Indexing the Signature Quadratic Form Distance for Efficient Content-Based Multimedia Retrieval

Christian Beecks • Jakub Lokoč ○ Thomas Seidl • Tomáš Skopal ○

•Data Management and Data Exploration Group
•RWTH Aachen University, Germany
•{beecks,seidl}@cs.rwth-aachen.de

○SIRET Research Group, Faculty of Mathematics and Physics
○Charles University in Prague, Czech Republic
○{lokoc,skopal}@ksi.mff.cuni.cz

ABSTRACT

The *Signature Quadratic Form Distance* has been introduced as an adaptive similarity measure coping with flexible content representations of various multimedia data. Although the Signature Quadratic Form Distance has shown good retrieval performance with respect to their qualities of effectiveness and efficiency, its applicability to index structures remains a challenging issue due to its dynamic nature. In this paper, we investigate the *indexability* of the Signature Quadratic Form Distance regarding *metric access methods*. We show how the distance's inherent parameters determine the indexability and analyze the relationship between effectiveness and efficiency on numerous image databases.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Retrieval Models, Search Process; H.3.1 [**Content Analysis and Indexing**]: Indexing Methods

General Terms

Theory, Experimentation, Performance

Keywords

Signature Quadratic Form Distance, content-based retrieval, metric access method, multimedia database

1. INTRODUCTION

The ever-increasing amount of complex multimedia data including images, videos, and music challenges the effectiveness and efficiency of today's content-based analysis and retrieval techniques and systems [6, 14, 24, 26] which support users in searching and browsing voluminous multimedia databases in an interactive and efficient manner.

Supported by such content-based retrieval systems, users frequently issue content-based similarity queries by selecting already displayed multimedia objects or by sketching the intended object contents. Given an example multimedia object or sketch, the retrieval system then searches for the most related database objects with respect to the query object. In case of content-based retrieval purpose this relationship is frequently obtained by measuring the similarity between the query and each database object by means of distance functions which finally determine the most similar multimedia objects returned to the user.

The effectiveness as well as efficiency of this content-based retrieval process depends on the applied retrieval model, comprising feature representation and similarity measure, and the query processing method. It has been shown that the combination of adaptive similarity measures, such as the *Hausdorff Distances* [12, 19], the *Earth Mover's Distance* [22], and the *Signature Quadratic Form Distance* [1, 3], and flexible feature representations, such as *feature signatures* [21], provides good retrieval performance and extensive applicability to nearly all kinds of multimedia data which can be expressed by the corresponding feature representation. Thus, the problem of retrieval performance in terms of effectiveness is attributed to the retrieval model, while the problem of efficiency is related to the way of processing content-based similarity queries.

In this paper, we aim at improving the efficiency of the content-based retrieval process by making use of *metric access methods* [5, 23, 29]. For this purpose, we focus on the recently introduced Signature Quadratic Form Distance showing good retrieval performance for various multimedia databases [2]. We investigate the distance's inherent similarity function, so far only examined to adapt the distance to specific domains, and show how this similarity function affects the *indexability* of the Signature Quadratic Form Distance. We include the following contributions:

- A brief overview of the *metric space approach* as a fundamental prerequisite for metric access methods.
- An investigation of the Signature Quadratic Form Distance's inherent similarity function and their relationship to indexability.

ICMR '11, April 17-20, Trento, Italy

- A simple approach to process content-based similarity queries efficiently.
- An evaluation on numerous benchmark image databases showing the benefit of our findings.

The structure of this paper is as follows: in Section 2 we describe the content-based similarity model including feature representation and similarity measure. In Section 3 we review basic principles of metric space approaches. In Section 4 we investigate the indexability of the Signature Quadratic Form Distance before we outline a simple query processing approach in Section 5. The experimental results are reported in Section 6 before we conclude our paper with an outlook on future research directions in Section 7.

2. CONTENT-BASED SIMILARITY MODEL

In this section, we present the used content-based similarity model involving feature representation and similarity measure.

Representing multimedia objects by features in some feature space is a challenging task of nearly all content-based analysis and retrieval techniques. Whereas specific object recognition tasks, such as *copy*, *duplicate*, or *near-duplicate* detection, require the features to be accessible in an unaggregated way, content-based retrieval approaches frequently require some degree of generalization in order to cope with different similarity notions. As a consequence, extracted features are aggregated and approximated by so-called *feature representations*.

While numerous approaches aim at aggregating extracted features into an equi-length feature vector [15], namely *feature histogram*, which can be compared by using adaptable distance functions [10, 20], recent approaches tend to approximate the object's contents via more flexible feature representations, so-called *feature signatures* [21]. This type of feature representation reasonably adjusts to the contents of individual multimedia objects and can be compared by making use of adaptive similarity measures [2], such as the Hausdorff Distances [12, 19], the Earth Mover's Distance [22], and the Signature Quadratic Form Distance [1, 3].

In general, feature signatures exhibit high applicability to any kind of local features [7, 17, 27] by aggregating the object's features according to a partitioning of the feature space $\mathcal{FS}$ for each multimedia object individually. Consequently, the contents of each multimedia object is reflected by a single feature signature. For this purpose, the features are frequently partitioned via a clustering algorithm, e.g. k-means, and each partition is stored by a representative with the corresponding weight reflecting the number of features belonging to the current partition. Frequently, centroids of the clustering algorithm are chosen as representatives. In general, a feature signature is defined as follows.

Definition 1. *Feature Signature*
Given a feature space $\mathcal{FS} \subseteq \mathcal{R}^n$, the feature signature S^o of object o is defined as a set of tuples from $\mathcal{FS} \times \mathcal{R}$ comprising representatives $r^o \in \mathcal{FS}$ and weights $w^o \in \mathcal{R}^+$:

$$S^o = \{\langle r^o, w^o \rangle | \; r^o \in \mathcal{FS} \;\wedge\; w^o \in \mathcal{R}^+\}.$$

We depict an example of feature signatures according to a feature space comprising position and color information in Figure 1 where we visualize three images and their corresponding feature signatures. The feature signatures' representatives are depicted via circles in the corresponding color. The weights are reflected via the diameter of the circles.

Figure 1: Three example images with their corresponding feature signature visualizations.

So far, we have defined feature signatures as a common feature representation form of multimedia data. In the remainder of this section we focus on the Signature Quadratic Form Distance. As mentioned above, this distance is an adaptive similarity measure introduced for the comparison of any kind of multimedia data which can be represented by feature signatures. For this purpose, the distance makes use of a *similarity function* that determines a similarity value between any two feature signatures' representatives. In general, this similarity function behaves inversely proportional to a distance function between two representatives, i.e. the lower the representatives' distance, the higher the corresponding similarity value and vice versa. In [3] three example similarity functions were proposed:

- *Minus function:* $f_-(r_i, r_j) = -d(r_i, r_j)$
- *Heuristic function:* $f_h(r_i, r_j) = \frac{1}{\alpha + d(r_i, r_j)}$
- *Gaussian function:* $f_g(r_i, r_j) = e^{-\alpha \cdot d^2(r_i, r_j)}$

It turns out that the Gaussian function with the parameter $\alpha \in \mathcal{R}^+$ adapted to the current multimedia database exhibits the highest retrieval performance in terms of effectiveness, while the minus function results in the lowest computation time [2]. In Section 4, we will show that these parameters strongly affect the indexability of the Signature Quadratic Form Distance which is defined as follows.

Definition 2. *Signature Quadratic Form Distance*
Given two feature signatures $S^q = \{\langle r_i^q, w_i^q \rangle, i = 1, \ldots, n\}$ and $S^p = \{\langle r_i^p, w_i^p \rangle, i = 1, \ldots, m\}$ and a similarity function f_s over some feature space $\mathcal{FS}$, the Signature Quadratic Form Distance $SQFD_{f_s}$ between S^q and S^p is defined as:

$$SQFD_{f_s}(S^q, S^p) = \sqrt{(w_q | - w_p) \cdot A_{f_s} \cdot (w_q | - w_p)^T},$$

where $A_{f_s} \in R^{(n+m)\times(n+m)}$ is the similarity matrix arising from applying the similarity function f_s to the corresponding representatives, i.e. $a_{ij} = f_s(r_i, r_j)$. Furthermore, $w_q = (w_1^q, \ldots, w_n^q)$ and $w_p = (w_1^p, \ldots, w_m^p)$ form weight vectors, and $(w_q | - w_p) = (w_1^q, \ldots, w_n^q, -w_1^p, \ldots, -w_m^p)$ denotes the concatenation of weights w_q and $-w_p$.

As can be seen in Definition 2, the Signature Quadratic Form Distance takes into account the similarity values between any two representatives according to the similarity function f_s. This similarity relationship is reflected within the similarity matrix A_{f_s} which has to be determined for each distance computation individually. Thus, the complexity of a single distance computation is in $\mathcal{O}\left((n+m)^2 \cdot \mathcal{O}(f_s)\right)$ where n and m denote the size of feature signatures S^q and S^p, respectively, and $\mathcal{O}(f_s)$ denotes the complexity of the similarity function f_s.

In order to process content-based similarity queries, the computation of the Signature Quadratic Form Distance has to be carried out for each database object individually. Although this process can be parallelized, query response times can grow from seconds to minutes when increasing the number of objects contained in the multimedia database. One promising approach to tackle this scalability issue is *metric access methods* which organize the data in some *metric space* implied by the similarity measure. We briefly outline the basic principles of the metric space approach in the following section.

3. METRIC SPACE APPROACH

In this section, we outline the basic principles of metric spaces and their ability to answer content-based similarity queries efficiently.

A *metric space* [5, 23, 29] consists of a feature representation domain, in the scope of this paper the feature representation domain denotes the set of all possible feature signatures, and a distance function δ which has to satisfy the metric postulates[1]: *non-negativity*, *identity of indiscernibles*, *symmetry*, and *triangle inequality*. In this way, metric spaces allow domain experts to model their notion of content-based similarity by an appropriate feature representation and distance function serving as similarity measure. At the same time, this approach allows database experts to design metric index structures, the so-called *metric access methods*, for efficient query processing of content-based similarity queries, which rely on the distance function δ only, i.e. these methods do not necessarily know the structure of the objects' feature representation.

According to the metric postulates mentioned above, metric access methods organize database objects by grouping them based on their distances with the aim of minimizing not only traditional database costs like I/O but also the number of costly distance function evaluations. For this purpose, nearly all metric access methods apply a simple filtering rule which can be directly derived from the triangle inequality: the *lower-bounding principle*.

We illustrate this fundamental principle in Figure 2 where we depict the query object q, some pivot element p, and a database object o over some metric space. Note that pivot elements are used to group database objects and to improve the efficiency of the search process by pruning whole parts of the metric index structure. Given a range query (q, r), we aim at estimating the distance $\delta(q, o)$ by making use of $\delta(q, p)$ and $\delta(o, p)$, the latter is already stored within the metric access method. Due to the triangle inequality, we can safely filter object o and also all objects o' contained in the same group for which holds that $\delta(o', p) \leq \delta(o, p)$ if the lower-bound $LB(\delta(q, o)) = |\delta(q, p) - \delta(o, p)| \geq r$.

So far, we have only considered the distance functions' metric properties in order to apply the lower-bounding principle and to obtain exact search results. However, the efficiency of metric access methods relies mainly on the data distribution. If the data objects are not naturally well clustered, then it might be impossible for metric access methods to process content-based similarity queries efficiently. This corresponds to the similar problem in high-dimensional vector spaces, the curse of dimensionality [4].

[1]Note that even non-metric distance functions can be turned into metric ones as shown in [25].

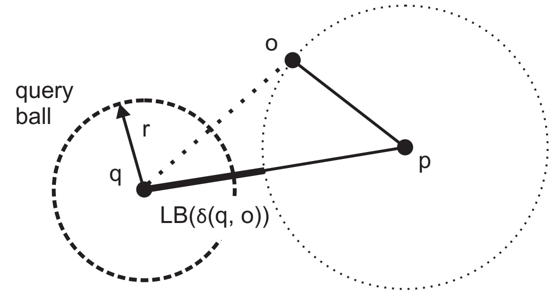


Figure 2: The lower-bounding principle.

In the following, we denote the probability of creating metric access methods which might improve the efficiency of content-based similarity queries as *indexability*. Metric spaces exhibit poor indexability if all distance values are nearly the same. In this case any filtering based on the triangle inequality cannot be successful, because the determined lower-bounds are always smaller than any meaningful range query radius and consequently processing content-based similarity queries with metric access methods deteriorates to the sequential scan.

One measure indicating the indexability of a given database $\mathbb{S}$ and a distance function δ is the *intrinsic dimensionality* [5] which is defined as follows:

$$\rho(\mathbb{S}, \delta) = \frac{E(\delta|\mathbb{S})^2}{2 \cdot Var(\delta|\mathbb{S})},$$

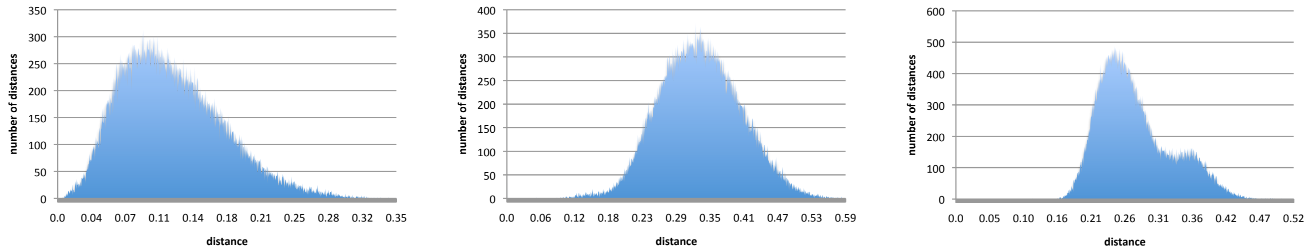
where $E(\delta|\mathbb{S})$ is the expected distance value and $Var(\delta|\mathbb{S})$ is the variance of distance values within the database $\mathbb{S}$. Intuitively, the intrinsic dimensionality ρ reflects the distance distribution in a single compressed value. The lower this value the better the indexability, and vice versa. We depict three example distance distributions of the *Corel Wang* database [28] and their intrinsic dimensionality ρ in Figure 3. As can be seen in the figure, shifting the distance distribution to smaller distance values will result in a lower value of the intrinsic dimensionality.

In the following section, we will investigate the indexability of the Signature Quadratic Form Distance and examine the relationship between the similarity functions' inherent parameters and the intrinsic dimensionality.

4. INDEXABILITY OF THE SIGNATURE QUADRATIC FORM DISTANCE

In this section, we investigate the indexability of the Signature Quadratic Form distance by focusing on the distance's inherent similarity function.

So far, the similarity function was only examined for the purpose of adapting the Signature Quadratic Form Distance to specific multimedia databases in order to maximize their qualities of effectiveness. In particular it turns out the Gaussian function $f_g(r_i, r_j) = e^{-\alpha \cdot d^2(r_i, r_j)}$ leads to the highest



(a) $\rho = 2.5,\ \alpha = 0.01$ (b) $\rho = 11.1,\ \alpha = 1.0$ (c) $\rho = 11.9,\ \alpha = 10.0$

Figure 3: Distance distributions of the *Corel Wang* database for the Signature Quadratic Form Distance using Gaussian similarity function with varying parameter α. The intrinsic dimensionality is denoted as ρ.

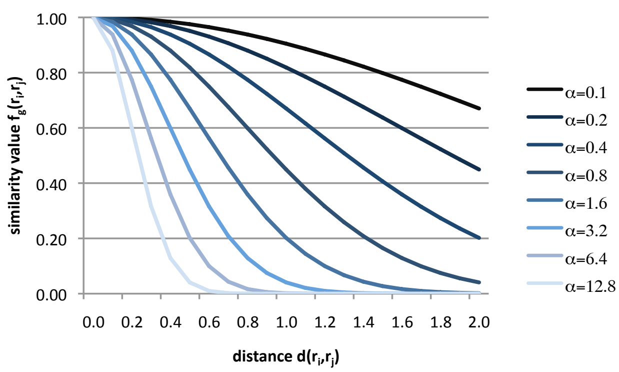


Figure 4: Similarity values of the Gaussian function $f_g(r_i, r_j) = e^{-\alpha \cdot d^2(r_i,r_j)}$ by varying the parameter the distance $d(r_i, r_j)$.

retrieval results in terms of *mean average precision* values [15] when adjusting the parameter $\alpha \in \mathcal{R}$ to the multimedia database accordingly.

Let us now take a closer look at this Gaussian function f_g for which we depict the function values regarding different choices of α in Figure 4. It can be seen in the figure, that the increase of α is accompanied by the increase of the similarity function's slope. Consequently, distance values among representatives can result in more different similarity values, depending on the value of α and the difference of the distance values. For instance, given two distance values $d_1 = 0.2$ and $d_2 = 0.8$ between some feature signatures' representatives, then it holds for the Gaussian functions $f_{g,\alpha}$ and $f_{g,\beta}$ with parameters $\alpha \leq \beta$ that $f_{g,\alpha}(d_1) - f_{g,\alpha}(d_2) \leq f_{g,\beta}(d_1) - f_{g,\beta}(d_2)$. In other words, decreasing the value of α will result in more indistinguishable similarity values. In general, adjusting the parameter α of the Gaussian function f_g to the database leads to smaller distance values of the Signature Quadratic Form Distance and thus reduces the expected distance value of the distance distribution which might result in a lower intrinsic dimensionality.

In order to elucidate our understanding of indexability, we depict three distance distributions of the *Corel Wang* database [28] in Figure 3, where we gathered 100,000 random Signature Quadratic Form Distance values using the Gaussian function f_g. It can be seen in the figure that the parameter α changes the distance distribution and thus the intrinsic dimensionality. While the intrinsic dimensionality stays above a value of 11 for $\alpha = 1.0$ and $\alpha = 10.0$, it improves to a value of 2.5 for $\alpha = 0.01$. As mentioned in the previous section, small values of intrinsic dimensionality will probably allow better metric indexing, while large values are difficult to index with metric access methods. We can see from this example that the indexability of the Signature Quadratic Form Distance is determined by the choice of the similarity function or rather the parameter α of a specific similarity function.

Before we continue with an in-depth empirical investigation of the relationship between intrinsic dimensionality and retrieval performance regarding the qualities of effectiveness and efficiency on different image databases, we describe the pivot table approach for efficient query processing of content-based similarity queries as an easy to understand example of a metric access method in the following section.

5. PIVOT TABLE FOR EFFICIENT QUERY PROCESSING

In this section, we describe how content-based similarity queries can be processed efficiently by making use of pivot tables.

The basic idea of a *pivot table*, which was originally introduced as LAESA [16], is to precompute distances $\delta(o_i, p_j)$ between all database objects $o_i \in \mathbb{S}$ and a selected static set of pivots elements $p_j \in \mathbb{S}$ and to represent each database object by means of its distances to the pivot elements. Given m pivot elements, the database objects are then represented by an m-dimensional vector in the m-dimensional pivot space, as illustrated in Figure 5.

Suppose a content-based similarity query q with range r is issued. This query will be processed in the following steps: first it is mapped to the point $q' = (\delta(q, p_1), \ldots, \delta(q, p_m))$ in the pivot space. Second the corresponding *maximum distance* L_{max} with the same range r centered at q' is issued in the pivot space and potential database objects are gathered. Finally the remaining non-filtered database objects are refined using the original distance function δ. By processing content-based similarity queries in this way, the retrieval results are guaranteed to be complete, i.e. non-approximate.

The performance of a pivot table in terms of efficiency depends on the number of pivot elements, the pivot selection strategy, the organization of the pivot table, and also the processing method for finding and refining database objects.

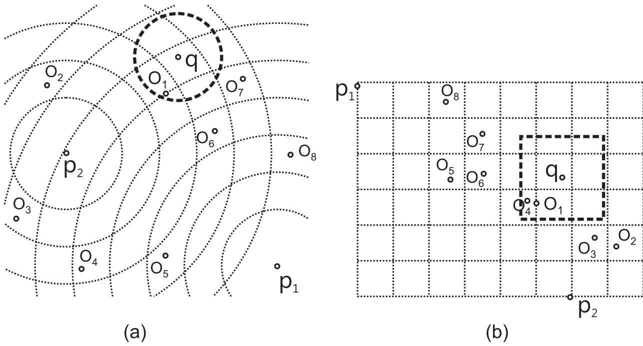


Figure 5: The basic idea of a pivot table: mapping the objects from the original space (a) to the pivot space (b).

However, tuning pivot tables in order to achieve the highest possible speed-up is meaningless, as long as the underlying multimedia database with the current distance function exhibits high intrinsic dimensionality. In order to evaluate the indexability of the Signature Quadratic Form Distance regarding the parameter modification of different similarity functions, we thus take the basic approach of pivot tables: finding and refining the database objects according to a sequential scan. The experimental evaluation is described in the following section.

6. EXPERIMENTAL EVALUATION

In this section, we evaluate the indexability of the Signature Quadratic Form Distance with respect to different benchmark image databases. For this purpose, we make use of the *Corel Wang* database [28] comprising 1,000 images from ten different topics, the *Coil 100* database [18] comprising 7,200 images of 100 different objects, the *101 objects* database [8] consisting of 9,196 images from 101 categories, the *MIR Flickr* database [11] including 25,000 web-images with textual annotations, and the *ALOI* database [9] which is similar to the *Coil 100* database but comprising 72,000 images. We depict some example images of the aforementioned databases in Figure 6.

We extracted feature signatures based on seven-dimensional features $(L, a, b, x, y, \chi, \eta) \in \mathcal{FS}$ including color information (L, a, b), position information (x, y), contrast information χ, and coarseness information η. These features were randomly extracted for each image and then aggregated by applying an adaptive variant of the k-means clustering algorithm described in [13]. Thus, we obtain one feature signature for each single image, which vary in size between 5 and 115 representatives. On average a feature signature consists of 54 representatives.

We perform all experiments on an Intel(R) Core(TM)2 Quad CPU Q9550 2.83GHz machine with 8 GB main memory based on a C++ implementation.

In order to evaluate the indexability of the Signature Quadratic Form Distance regarding the heuristic and Gaussian similarity function, we first measure the distance's metric behavior. We empirically evaluated the distance's metric behavior by taking 100,000 random triplets of feature signatures for each combination of image database, similarity function and their parameter α and checked whether all possible triangle inequalities are satisfied, i.e. if it holds

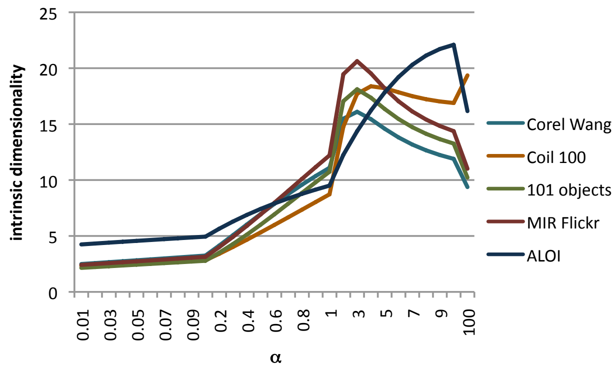
Figure 6: Example images of the *Corel Wang*, *Coil 100*, *101 objects*, *MIR Flickr*, and *ALOI* databases (from left to right).

that $SQFD_{f_s}(S^i, S^j) \leq SQFD_{f_s}(S^i, S^l) + SQFD_{f_s}(S^l, S^i)$ for all feature signatures S^i, S^j, S^l. We observed that all triplets satisfy the triangle inequality and that the Signature Quadratic Form Distance completely shows metric behavior. Consequently, we state that metric access methods will not affect the retrieval performance in terms of effectiveness and that the retrieval results obtained by such methods are non-approximate, i.e. exact.

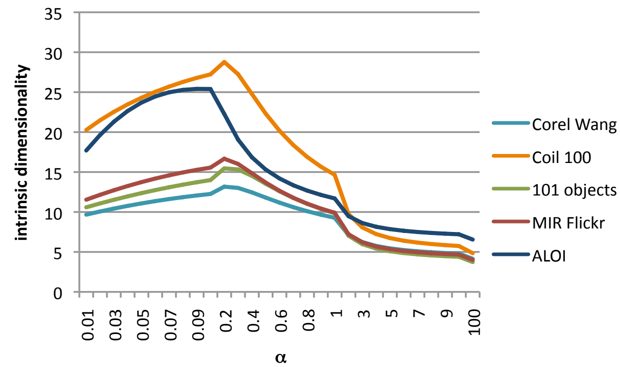
In the remainder of this section, we first analyze the indexability in terms of *intrinsic dimensionality* and *mean average precision* values [15] as they indicate whether the Signature Quadratic Form Distance allows for efficient and effective metric indexing. Then, we study the increase in efficiency in terms of query response times and number of distance computations.

In order to measure the intrinsic dimensionality, we evaluated 100,000 random distance computations for each of the aforementioned databases. The results are shown in Figure 7. As can be seen in Figure 7(a), all databases show a similar behavior: the intrinsic dimensionality decreases by decreasing the value of parameter α of the Gaussian function. In case of the heuristic function, c.f. Figure 7(b), the intrinsic dimensionality decreases with an increasing value of parameter α. Both figures indicate a significant improvement of the intrinsic dimensionality when decreasing or increasing the value of parameter α.

Let us now link these results regarding the intrinsic dimensionality to the retrieval quality in terms of effectiveness. For this purpose, we evaluated mean average precision values for 100 random queries for the same databases. We took the provided class/category/annotation information of each image database as ground truth. The results are depicted in Figure 8 for the Signature Quadratic Form Distance applying the Gaussian function f_g and the heuristic function f_h, respectively. Both similarity functions show a very similar behavior: the mean average precision values decrease when the intrinsic dimensionality improves. Thus it seems that indexability of the Signature Quadratic Form Distance comes at costs of low retrieval quality. In general this is true, because lower intrinsic dimensionality is accompanied by lower mean average precision values. However, it can be recognized that this trade-off is well-natured. For instance, consider the use of the Signature Quadratic Form Distance with the Gaussian function f_g on the *Corel Wang* database. For a value of parameter $\alpha = 1.0$ we observe a mean average precision value of approximately 0.5 while the intrinsic dimensionality is approximately 11.1. Decreasing the pa-

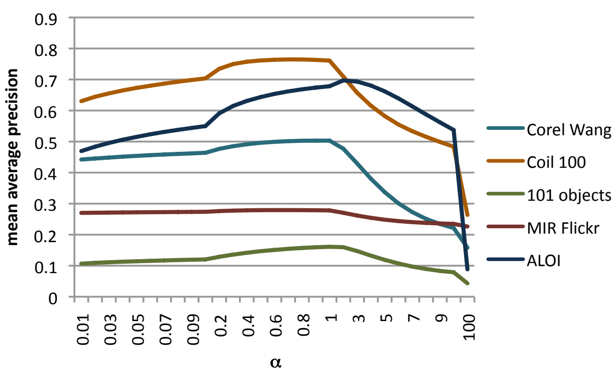


(a) Gaussian function f_g

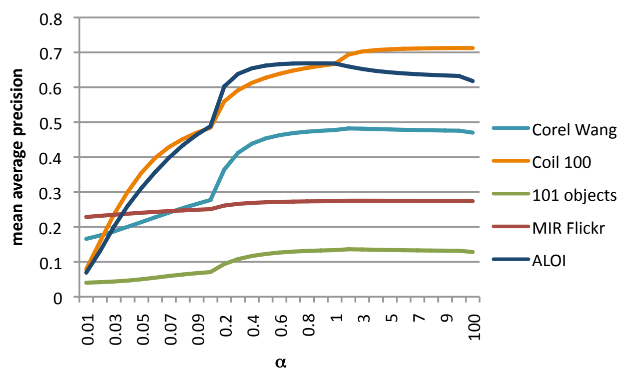


(b) Heuristic function f_h

Figure 7: Intrinsic dimensionality of the Signature Quadratic Form Distance by making use of (a) the Gaussian function $f_g(r_i, r_j) = e^{-\alpha \cdot d^2(r_i, r_j)}$ and (b) the heuristic function $f_h(r_i, r_j) = \frac{1}{\alpha + d(r_i, r_j)}$.



(a) Gaussian function f_g



(b) Heuristic function f_h

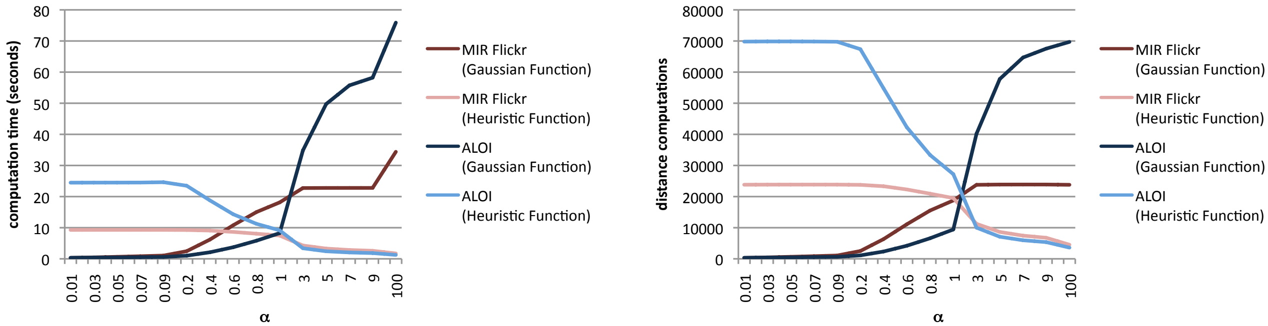
Figure 8: Mean average precision values of the Signature Quadratic Form Distance by making use of (a) the Gaussian function $f_g(r_i, r_j) = e^{-\alpha \cdot d^2(r_i, r_j)}$ and (b) the heuristic function $f_h(r_i, r_j) = \frac{1}{\alpha + d(r_i, r_j)}$.

rameter α to a value of 0.1 reduces the intrinsic dimensionality to a value of 3.2 while the mean average precision value only decreases to 0.46. Thus we improve the indexability by a factor of approximately 3.4 while the retrieval quality is only reduced by a factor of 1.1. This behavior is also observable for other combinations of databases and similarity functions.

As a first result, we claim that the indexability strongly depends on the parameters of the similarity function of the Signature Quadratic Form Distance. In the following, we report query response times for the two largest databases, *MIR Flickr* and *ALOI*, in order to see how the intrinsic dimensionality affects the query processing behavior. For this purpose we organize the databases via a pivot table comprising 100 pivot elements (c.f. Section 5).

The computation time values in seconds and the corresponding number of distance computations are depicted in Figure 9(a) and Figure 9(b), respectively. The measured values are averaged over 100 randomly chosen 10-nearest-neighbor queries. Corresponding to the intrinsic dimensionality evaluated above, the computation time values for the Signature Quadratic Form Distance applying a Gaussian function decrease when the value of parameter α becomes smaller. The heuristic function behaves inversely: increasing the value of α decreases the computation time values. Thus, choosing the value of parameter α smaller than 0.2 and 0.09 for the *ALOI* and *MIR Flickr* database, respectively, results in query response times below one second when carrying out the Signature Quadratic Form Distance computation with the Gaussian function. As can be seen in the figures, the number of distance computations is proportional to the computation time values.

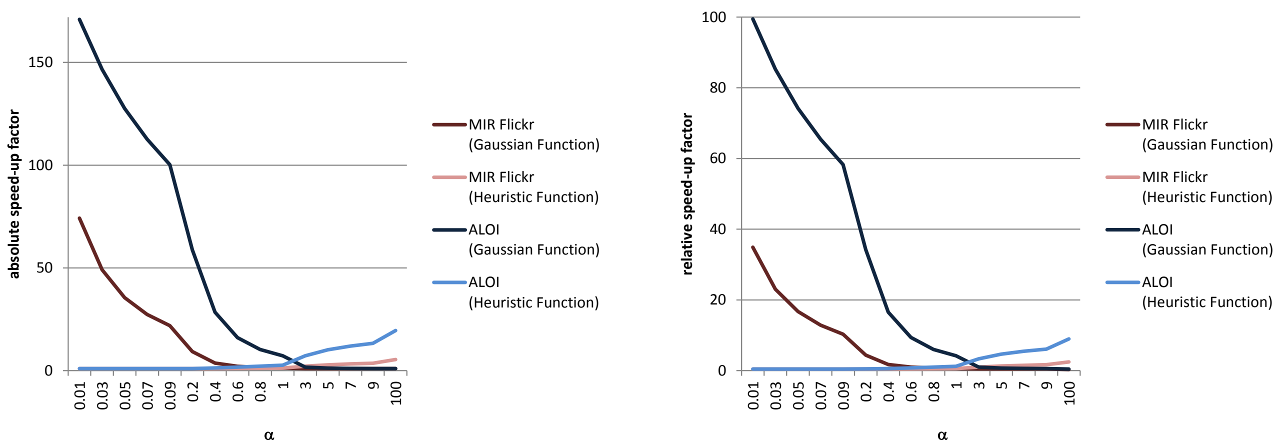
Based on the computation time values needed to answer 10-nearest-neighbor queries, we depict the speed-up factor in Figure 10. We evaluated the absolute speed-up factor by comparing the pivot table approach with a naive sequential scan in Figure 10(a) and the relative speed-up factor by comparing the computation time values of the pivot table approach needed for the best mean average precision values with the computation time values needed for the current value of parameter α in Figure 10(b). Both speed-up factors increase significantly for the Gaussian function by decreasing



(a) computation time (seconds)

(b) distance computations

Figure 9: Computation time values in seconds (a) and number of distance computations (b) needed to answer 10-nearest-neighbor queries by making use of pivot tables as metric access method.



(a) absolute speed-up factor

(b) relative speed-up factor

Figure 10: Speed-up factor compared to the sequential scan (a) and compared to pivot table used for the best mean average precision values (b).

the value of parameter α. The maximum absolute speed-up factor of 170 is reached for the *ALOI* database which is approximately twice as high as the corresponding relative speed-up factor. At the same time, the retrieval quality in terms of mean average precision value does not fall below 68%. This mean average precision value might be acceptable depending on the current application. Similar results can be observed for the *MIR Flickr* database where a maximum absolute speed-up factor of 74 is reached while maintaining a high retrieval quality of more than 97%.

To sum up, we have shown that the Signature Quadratic Form Distance is well indexable via metric access methods by varying the similarity function's parameter. By adjusting the Signature Quadratic Form Distance's parameters according to individual user's needs, the trade-off between efficiency and effectiveness can be balanced.

7. CONCLUSIONS AND FUTURE WORK

In this paper, we have investigated the indexability of the Signature Quadratic Form Distance for efficient content-based multimedia retrieval. By making use of a simple metric access method, we have shown that the indexability of the Signature Quadratic Form Distance depends on its inherent similarity function which, so far, was only examined for the purpose of adapting the distance to specific domains. As a result, we have reached a promising trade-off between indexability and retrieval quality: we have shown how to improve the efficiency of the content-based retrieval process by a factor of more than 170 while maintaining a retrieval quality of more than 68%.

As future work, we plan to investigate other metric access methods in order to improve the efficiency of the Signature Quadratic Form Distance even further. Additionally, we also plan to study the indexability of the other state-of-the-art similarity measures.

Acknowledgments

The projects underlying this report were funded by the German Federal Ministry of Economics and Technology under project funding reference number 01MQ09014 and by the

Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center (SFB) 686. The responsibility for the content of this publication lies with the authors. In addition, this research has been supported in part by Czech Science Foundation project GAČR 201/09/0683 and by the grant SVV-2011-263312 (second and fourth author).

8. REFERENCES

[1] C. Beecks, M. S. Uysal, and T. Seidl. Signature Quadratic Form Distances for Content-Based Similarity. In *Proc. ACM International Conference on Multimedia*, pages 697–700, 2009.

[2] C. Beecks, M. S. Uysal, and T. Seidl. A Comparative Study of Similarity Measures for Content-Based Multimedia Retrieval. In *Proc. IEEE International Conference on Multimedia & Expo*, pages 1552–1557, 2010.

[3] C. Beecks, M. S. Uysal, and T. Seidl. Signature Quadratic Form Distance. In *Proc. ACM International Conference on Image and Video Retrieval*, pages 438–445, 2010.

[4] C. Böhm, S. Berchtold, and D. Keim. Searching in High-Dimensional Spaces – Index Structures for Improving the Performance of Multimedia Databases. *ACM Computing Surveys*, 33(3):322–373, 2001.

[5] E. Chávez, G. Navarro, R. Baeza-Yates, and J. L. Marroquín. Searching in Metric Spaces. *ACM Computing Surveys*, 33(3):273–321, 2001.

[6] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, 40(2):1–60, 2008.

[7] T. Deselaers, D. Keysers, and H. Ney. Features for image retrieval: an experimental comparison. *Information Retrieval*, 11(2):77–107, 2008.

[8] L. Fei-Fei, R. Fergus, and P. Perona. Learning generative visual models from few training examples an incremental bayesian approach tested on 101 object categories. In *Proc. of the Workshop on Generative-Model Based Vision*, 2004.

[9] J.-M. Geusebroek, G. J. Burghouts, and A. W. M. Smeulders. The Amsterdam Library of Object Images. *International Journal of Computer Vision*, 61(1):103–112, 2005.

[10] R. Hu, S. Rüger, D. Song, H. Liu, and Z. Huang. Dissimilarity measures for content-based image retrieval. In *Proc. IEEE International Conference on Multimedia & Expo*, pages 1365 –1368, 2008.

[11] M. J. Huiskes and M. S. Lew. The MIR flickr retrieval evaluation. In *Proc. ACM International Conference on Multimedia Information Retrieval*, pages 39–43, 2008.

[12] D. P. Huttenlocher, G. A. Klanderman, and W. A. Rucklidge. Comparing Images Using the Hausdorff Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, 1993.

[13] W. K. Leow and R. Li. The analysis and applications of adaptive-binning color histograms. *Computer Vision and Image Understanding*, 94(1-3):67–91, 2004.

[14] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain. Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2(1):1–19, 2006.

[15] C. D. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. Cambridge University Press, New York, NY, USA, 2008.

[16] M. L. Micó, J. Oncina, and E. Vidal. A New Version of the Nearest-neighbour Approximating and Eliminating Search Algorithm (AESA) with Linear Preprocessing Time and Memory Requirements. *Pattern Recognition Letters*, 15(1):9–17, 1994.

[17] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1615–1630, 2005.

[18] S. Nene, S. K. Nayar, and H. Murase. Columbia Object Image Library (COIL-100). Technical report, Department of Computer Science, Columbia University, 1996.

[19] B. G. Park, K. M. Lee, and S. U. Lee. Color-based image retrieval using perceptually modified Hausdorff distance. *Journal on Image and Video Processing*, 2008:1–10, 2008.

[20] J. Puzicha, J. Buhmann, Y. Rubner, and C. Tomasi. Empirical evaluation of dissimilarity measures for color and texture. In *Proc. IEEE International Conference on Computer Vision*, pages 1165–1172, 1999.

[21] Y. Rubner and C. Tomasi. *Perceptual Metrics for Image Database Navigation*. Kluwer Academic Publishers, Norwell, MA, USA, 2001.

[22] Y. Rubner, C. Tomasi, and L. J. Guibas. The Earth Mover's Distance as a Metric for Image Retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.

[23] H. Samet. *Foundations of Multidimensional and Metric Data Structures*. Morgan Kaufmann, 2006.

[24] N. Sebe, M. S. Lew, X. Zhou, T. S. Huang, and E. M. Bakker. The state of the art in image and video retrieval. In *Proc. ACM International Conference on Image and Video Retrieval*, pages 1–8, 2003.

[25] T. Skopal. Unified framework for fast exact and approximate search in dissimilarity spaces. *ACM Trans. Database Syst.*, 32(4):46, 2007.

[26] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

[27] K. van de Sande, T. Gevers, and C. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 2010.

[28] J. Z. Wang, J. Li, and G. Wiederhold. SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture LIbraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(9):947–963, 2001.

[29] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search: The Metric Space Approach*. Springer-Verlag New York, Inc., 2005.