

# OPTIMAL CONFIGURATION OF HASH TABLE BASED MULTIMEDIA FINGERPRINT DATABASES USING WEAK BITS

Claus Bauer, Regunathan Radhakrishnan and Wenyu Jiang

Dolby Laboratories, San Francisco, CA, 94103, USA  
Email: cb@dolby.com, regu.r@dolby.com, wzj@dolby.com

## ABSTRACT

The increasingly large amount of digital multimedia content has created a need for technologies to search and identify multimedia files. Multimedia fingerprinting has been widely researched and successfully commercialized as a technology to trace and recognize both audio and video content.

Most published research on multimedia fingerprinting focuses on the design of the fingerprinting extraction algorithm. The architecture of the fingerprinting database and the database query search algorithm has only been investigated in very few publications. In particular, no comprehensive mathematical models describing the accuracy of database query algorithms have been developed. Therefore, most performance evaluations of fingerprint system rely on experimental data.

In this paper, we develop an analytic model to describe the operation of hash-table based multimedia fingerprinting databases. It allows to predict the accuracy of a fingerprinting query algorithm as a function of the statistical distribution of the fingerprints and the specific design of the fingerprint database.

**Keywords**— Multimedia fingerprints, database configuration, analytic model

## 1. INTRODUCTION

The explosive growth of digital multimedia content and its rapid distribution over the Internet have created a need for technologies to classify, trace, and identify multimedia files. Multimedia fingerprinting is considered as one of the most promising technologies that can provide these functionalities. For both audio [1] - [5] and video [6] - [10], substantial research on the design of multimedia fingerprinting systems has been performed and various companies have developed complete fingerprint systems [11] - [14].

Multimedia fingerprint extraction algorithms capture perceptual features of the multimedia object and store them in a compact bit sequence called the *fingerprint*. The accuracy of a fingerprint system is defined via the notions of *robustness* and *sensitivity*. Depending on the specific application of multimedia fingerprinting, the fingerprints are required to be robust against changes of the multimedia object from which they are derived. For an audio file, a modification could be a transcoding of the content. For a video file, a modification could be a rotation or

a cropping of the video. A modification of the multimedia object that does not prevent a human being from recognizing the modified and the unmodified multimedia object contain the same content should only lead to a relatively small change of the fingerprint. Fingerprints should be sensitive to content, i.e., fingerprints derived from different multimedia content must differ significantly. Indeed, fingerprint should not only be able to uniquely identify a multimedia file, but should moreover allow to identify a certain point in time in a multimedia file with high timely accuracy. This characteristic is required to identify parts of multimedia files illegally posted on Web 2.0 web pages like YouTube, MySpace, etc.

Most applications of fingerprints rely on the following fingerprint system architecture. Initially, a *reference* database of multimedia fingerprints is built. It contains the fingerprints of all multimedia objects the application shall recognize. During operation, a fingerprint of a query multimedia object is compared against the reference database using a specific database query algorithm. The algorithm decides if the fingerprint of the query multimedia object corresponds to a fingerprint in the reference database. Fingerprint matching algorithms use techniques from statistics and fuzzy logic.

Large scale multimedia fingerprint applications require highly scalable database design and query techniques. Most fingerprinting databases use either a tree structure [15] or a hash-table based design [16]. For both types of databases, the database design requires the configuration of specific parameters determining the exact structure of the database. In this paper, we investigate a way to model and optimally design a hash-table based multimedia fingerprint database. In particular, we develop a model to predict the performance of a query algorithm searching through a hash-based database as a function of both the statistical distribution of the fingerprints and the actual values taken by the database design parameters. The architecture of the database and the choice of the query algorithm determine the accuracy of the fingerprint matching algorithm as well the storage requirements of the database and the speed of the database search. The model proposed in this paper helps the system designer to methodically evaluate the different design trade-offs. In sec. 2, we define some terminology used in this paper. The concept of hash-based databases using weak bits is explained in sec. 3. In sec. 4, we develop an analytic model of the performance of the database query algorithm. In sec. 5, we describe

how this model can be used to dynamically configure a fingerprint database.

## 2. TERMINOLOGY

A multimedia fingerprint  $h$  is derived from a multimedia object  $x$  using a hash function  $H(x)$ , i.e.,  $h = H(x)$ . The fingerprint  $h$  is a sequence of bits of length  $L$ . We divide the fingerprint into an index of length  $I < L$  and a remainder of length  $L - I$ . We define the following probabilities to analytically describe the notions of robustness and sensitivity:

- $P_f$  Probability of false positive: A query fingerprint which does not correspond to any fingerprint in the reference fingerprinting database is erroneously declared to correspond to a fingerprint in the database.
- $P_c$  Probability of correct identification: A query fingerprint which corresponds to a fingerprint in the fingerprint reference database is correctly declared to correspond to this fingerprint in the reference database.
- $P_m$  Probability of misclassification: A query fingerprint which corresponds to a fingerprint in the fingerprint reference database is erroneously declared to correspond to a different fingerprint in the reference database.
- $P_d$  Probability of detection: A query fingerprint which corresponds to a fingerprint in the fingerprint reference database is declared to correspond to a fingerprint in the reference database. Obviously,

$$P_d = P_c + P_m. \quad (1)$$

Further, we introduce the following terminology:

- Collision: Depending on the fingerprint extraction algorithm and the way the index is chosen as a part of the fingerprint, the fingerprints of different multimedia objects might be indexed by the same database index. This scenario, where indexes of fingerprints of different multimedia objects *collide*, is called a *collision*.
- In order to simplify the calculations in this paper, we assume that for each index the expected number of collisions is identical. We denote the average number of collisions per fingerprint as  $D$ .
- Weak bits: Fingerprints are typically designed such that the fingerprint of an original multimedia object and the fingerprint of a modified version of the multimedia object differ only by few bits. When extracting the fingerprint of the query multimedia object, each of the  $I$  bits of the index of the fingerprint of the query multimedia object is assigned a certain probability. Using techniques from [17], this probability describes the likelihood with which the respective bit would change if the query multimedia object would be modified. The bits which are assigned a high(low) probability of change are called weak(reliable) bits.

- $P_W$  Probability of correct weak bit prediction: Assuming that a query multimedia fingerprint corresponds to a fingerprint in the reference database, this probability denotes the probability that the reliable bits of the index of the fingerprint of the query multimedia object are identical with the corresponding bits of the index of the fingerprint of the reference multimedia object. In other words,  $P_W$  denotes the probability that the weak bit prediction is correct. Using common notation, we set  $P_{W^c} = 1 - P_W$ . For any event  $A$ , we denote by  $P_{A|W}$  ( $P_{A|W^c}$ ) the conditional probability of  $A$  given that the weak bit prediction is correct (wrong).
- $p$  :  $p$  defines the probability that an individual bit of the  $L - I$  bits of the remainder of the fingerprint changes when the multimedia object, from which the fingerprint is derived, is modified. The value of the variable  $p$  depends on the design of the fingerprint extraction algorithm and the considered signal modifications.

**Note:** In practice, the quantities  $p$  and  $P_W$  tend to be related. In this paper, we do not consider this possible correlation, but assume that both quantities can be independently measured and used as input parameters for the system model described in sect. 4.

## 3. MULTIMEDIA FINGERPRINT DATABASES USING WEAK BITS

We first summarize the concept of multimedia fingerprint databases using weak bits described in [17]. When a query multimedia object is queried against the database, the fingerprint of the multimedia object is extracted. According to methods as described in [17], the  $I$  bits of the query fingerprint that constitute the index are divided into  $W \leq I$  weak and  $I - W$  reliable bits. We then consider all indexes that have the same  $I - W$  reliable bits as the index of the fingerprint of the query multimedia object. The number of these index equals  $2^W$  as we consider all possible permutations of the  $W$  unreliable bits. We recall the assumption in sec. 2 that the average number of collisions for each of the  $2^W$  indexes is  $D$ . Thus, the  $2^W$  indices index

$$N = D2^W. \quad (2)$$

fingerprints. The fingerprint matching algorithm calculates the  $N$  Hamming distances between the fingerprint of the multimedia query object and each of the  $N$  indexed fingerprints. Subsequently, the fingerprint among the  $N$  fingerprints with the lowest Hamming distance to the fingerprint of the multimedia query object is chosen. If the lowest Hamming distance is achieved by more than one of the  $S$  fingerprints, then one of these fingerprints is picked randomly. If this Hamming distance is not bigger than a pre-defined threshold  $\tau$ , then the corresponding indexed fingerprint is declared to present the query multimedia object. If this Hamming distance is bigger than a pre-defined threshold  $\tau$ , then the query multimedia object is declared not to

correspond to any multimedia object the fingerprint of which is included in the reference database.

#### 4. ANALYTIC MODEL OF THE DATABASE QUERY ALGORITHM

In sec. 4.1, we describe a framework to analyze fingerprint systems from [18]. In sec. 4.2 we derive an analytic model model of a multimedia fingerprint database using weak bits.

##### 4.1. Decision theoretic framework for analyzing binary hash-based content identification systems

In [18], a general model to analyze the performance of binary hash-based content identification systems is considered. The reference database consists of  $U$  copyrighted multimedia objects. However, as outlined in sec. 3, for each database query we only consider  $N \leq U$  multimedia objects  $V_1, \dots, V_N$  with fingerprints  $X_1, \dots, X_N$ . In order to simplify our presentation, we assume that no two fingerprints  $X_i$  are identical. When a user uploads a query multimedia object  $Z$  the fingerprint  $Y$  of  $Z$  is computed. The fingerprint matching algorithm decides if the multimedia object  $Z$  corresponds to any of the  $N$  multimedia objects  $V_1, \dots, V_N$ . The input to the fingerprint matching algorithm are the  $N$  Hamming distances  $d_i = d(V_i, Y)$ ,  $i = 1, \dots, N$  which respectively denote the  $N$  Hamming distances between the fingerprints of the query multimedia object  $Y$  and the  $N$  multimedia objects  $V_1, \dots, V_N$ . The database query algorithm uses a threshold parameter  $\tau$  as follows:

- If  $i = \arg_{j=1,2,\dots,N} \min d_j$  and  $d_i \leq \tau$ , the algorithm decides that the multimedia object  $Y$  corresponds to the multimedia object  $V_i$ . If the minimum value of all  $N$  Hamming distances  $d_i$  is attained by more than one index  $i$ , one of them is chosen randomly.
- Else, the algorithm decides that the multimedia object  $Y$  does not correspond to any of the  $N$  multimedia objects  $V_N$ .

In order to model the database query algorithm as a multiple hypothesis test, we define the following hypotheses:

- $H_0$ :  $Z$  is not from the database  $V_1, \dots, V_N$
- $H_i$ ,  $1 \leq i \leq N$  :  $Z$  corresponds to the multimedia object  $V_i$

As described in sec. 3, the  $L$  bits of the fingerprint consist of the  $I$  index bits and the  $L - I$  remainder bits. Thus,  $d_i = d_{I,i} + d_{R,i}$ ,  $1 \leq i \leq N$ , where

- $d_{I,i}$ : Hamming distance between the index of the fingerprint of multimedia object  $V_i$  and the index of the fingerprint of the query multimedia object  $Z$
- $d_{R,i}$ : Hamming distance between the remainder of the fingerprint of multimedia object  $V_i$  and the remainder of the fingerprint of the query multimedia object  $Z$

We define  $A_{min}$  as the number of  $j$  for which  $d_j = \min_{i=1,\dots,N} d_i$ . Further, we will make use of the parameter  $B_j$ ,  $0 \leq j \leq I$ , which denotes the number of  $d_i$  satisfying  $d_{I,i} = j$ , i.e.,  $B_j = \text{number}\{i | 1 \leq i \leq N, d_{I,i} = j\}$ . According to [18], [19] signal modifications lead to changes in the fingerprint space which can be modeled as additive noise. Here, the effect of signal modifications on fingerprint bits is modeled by the probability  $p$  defined in sec. 2.

##### 4.2. Calculation of $P_f, P_c, P_m$ , and $P_d$

We first introduce some auxiliary notation similar to the one used in [18]. We abbreviate the multinomial coefficients as  $C_M^{a_1 \dots a_h} := \binom{N}{a_1 \dots a_h}$ , where  $a_1 + \dots + a_h = M$ . Further, we define:  $f_0(k) = 2^{-L+I} C_{L-I}^k$ ,  $f_1(k) = C_{L-I}^k p^k (1-p)^{L-k}$ ,  $F_u(k) = \sum_{j=k}^L f_u(j)$  for  $0 \leq k \leq L - I$ ,  $u = 1, 2$ .

###### 4.2.1. Calculation of $P_f$

$$\begin{aligned} P_f &= P(\text{at least one of the } d_1, d_2, \dots, d_N \leq \tau | H_0) \\ &= 1 - P(\text{none one of the } d_1, d_2, \dots, d_N \leq \tau | H_0) \\ &= 1 - \prod_{m=1}^N G_0(\tau + 1 - d_{I,m}) \\ &= 1 - \prod_{u=0}^W G_0^{DC_u^w}(\tau + 1 - u), \end{aligned}$$

where  $G_0(\cdot)$  is defined as follows: We first note that under the hypothesis  $H_s$  for any  $i \in \{1, \dots, N\}$ ,  $i \neq s$  satisfying  $d_{I,i} \leq j + 1$  for a fixed  $j$ , we have  $P(d_i > j) = F_0(j + 1 - d_{I,i})$ , whereas for  $d_{I,i} > j + 1$ ,  $P(d_i > j) = 1$ . Thus

$$\begin{aligned} P(d_i > j) &= G_0(j + 1 - d_{I,i}) \\ &= \begin{pmatrix} F_0(j + 1 - d_{I,i}) & , \text{if } d_{I,i} \leq j + 1, \\ 1, & , \text{if } d_{I,i} > j + 1. \end{pmatrix}. \end{aligned} \quad (3)$$

###### 4.2.2. Calculation of $P_c$

We note that if the weak bit prediction is not correct, then the fingerprint in the reference database that corresponds to the fingerprint of the query multimedia object will not be identified. Thus,  $P_{c|W^c} = 0$  which implies  $P_c = P_{c|W} P_W$ . If the weak bit prediction is correct, then a specific hypothesis  $H_s$ ,  $1 \leq s \leq N$ , is correct, i.e., we need to calculate

$$P_c = P_{c|W} P_W = \sum_{s=1}^N P_{c|H_s} P_{H_s}. \quad (4)$$

The following calculations will show that the value of  $P_{c|H_s}$  depends indeed on the actual hypothesis  $H_s$  and in particular

depends on the quantity  $d_{I,s}$ . We first see

$$\begin{aligned} P_{c|H_s} &= P(\text{deciding } s|H_s) \\ &= P(d_s \leq \tau \wedge d_s < \min_{i \neq s} d_i | H_s) \\ &+ P(\min_{i \neq s} d_i = d_s \leq \tau \wedge s \text{ is decided} | H_s). \end{aligned} \quad (5)$$

We now calculate the first probability on the RHS of (5):

$$\begin{aligned} &P(d_s \leq \tau \wedge d_s < \min_{i \neq s} d_i | H_s) \\ &= \sum_{j=0}^{\tau} P(d_s = j \wedge j < \min_{i \neq s} d_i | H_s) \\ &= \sum_{j=0}^{\tau} P(d_{R,s} = j - d_{I,s} \wedge j < \min_{i \neq s} d_i | H_s). \end{aligned} \quad (6)$$

Under the condition  $H_s$ , the events  $d_{R,s} = j - d_{I,s}$  and  $j < \min_{i \neq s} d_i$  are independent, i.e.,

$$\begin{aligned} &P(d_{R,s} = j - d_{I,s} \wedge j < \min_{i \neq s} d_i | H_s \wedge) \\ &= P(d_{R,s} = j - d_{I,s} | H_s \wedge) P(j < \min_{i \neq s} d_i | H_s \wedge). \end{aligned} \quad (7)$$

It is easy to see that

$$P(d_{R,s} = j - d_{I,s} | H_s \wedge W) = g_1(j - d_{I,s}), \quad (8)$$

where

$$g_1(j - d_{I,s}) = \begin{cases} f_1(j - d_{I,s}), & \text{if } j \geq d_{I,s}, \\ 0, & \text{else.} \end{cases}$$

Using (3), we obtain

$$P(j < \min_{i \neq s} d_i | H_s) = \prod_{m=1, m \neq s}^N G_0(j + 1 - d_{I,m}). \quad (9)$$

We now calculate the second probability on the RHS of (5):

$$\begin{aligned} &P(\min_{i \neq s} d_i = d_s \leq \tau \wedge s \text{ is decided} | H_s) \\ &= \sum_{k=1}^{N-1} P(\min_{i \neq s} d_i = d_s \leq \tau \wedge A_{min} = k + 1 \wedge \\ &\quad s \text{ is decided} | H_s) \\ &= \sum_{k=1}^{N-1} \frac{1}{k+1} P(\min_{i \neq s} d_i = d_s \leq \tau \wedge A_{min} = k + 1 | H_s) \\ &= \sum_{k=1}^{N-1} \frac{1}{k+1} \sum_{j=0}^{\tau} P(d_{R,s} = j - d_{I,s} \wedge \min_{i \neq s} d_i = j \wedge \\ &\quad A_{min} = k + 1 | H_s). \end{aligned} \quad (10)$$

The last probability can be written as

$$\begin{aligned} &P(d_{R,s} = j - d_{I,s} \wedge \min_{i \neq s} d_i = j \leq \wedge \\ &\quad A_{min} = k + 1 | H_s) \\ &= \sum_{b_1=1}^N \dots \sum_{\substack{b_k=1 \\ b_1 < \dots < b_k, b_i \neq s \\ \#\{b_i | d_{I,b_i} = d_{I,s}\} \leq B_{d_{I,s}-1}}}^N \\ &P(d_{R,s} = j - d_{I,s} \wedge \min_{i \neq s} d_i = j \text{ attained} \\ &\quad \text{by } b_1, \dots, b_k | H_s). \end{aligned} \quad (11)$$

Arguing as in (7) - (9), we write the last probability in (11) as

$$\begin{aligned} &P(d_{R,s} = j - d_{I,s} \wedge \min_{i=1, \dots, N} d_i = j \text{ attained} \\ &\quad \text{by } s, b_1, \dots, b_k | H_s) = g_1(j - d_{I,s}) \prod_{i=1}^k g_0(j - d_{I,b_i}) \\ &\quad \times \prod_{m=1, m \neq s, b_1, \dots, b_k}^N G_0(j + 1 - d_m), \end{aligned} \quad (12)$$

where  $g_0(\cdot)$  is defined via  $f_0(\cdot)$  in the same way  $g_1(\cdot)$  is defined via  $f_1(\cdot)$  above. From(5) - (8) and (9) - (12), we get

$$\begin{aligned} P_{c|H_s} &= \sum_{j=0}^{\tau} g_1(j - d_{I,s}) \prod_{m=1, m \neq s}^N G_0(j + 1 - d_m) \\ &+ \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \frac{1}{k+1} \sum_{b_1=1}^N \dots \sum_{\substack{b_k=1 \\ b_1 < \dots < b_k, b_i \neq s \\ \#\{b_i | d_{I,b_i} = d_{I,s}\} \leq B_{d_{I,s}-1}}}^N \\ &\quad \times g_1(j - d_{I,s}) \prod_{i=1}^k g_0(j - d_{I,b_i}) \prod_{m=1, m \neq s, b_1, \dots, b_k}^N \\ &\quad \times G_0(j + 1 - d_m) \\ &= \sum_{j=0}^{\tau} g_1(j - d_{I,s}) G_0^{-1}(j + 1 - d_{I,s}) \prod_{u=0}^W \\ &\quad \times G_0^{DC^u}(j + 1 - u) \\ &+ \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \frac{1}{k+1} g_1(j - d_{I,s}) G_0^{-1}(j + 1 - d_{I,s}) \\ &\quad \times \sum_{a_0=0}^{C_W^0 - \delta_{d_{I,s},0}} \dots \sum_{\substack{a_W=0 \\ a_1 + \dots + a_W = k}}^{C_W^W - \delta_{d_{I,s},W}} C_W^{a_0, \dots, a_W} \\ &\quad \times \prod_{i=0}^W g_0^{a_i}(j - i) \prod_{u=0}^W G_0^{DC^i - a_i}(j + 1 - u), \end{aligned} \quad (13)$$

where  $\delta_{a,b}$  is the Kronecker delta. Inserting (13) into (4),

$$\begin{aligned}
P_{c|W} &= \sum_{v=0}^W \frac{C_W^v}{2^W} \sum_{j=0}^{\tau} g_1(j-v) G_0^{-1}(j+1-v) \\
&\times \prod_{u=0}^W G_0^{DC_W^u}(j+1-u) \\
&+ \sum_{v=0}^W \frac{C_W^v}{2^W} \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \frac{1}{k+1} g_1(j-v) G_0^{-1}(j+1-v) \\
&\sum_{a_0=0}^{C_W^0-\delta_{v,0}} \dots \sum_{\substack{a_W=0 \\ a_1+\dots+a_W=k}}^{C_W^W-\delta_{v,W}} C_W^{a_0,\dots,a_W} \prod_{i=0}^W g_0^{a_i}(j-i) \\
&\times \prod_{i=0}^W G_0^{DC_W^i-a_i}(j+1-i). \tag{14}
\end{aligned}$$

#### 4.2.3. Calculation of $P_m$

We note that

$$P_m = P_{m|W} P_W + P_{m|W^c} P_{W^c}. \tag{15}$$

We first calculate  $P_{m|W}$ . Similarly to the calculation  $P_{c|W}$  we first consider the conditional probability  $P_{m|H_s}$  for a fixed  $s$ .

$$\begin{aligned}
P_{m|H_s} &= P(\text{deciding } i \neq s \neq 0 | H_s) \\
&= P(\min_{i \neq s} d_i \leq \tau \wedge \min_{i \neq s} d_i < d_s | H_s) \\
&+ P(\min_{i \neq s} d_i = d_s \leq \tau \wedge i \neq s \text{ is decided} | H_s). \tag{16}
\end{aligned}$$

The first probability on the RHS of (16) is transformed as

$$\begin{aligned}
&P(\min_{i \neq s} d_i \leq \tau \wedge \min_{i \neq s} d_i < d_s | H_s) \\
&= \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} P(\min_{i \neq s} d_i = j \wedge \min_{i \neq s} j < d_s \wedge \\
&\quad A_{min} = k+1 | H_s) \\
&= \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \sum_{b_1=1}^N \dots \sum_{\substack{b_k=1 \\ \#\{b_i | d_{I,b_i}=d_{I,s}\} \leq B_{d_{I,s}}-1}}^N \\
&\quad P(d_{R,s} > j - d_{I,s} \wedge \min_{i=1,\dots,N} d_i = j \text{ attained} \\
&\quad \text{by } s, b_1, \dots, b_k | H_s). \tag{17}
\end{aligned}$$

Arguing as in (7) - (9), we write the last expression in (17) as

$$\begin{aligned}
&P(d_{R,s} > j - d_{I,s} \wedge \min_{i=1,\dots,N} d_i = j \text{ attained} \\
&\text{by } s, b_1, \dots, b_k | H_s) = G_1(j+1-d_{I,s}) \prod_{i=1}^k g_0(j-b_i) \\
&\times \prod_{m=1, m \neq s, b_1, \dots, b_k}^N G_0(j+1-d_m), \tag{18}
\end{aligned}$$

where  $G_1(\cdot)$  is defined via  $F_1(\cdot)$  in the same way  $G_1(\cdot)$  is defined via  $F_1(\cdot)$  in (3). The second probability in (16) is calculated as follows:

$$\begin{aligned}
&P(\min_{i \neq s} d_i = d_s \leq \tau \wedge i \neq s \text{ is decided} | H_s) \\
&= \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \frac{k}{k+1} P(\min_{i \neq s} d_i = d_s = j \wedge \\
&\quad A_{min} = k+1 | H_s). \tag{19}
\end{aligned}$$

This probability has been calculated in (10) - (12). Summarizing (16) - (19), we obtain

$$\begin{aligned}
P_{m|H_s} &= \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \left( G_1(j+1-d_{I,s}) + g_1(j-d_{I,s}) \frac{k}{k+1} \right) \\
&\sum_{b_1=1}^N \dots \sum_{\substack{b_k=1 \\ b_1 < \dots < b_k, b_i \neq s \\ \#\{b_i | d_{I,b_i}=d_{I,s}\} \leq B_{d_{I,s}}-1}}^N \prod_{i=1}^k g_0(j-I_{b_i}) \\
&\times \prod_{m=1, m \neq s, b_1, \dots, b_k}^N G_0(j+1-d_{I,m}) \\
&= \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \left( G_1(j+1-d_{I,s}) + g_1(j-d_{I,s}) \frac{k}{k+1} \right) \\
&\times G_0^{-1}(j+1-d_{I,s}) \sum_{a_0=0}^{C_W^0-\delta_{d_{I,s},0}} \dots \sum_{\substack{a_W=0 \\ a_1+\dots+a_W=k}}^{C_W^W-\delta_{d_{I,s},W}} \\
&\times C_W^{a_0,\dots,a_W} \prod_{i=0}^W g_0^{a_i}(j-i) \prod_{i=0}^W G_0^{DC_W^i-a_i}(j+1-i).
\end{aligned}$$

Arguing as in (14), we obtain

$$\begin{aligned}
P_{m|W} &= \sum_{v=0}^W \frac{C_W^v}{2^W} \sum_{j=0}^{\tau} \sum_{k=1}^{N-1} \left( G_1(j+1-v) + g_1(j-v) \frac{k}{k+1} \right) \\
&\times G_0^{-1}(j+1-v) \sum_{a_0=0}^{C_W^0-\delta_{v,0}} \dots \sum_{\substack{a_W=0 \\ a_1+\dots+a_W=k}}^{C_W^W-\delta_{v,W}} C_W^{a_0,\dots,a_W} \\
&\times \prod_{i=0}^W g_0^{a_i}(j-i) \prod_{i=0}^W G_0^{DC_W^i-a_i}(j+1-i). \tag{20}
\end{aligned}$$

Now, we calculate  $P_{m|W^c}$ . The condition  $W^c$  implies that the fingerprint of the query multimedia object does not correspond to any of the  $N$  fingerprints indexed as described in sec. 3. Thus, the hypothesis  $H_0$  holds. The conditional probability of the erroneous identification of the query fingerprint with any of the  $N$  fingerprints given that the weak bit prediction is wrong is therefore equal to the probability of a false positive, i.e.,  $P_{m|W^c} = P_f$ .

#### 4.2.4. Calculation of $P_d$

We note that  $P_d = P_{d|W}P_W + P_{d|W^c}P_{W^c}$ . To calculate  $P_{d|W}$ , we first consider  $P_{d|H_s}$ . We see

$$\begin{aligned} P(d|H_s) &= P(\min_{i=1,\dots,N} d_i \leq \tau | H_s) \\ &= 1 - \prod_{i=1}^N P(d_i \geq \tau + 1 | H_s) \\ &= 1 - G_1(\tau + 1 - d_{I,s}) \prod_{m=1, m \neq s}^N G_0(\tau + 1 - d_m). \end{aligned}$$

Averaging as in (14), we now derive  $P_{d|W}$  as follows:

$$\begin{aligned} P_{d|W} &= 1 - \sum_{v=0}^W \frac{C_W^v}{2^W} \sum_{j=0}^{\tau} G_1(j - v) G_0^{-1}(j + 1 - v) \\ &\quad \times \prod_{u=0}^W G_0^{DC_W^u}(j + 1 - u). \end{aligned} \quad (21)$$

Last, we conclude from (1) and  $P_{c|W^c} = 0$  and  $P_{m|W^c} = P_f$ ,  $P_{d|W^c} = P_{m|W^c} = P_f$ . Alternatively, using (1),  $P_d$  can be derived from the formulas for  $P_c$  and  $P_m$  given above.

## 5. DYNAMIC DATABASE CONFIGURATION

When the system is initialized, a maximum fingerprint length  $L_{max}$  is defined. Using maximum values of  $N$  and  $p$  and a minimum value of  $P_W$  envisioned for the target application, one uses the probability formula derived in sec. 4 and a series of values for  $L, I, \tau, W$  to find the smallest value of  $L = L_{max}$  for which one obtains the target values for  $P_f, P_c, P_m$ , and  $P_d$ . During operation, one similarly uses the current values of  $N, p, P_W$  to determine the actual values of  $L_{actual}, I_{actual}, \tau_{actual}, W_{actual}$  such that the resulting values of  $P_f, P_c, P_m$ , and  $P_d$  best match the target values for the application. For each reference multimedia object, a fingerprint of length  $L_{max}$  is extracted. The  $L_{max}$  bits are divided in two  $S_1$  and  $S_2$  containing  $L_{actual}$  and  $L_{max} - L_{actual}$  bits, respectively. For any database query, only the  $L_{actual}$  bits in  $S_1$  are considered. The sets  $S_1$  and  $S_2$  are updated whenever one of the input parameters  $N, p, P_W$  or one of the target probabilities changes.

## 6. REFERENCES

- [1] Cano, P. et al; *A Review of Algorithms for Audio Fingerprinting*, IEEE MMSP 2002.
- [2] J. Haitsma, T. Kalker, and J. Oostveen, *Robust audio hashing for content identification*, Int. Workshop on Content Based Multimedia Indexing, Brescia, Italy, September 19-21, 2001.
- [3] E. Allamanche, J. Herre, O. Helmuth, B. Froba, T. Kasten, and M. Cremer, *Content-based identification of audio material using mpeg-7 low level description*, Proc. of ISMIR 2001.
- [4] F. Kurth, A. Ribbrock, and M. Clausen, *Identification of highly distorted audio material for querying large scale databases*, in Proc. AES 112th Int. Conv., May 2002.
- [5] C. Burges, J. Platt, and S. Jana, *Extracting noise-robust features from audio data*, Proc. of ICASSP 2002.
- [6] J.S.Seo, J.Haitsma, T. and C.D.Yoo, *A robust image fingerprinting system using the radon transform*, Signal Processing:Image Communication, 325339 (2004).
- [7] F.Lefbvre, J. and B.Macq, *A robust soft hash algorithm for digital image signature*, EUSIPCO 2002.
- [8] A.Swaminathan, Y. and M.Wu, *Image hashing resilient to geometric and filtering operations*, IEEE MMSP 2004.
- [9] S.Mavandadi and P.Aarabi, *Rotation invariance in images*, Proc. of ICASSP 2007.
- [10] C. Lu, C.Y.Hsu, S. and Chang, P., *Robust mesh based hashing for copy detection and tracing of images*, Proc. of ICME 2004.
- [11] <http://www.advestigo.com>
- [12] <http://www.auditude.com>
- [13] <http://www.civolution.com>
- [14] <http://www.vobileinc.com>
- [15] Miller, M.L.; Rodriguez, M.A.; Cox, I.J; *Audio Fingerprinting: Nearest Neighbor Search in High Dimensional Binary Spaces*. The Journal of VLSI Signal Processing, Volume 41, Number 3.
- [16] Haitsma, J.; Kalker, T.; *A Highly Robust Audio Fingerprinting System*, Prof. of ISMIR 2002.
- [17] US Patent 2002/0178410.
- [18] Varna, A.L.; Swaminathan, A.; Wu, M.; *A Decision Theoretic Framework for Analyzing Binary Hash-based Content Identification Systems*. Proc. of DRM'08.
- [19] McCarthy, E.; Balado, F.; Silvestre, G.; Hurley, N.; *A framework for soft hashing and its applications to robust image hashing*. Proc. of ICIP 2004.