, $H_2^{-1} : [0,1] \to [0,1/2]$ is its inverse mapping, and $(a \star b) = a(1-b) + b(1-a)$.

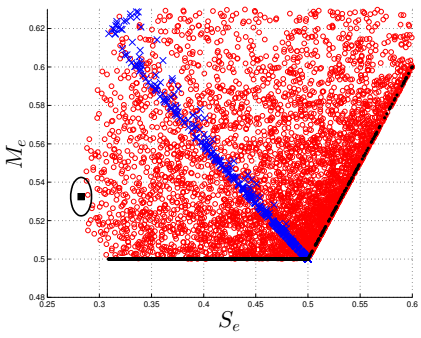

Fig. 2: (red o) indicate the exponent of the search–memory complexity of the proposed scheme using the conditional distribution $P(u|x,y)$. (blue $\times$) and (black $\cdot$) show the complexity using the conditional distribution $P(u|y)$ and $P(u|x)$ under the condition $X \leftrightarrow Y \leftrightarrow U$ and $U \leftrightarrow X \leftrightarrow Y$, respectively. (Circled ■) shows the minimum achievable search complexity.

Proof. We only give an outline proof. First, let's consider minimization of the search complexity $S_e = I(U;X,Y)$ subject to $g_1(\mathbf{p}) = R - I(U;X) - I(U;Y) \le 0$ in (13). This is a non-convex optimization. The point $\mathbf{p}^* = (p_1^*, p_2^*, p_3^*, p_4^*) = (p_1^*, 1/2, 1/2, 1-p_1^*)$ satisfies Karush–Kuhn–Tucker (KKT) conditions [12] concerning the Lagrangian function $\mathrm{L}(\mathbf{p},\mu) = S_e(\mathbf{p}) + \mu_1 g_1(\mathbf{p})$. Consequently, $\mathbf{p}^*$ is a local minimum. Then, the Second-Order-Sufficient-Condition [12] shows that $\mathbf{p}^*$ is a strict local minimum. To prove that $\mathbf{p}^*$ is a global minimum, we show as follows that $\mathbf{p}^*$ is the only point satisfying KKT conditions. If $\mu \neq 0$, $\nabla \mathrm{L}(\mathbf{p},\mu) = \mathbf{0}$ and $\mu g_1(\mathbf{p}) = 0$ leads to five linearly independent equations which has only one solution $\mathbf{p}^*$ and μ^*. On the other hand, if $\mu = 0$, $\nabla S_e = \mathbf{0}$ has solutions where $p_1 = p_2 = p_3 = p_4$. However, under this solution the primal feasibility condition $g_1(\mathbf{p}) = R < 0$ can not be satisfied since $R > 0$. Finally, following the fact that $\mu^* = (\log \frac{1-p_1^*}{p_1^*})/(2\log \frac{1-p_1^* \star q/2}{p_1^* \star q/2}), 0 \le p_1^* \le 1/2$ is a monotonically decreasing function [13], the Hessian of the Lagrangian function is positive-definite over the subspace $\mathcal{S} = \{(a_1,a_2,a_3,a_1) : a_1,a_2,a_3 \in \mathbb{R}\}$. Similarly, $\mathbf{p}^*$ minimizes the other non-convex minimization problem $S_e = R - I(U;X) + I(U;X|Y)$ subject to $g_1(\mathbf{p}) \ge 0$. Note, $g_1(\mathbf{p})$ is active for both problems, i.e., $g_1(\mathbf{p}) = 0$. Thus, both minimizations lead to the same solution. □

Remark 3. The optimal solution $\mathbf{p}^*$ is achieved when $I(U;X) = I(U;Y)$ and equivalently $I(U;X|Y) = I(U;Y|X)$.

Table 1: Performance and search complexity analysis of two–stage identification setup.

Distortion		$P_{\mathcal{E}}(\%)$	clustering						Search comp.				M_e	
model	Param.		k-medians			BBMM			k-medians		BBMM		k-medians	BBMM
			M_1	$M_2 \approx$	M_3	M_1	$M_2 \approx$	M_3	S_e	usage (%)	S_e	usage (%)		
AWGN (PSNR)	40 dB	0.019	180	115	30	220	473	5	0.19	16.73	0.17	10.68	0.22	0.26
	30 dB	0.024	180	115	70	220	568	6	0.21	38.85	0.18	14.95	0.22	0.26
	20 dB	0.389	180	115	125	220	946	10	0.22	69.33	0.2	35.14	0.22	0.27
JPEG (QF)	75	0.010	180	115	27	220	378	4	0.18	15.06	0.16	8.67	0.22	0.25
	50	0.014	180	115	30	220	568	5	0.19	16.73	0.17	12.64	0.22	0.26
	25	0.016	180	115	50	220	662	7	0.20	27.79	0.19	19.62	0.22	0.27
HistEQ		3.140	180	115	140	220	1236	13	0.22	87.32	0.21	50.91	0.22	0.28

In other words, the minimum search complexity in binary scheme can be realized when $M_3 = 2^{NI(U;X|Y)}$, the number of detected clusters by the first decoder, is equal to $2^{NI(U;Y|X)}$, the number of clusters contain every item.

Remark 4. The memory complexity exponent corresponding to the minimum search complexity exponent is $M_e = R/2 + (1-q)(1-H_2(p_1^*))$. Fig.2 (Circled ■) shows the minimum achievable search complexity exponent and its corresponding memory complexity exponent related to Example 1.

4. NUMERICAL EVALUATION

It should be noted that our theoretical findings accurately hold as N, the dimension of the feature space, becomes very large. However in practice, we usually deal with finite data of finite dimensionality. Consequently, real datasets follow some specific structure in a finite dimensional space. To capture this structure, we exploit two different binary clustering approaches, as usually fingerprints for identification are binary. One clustering approach is k-medians [14] that results in disjoint clusters. The other one is a Bayesian Bernoulli Mixture Model (BBMM) [6] that naturally has overlapped clusters.

Using BBMM clustering, the model which generates a binary database C in (1) is assumed to be

$$P\left(C,\, B \,|\, C_u\right) = \prod_{i=1}^{M}\prod_{j=1}^{M_1}\left[\frac{1}{M_1}\prod_{k=1}^{N} u_k(j)^{x_k(i)}(1-u_k(j))^{1-x_k(i)}\right]^{b_{ij}} \tag{14}$$

$$p(C_u) = \prod_{j=1}^{M_1}\prod_{k=1}^{N} \texttt{Beta}(u_{jk} \,|\, \alpha_{jk}, \beta_{jk}), \tag{15}$$

where $B_{M\times M_1} = (b_{ij})$ is a binary assignment matrix of size $M \times M_1$, $b_{ij} = 1$ specifies that the i'th item belongs to the j'th cluster, $x_k(i)$ denotes the k'th sample of the i'th item and $u_k(j)$ denote the k'th sample and the j'th cluster centroid. `Beta` indicates the beta distribution, and α_{jk} and β_{jk} are hyper-parameters (both are set to 1 here). In this scheme, all clusters are weighted equally, i.e. $1/M_1$. In this way, the cluster sizes are more balanced, which is closer to the theoretical results in Theorem 1. To fit the model to the data, we follow the Variational-Bayes approach with a mean-filed approximation [6], which alternatively updates until convergence $\tilde{B}$ and $\{\tilde{u}_{jk}\}$, the learning result of B and C_u, respectively. The benifit of such learning is two-fold. First, a codebook can be generated by binarizing $\tilde{B}$ so that each sample is assgined to M_3 clusters. Second, the most-likely cluster(s) which generate a given query y can be estimated using the learned $\{\tilde{u}_{jk}\}$.

We employ a real image database consisting of $20,828$ gray–scaled images from different categories of ImageNet (`http://www.image-net.org/`). All images are resized to 384×512 pixels. We extract binary fingerprints from each image as follows (see [2, 15]). First, each image is divided into 16×16 blocks. Then, the 2D DCT of each block is computed. The feature vector is constructed by concatenating the DCT coefficients at the coordinates $(1, 2)$ inside each block, resulting in a vector of length 768. Finally, the fingerprint of length $N = 64$ from each feature vector is extracted by using Random Projections [2] followed by a one-level scale quantizer. The identification rate is $R = (log_2 M)/N = 0.22$.

Table 1 shows the performance of the identification system based on the two–stage decoding with different M_1 and M_2, under various distortions such as Additive White Gaussian Noise (AWGN) with different Peak Signal to Noise Ratios (PSNR) PSNR$= 10\log_{10}(255^2/\sigma_Z^2)$, JPEG compression with different Quality Factors (QF) and Histogram Equalization (HistEQ). To highlight the search complexity reduction, we added the column usage(%) which indicate the percentage of the whole database used to identify a query. We keep the error probabilities $P_{\mathcal{E}}$ of k-medians and BBMM the same as the exhaustive search. Note, search and memory complexity exponents of exhaustive search are equivalent to R, i.e., $M_e = S_e = R$. And, the database usage in exhaustive search is 100%. Table 1 shows that clustering approaches like BBMM possessing overlapped clusters reduce search complexity considerably more than disjoint clustering like k-medians (see usage(%) columns).

5. CONCLUSIONS

We investigate the search and memory complexity of identification systems using a two–stage decoding strategy. We evaluate lower bounds to achieve the identification capacity on the necessary number of clusters, number of items in each cluster, and number of detected clusters at the first stage of decoding. We derive specific conditions to reach minimum search complexity for binary symmetric sources. We design a search scheme based on k-medians and Bayesian BMM to testify our theoretical findings. Numerical evaluation shows that using overlapped clustering approaches reduce search complexity more than disjoint clustering.

6. REFERENCES

[1] F. Willems, T. Kalker, J. Goseling, and J.-P. Linnartz, "On the capacity of a biometrical identification system," in *IEEE Int. Symp. Inform. Th.*, june-4 july 2003, p. 82.

[2] Farzad Farhadzadeh, Sviatoslav Voloshynovskiy, and Oleksiy Koval, "Performance analysis of content-based identification using constrained list-based decoding," *IEEE Trans. on Inf. Foren. and Sec.*, vol. 7, no. 5, pp. 1652–1667, 2012.

[3] J.A. O'Sullivan and N.A. Schmidt, "Large deviations performance analysis for biometrics recognition," in *Proc. 40th Ann. Allerton Conf. Comm. Control, and Comput.*, Oct. 2-4 2002.

[4] R.A. Finkel and J.L. Bentley, "Quad trees a data structure for retrieval on composite keys," *Acta Informatica*, vol. 4, no. 1, pp. 1–9, 1974.

[5] Antonin Guttman, "R-trees: a dynamic index structure for spatial searching," *SIGMOD Rec.*, vol. 14, no. 2, pp. 47–57, June 1984.

[6] Christopher M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[7] Roger Weber, Hans-Jörg Schek, and Stephen Blott, "A quantitative analysis and performance study for similarity-search methods in high-dimensional spaces," in *24th VLDB Conf.*, August 24-27 1998, pp. 194–205.

[8] F.M.J. Willems, "Searching methods for biometric identification systems: Fundamental limits," in *IEEE Int. Symp. Inform. Th.*, july 2009, pp. 2241 –2245.

[9] Farzad Farhadzadeh, Frans M.J. Willems, and Sviatoslav Voloshynovskiy, "Fundamental limits of identification: Identification rate, search and memory complexity trade–off," in *IEEE International Symposium on Information Theory (ISIT)*, Istanbul, Turkey, July 7–12 2013.

[10] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley-Interscience, New York, NY, USA, 1991.

[11] Hervé Jégou, Matthijs Douze, and Cordelia Schmid, "Improving bag-of-features for large scale image search," *Int. J. Comput. Vision*, vol. 87, no. 3, pp. 316–336, may 2010.

[12] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[13] E. Tuncel, "Capacity/storage tradeoff in high-dimensional identification systems," in *IEEE Int. Symp. Inform. Th.*, july 2006, pp. 1929 –1933.

[14] P. S. Bradley, O. L. Mangasarian, and W. N. Street, "Clustering via concave minimization," in *Advances in Neural Information Processing Systems -9*. 1997, pp. 368–374, MIT Press.

[15] Farzad Farhadzadeh, Sviatoslav Voloshynovskiy, Taras Holotyak, and Fokko Beekhof, "Active content fingerprinting: Shrinkage and lattice based modulations," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, 2013.