UNIVERSITE DE GENEVE

CENTRE UNIVERSITAIRE
D'INFORMATIQUE
COMPUTER VISION AND
MULTIMEDIA LABORATORY

Date: 10 July 2006
N° 06.01

TECHNICAL REPORT

Dual diffusion model of spreading activation for content-based image retrieval

S. Kosinov

Computer Vision Group
Computing Science Center, University of Geneva
24 rue du Général Dufour, CH - 1211 Geneva 4, Switzerland

e-mail: kosinov@cui.unige.ch

Abstract

This paper introduces a content-based information retrieval method inspired by the ideas of spreading activation models. In response to a given query, the proposed approach computes document ranks as their final activation values obtained upon completion of a diffusion process. This diffusion process, in turn, is dual in the sense that it models the spreading of the query's initial activation simultaneously in two similarity domains: low-level feature-based and high-level semantic.

The formulation of the diffusion process relies on an approximation that makes it possible to compute the final activation as a solution to a linear system of differential equations via a matrix exponential without the need to resort to an iterative simulation. The latter calculation is performed efficiently by adapting a sparse routine based on Krylov subspace projection method.

The empirical performance of the described dual diffusion model has been evaluated in terms of precision and recall on the task of content-based digital image retrieval in query-by-example scenario. The obtained experimental results demonstrate that the proposed method achieves better overall performance compared to traditional feature-based approaches. This performance improvement is attained not only when both similarity domains are used, but also when a diffusion model operates only on the feature-based similarities.

1 Introduction

There have been developed numerous contributions that aim to integrate higher-level semantic information into the information retrieval process so as to overcome problems arising due to synonymy, polysemy, data sparsity and semantic gap. A broad range of research efforts in this field share a common modeling apparatus of the vector space model and rely on the term-document matrix as a basic tool. Among these approaches are the latent semantic analysis [8], methods for query vector expansion [14], techniques for computing semantic kernels [7] and term proximities [23], as well as many other prominent contributions.

In our work we attempt to extend the traditional term-document matrix paradigm in the spreading activation model context to make it more suitable for content-based digital media retrieval. We consider the case when the documents in the collection at hand are associated with two classes of descriptors: low-level feature-based, such as dominant color in a digital image, and high-level semantic, such as keywords describing objects, location or a person shown. These two classes of document descriptors establish two separate similarity domains in which documents can be compared. Arguably, using the similarities in both of these domains and taking into account their inter-dependencies effectively may contribute substantially to the successful retrieval of relevant documents. And this is what the method described in this paper strives to accomplish. Below, we present an approach inspired by the spreading activation modeling ideas. The method processes a search query by computing document ranks as their final activation values obtained upon completion of a diffusion process. This diffusion process, in turn, is dual in the sense that it models the spreading of the query's initial activation simultaneously in two above mentioned similarity domains: low-level feature-based and high-level semantic.

Naturally, there exist many relevant contributions that are focused on achieving similar goals, based on analogous models or even share the same terminology [6, 12, 13]. One of these techniques has introduced semantic diffusion kernels [12] derived from the duality of the vector space representation model where a document can be seen as the counts of terms that appear in it, while a term can be regarded as the counts of documents in which it appears. The method reuses the very same source of information (the original term-document matrix) to augment the similarity of documents from the co-occurrence information of terms and vice versa, providing the final document similarity matrix as an equilibrium point of such an augmentation process. In contrast, the approach described in this paper does not specifically look to derive the ultimate document similarity matrix from the term-document one. Instead, we introduce a diffusion process that aggregates the information from two disparate external similarity sources with the term-document co-occurrence data, and directly produces the ranking of documents in response to a query.

Some other important differences of the proposed method can be highlighted by comparing it with the traditional spreading activation models [5, 6, 17]. These techniques have been designed to propagate initial activation through a network of connected nodes, and are set to proceed repeatedly through a sequence of steps of preadjustment, spreading, postadjustment and termination. The iterative nature of such an approach is likely to introduce a number of difficulties, especially in the target application domain of content-based digital media retrieval. For instance, the solution may lose accuracy quickly by accumulating errors from previous iterations; also, given the size of a typical collection containing thousands of images the computation may take too long to be applicable in interactive query-response scenarios. In addition to that, pure spreading activation has limited practical use unless extra constraints on distance, fan-out, path and activation are introduced [6]. This makes the closed-form mathematical modeling of the process more complicated. In our method, we formulate the spreading activation as a diffusion process by means of an appropriate approximation. This choice makes it possible to compute the final activation as a solution to a linear system of differential equations via a matrix exponential without the need to resort to an iterative simulation. The latter calculation is performed efficiently by adapting a sparse routine based on Krylov subspace projection method. Finally, we would like to mention a technique that attempts to improve the spreading activation model by taking into account second order term interactions derived from two-level co-occurrence data [1]. The proposed approach differs from this method in the important respect that not only second order but also all the higher order term interactions can be modeled in a unified fashion via a diffusion process.

In the sections that follow, we provide a more detailed description of the proposed dual diffusion spreading activation (DDSA) method. Section 2 will elaborate the main ideas behind the approach, derive its problem formulation, and address some concerns as for its efficiency and parametrization. This will be followed by the experimental results discussed in Section 3. Finally, conclusions and future perspectives will be presented in Section 4.

2 Dual diffusion model of spreading activation

As mentioned previously, our work has been inspired by the model derived in the classical studies of the associative retrieval [20] and mechanisms of human memory operations [19], known as spreading activation model.

In its original form, the spreading activation model is a simple processing technique applied to a network structure where related concepts are represented as connected nodes [5, 6, 17]. It works by propagating an initial activation from some node(s) through the network via the association links according to the formula:

$$In_j = \sum_i Out_i \, W_{ij}, \tag{1}$$

where In_j is the total input at node j, Out_i is the output of node i connected to node j, and W_{ij} is the weight describing the strength of association of the link connecting node i to node j. Subsequently, the output of the node in question Out_j is calculated as some function of its input In_j, and the process is repeated for the remaining nodes in the network.

Our method builds upon the foundations of the spreading activation model in order to improve information retrieval performance. The approach described below extends the framework so as to effectively combine the similarity data from two disparate similarity domains corresponding to the two classes of descriptors associated with the documents. This setup is motivated by the need to accommodate the fact that in multimedia retrieval the documents are usually represented by low-level (image color, motion vectors, etc.) and high-level (regular keywords, names, locations) features. The said combination of the two similarity domains is accomplished by modeling the spreading activation as a diffusion process that occurrs simultaneously in both of them.

An illustration of this idea is shown in Figure 1. Here, the document collection is shown to have 3 documents $D_1 \ldots D_3$ indexed by 4 terms $T_1 \ldots T_4$. The initial activation is located in node $a_{1,2}$ corresponding to term T_2 in document D_1. The diffusion process propagates the initial activation accross the terms of D_1 taking into account the high-level semantic similarities among $T_1 \ldots T_4$. At the same time, an identical diffusion process propagates the initial activation value for term T_2 across the documents $D_1 \ldots D_3$ in accordance with their low-level feature-based similarity, for instance, the dominant color in a digital image. Both low-level and high-level similarity domains are represented by their corresponding pairwise similarity matrices (not shown in Figure 1). The two diffusion processes, shown as arrows in Figure 1, compete simultaneously to propagate the initial activation throughout the term-document matrix. Much in the same way, the process illustrated for $a_{1,2}$ occurrs at every other entry of the matrix as well. Finally, upon completion of this dual diffusion process, the resulting activation values in the term-document matrix provide a natural way of ranking documents, ordering terms, performing document autoannotation and quantitative assessment of degree of association between a given semantic concept and alternative low-level similarity feature spaces, etc. However, in order to maintain the discourse focused and make the experimental evaluation

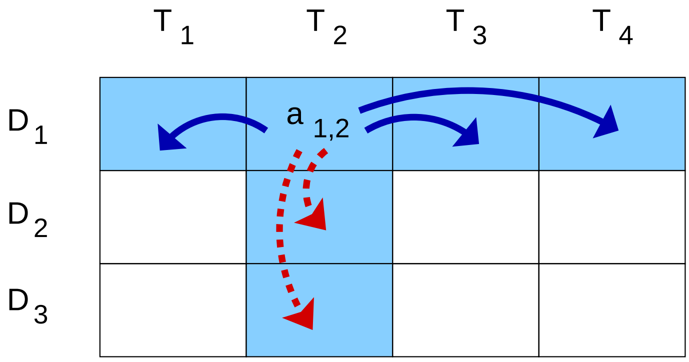

Figure 1: An simplified illustration of the duality of the spreading activation in the proposed method. Initial activation $a_{1,2}$ is propagated via a diffusion process accross terms within document D_1 according to the high-level semantic similarities of terms $T_1 \ldots T_4$ (shown as solid line arrows), as well as across the documents through the low-level feature-based similarities of documents $D_1 \ldots D_3$, (shown as dotted line arrows). The arrows show the direction of the propagation of initial activation.

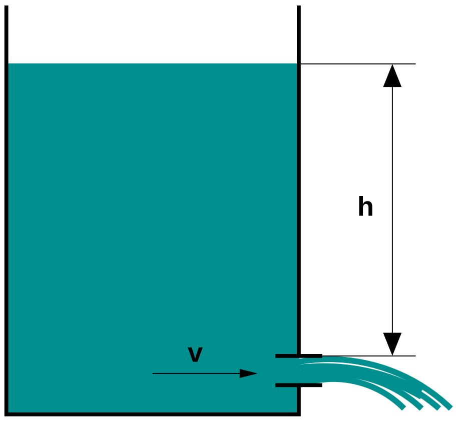

Figure 2: An illustration of the Torricelli's Theorem that relates the velocity of efflux, v, through a hole in the side of a filled tank and the current level in the tank, h, according to the formula: $v = \sqrt{2gh}$.

coherent, this paper we will only concentrate on the first possible application mentioned in the above list, and hence will consider a single-term query-by-example retrieval scenario, where the relevance of a given document to a certain query term is judged by the magnitude of its corresponding final activation value.

In the section that follows, we are going to describe the details of implementing such a dual diffusion spreading activation method, including an appropriate modeling approximation, non-iterative computation and estimation of the stopping criteria.

2.1 Fluid mechanics intuition

A natural way of modeling the diffusion process is to borrow some intuition from the disciplines of fluid mechanics and hydraulics. Indeed, there exists a substantial body of research that can be applied in our case. One of such research results is known as the Torricelli's Theorem [16] that establishes the link between the velocity of the outflow from, and the current level in a filled tank, as depicted in Figure 2. Using the above theorem would definitely allow to model a diffusion process well in accordance with the fluid mechanics paradigm. However, the non-linear dependence of the efflux velocity v on the current level h could lead to a formulation complicated enough to prevent efficient computation. This is why we have chosen to assume a linear dependence between the above two parameters, which has lead to the model that we discuss in the following section.

2.2 Problem formulation

Consider a collection of m documents $\{D_i | i = 1 \dots m\}$ indexed by n high-level semantic terms $\{T_j | j = 1 \dots n\}$ and some low-level features. Let the low-level feature-based similarity domain be represented by matrix $K \in \mathbb{R}^{m \times m}$ such that:

$$K = \begin{bmatrix} k(D_1, D_1) & \cdots & k(D_1, D_m) \\ k(D_2, D_1) & \cdots & k(D_2, D_m) \\ \vdots & \ddots & \vdots \\ k(D_m, D_1) & \cdots & k(D_m, D_m) \end{bmatrix}, \tag{2}$$

where $k(D_i, D_j) \in [0, 1]$ specifies the similariy of documents D_i and D_j according to their feature-based representation extracted automatically, e.g. from the analysis of image pixel colors, textures, etc. Also, let the high-level semantic similarity domain be defined by matrix $S \in \mathbb{R}^{n \times n}$ such that:

$$S = \begin{bmatrix} s(T_1, T_1) & \cdots & s(T_1, T_n) \\ s(T_2, T_1) & \cdots & s(T_2, T_n) \\ \vdots & \ddots & \vdots \\ s(T_n, T_1) & \cdots & s(T_n, T_n) \end{bmatrix}, \tag{3}$$

where $s(T_i, T_j) \in [0, 1]$ specifies the similariy of terms T_i and T_j according to their semantics, derived from some external source, such as WordNet [15].

Then, let $A \in \mathbb{R}^{m \times n}$ be the activation matrix whose initial non-zero values are set to a predefined constant $\alpha > 0$ in accordance with the query-by-example paradigm. That is, for a single-word query defined by term $\{T_q | q \in 1 \dots n\}$, we establish a corresponding set of documents Q from the training data set whose annotation includes T_q. Then, an element a_{ij} of A is assigned α if the training set document D_i belongs to query set Q and has term T_j in its annotation:

$$a_{ij} = \begin{cases} \alpha & \text{if } D_i \in Q \text{ and } D_i \text{ has } T_j; \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Alternatively, A can be thought of as a traditional term-document matrix, whose entries are initially zeroed out with the exception of the rows that correspond to the query set Q and are set according to the available annotation from the training data.

Finally, given the linearity assumption discussed in the previous section, we can model the dual diffusion process by specifying the rate of change of the activation level at every entry of A in a way somewhat similar in spirit to the traditional spreading activation model (1), as follows:

$$\frac{da_{ij}}{dt} = \sum_{l=1}^{n} a_{il} s(T_j, T_l) + \sum_{l=1}^{m} a_{lj} k(D_l, D_i). \tag{5}$$

Here, (5) states that a given activation value a_{ij} will increase rapidly if there are high activation entries in A corresponding to terms related to T_j within the same document D_i. In addition to that, (5) indicates that a_{ij} will also increase if there are other documents similar to D_i in the low-level feature-based representation that have high activation values corresponding to T_j.

We can rewrite (5) more compactly in the matrix form:

$$\frac{dA}{dt} = AS_B + K_B A, \tag{6}$$

where S_B and K_B are the balanced versions of matrices S and K given in (3) and (2), respectively. The balancing of the two matrices is necessary to ensure that the conservation law holds, i.e. the total activation in the system remains constant over time throughout the diffusion process. This constraint is enforced by modifying the diagonal elements of matrices S and K such that:

$$s_{jj} = -\sum_{i=1, i \neq j}^{n} s_{ij}, \tag{7}$$

$$k_{jj} = -\sum_{i=1, i \neq j}^{m} k_{ij}. \tag{8}$$

Apparently, (6) can be recognized as equivalent to a very popular problem frequently encountered in mathematics, physics and engineering applications:

$$\begin{cases} \boldsymbol{a}(t)' &= M\boldsymbol{a}(t),\ t \in [0,\tau] \\ \boldsymbol{a}(0) &= \boldsymbol{a}_0, \end{cases} \tag{9}$$

where we reassign $\boldsymbol{a} = vec(A)$, $M = I_n \otimes K_B + S_B \otimes I_m$ for appropriately sized identity matrices I_m, I_n and initial activation written in the form of a column vector $\boldsymbol{a}_0$. The analytic solution of (9) is given by

$$\boldsymbol{a}(t) = e^{tM}\boldsymbol{a}_0, \tag{10}$$

which implies that the final activation at any termination time τ may be obtained by immediately computing the solution via a matrix exponential, instead of running a costly and usually error-prone iterative simulation that traditional spreading activation models have to resort to. Once the final activation has been computed, the ranking of the of the documents with respect to query T_q is obtained simply by sorting the collection by the magnitude of the values in q-th column of the activation matrix.

In the two sections that follow, we discuss some techniques for determining when the dual diffusion process defined by (6) needs to be stopped, i.e. finding the termination time τ, and for computing solution (10) efficiently.

2.3 Stopping criteria

In order to specify the stopping criterion of the above described diffusion process, it is helpful to consider its behavior with respect to the admissible range of values of the termination time τ. On one hand, if τ is chosen too close to zero, the diffusion barely has a chance to propagate any activation, because it stops too soon. On the other hand, if τ is set too high, the result will not be of any practical use either, since the initial activation will have propagated and dissipated evenly throughout the system, rendering the activation process meaningless. Thus, the value of τ is to be selected so as to avoid the above two extremes.

While it is always possible to suggest a parameter estimation technique through the use of a validation data set, in our approach we make a conjecture as for the appropriate choice of τ that has proved to be quite accurate in practice and was confirmed by the experimental results discussed in section 3. Namely, we set τ to the shortest possible time necessary for the diffusion process to deplete any single activation source almost completely, until it reaches $\gamma = 1\% \ldots 5\%$ of its initial value:

$$\tau = \min_{a_{ij} \neq 0} \left(\frac{\log \gamma}{s_{jj} + k_{ii}} \right), \tag{11}$$

where $i = 1..m$, $j = 1..n$, γ is the target residual percentage level, and $a_{ij} \neq 0$ is the condition that ensures that only the sources of initial activation are considered.

2.4 Efficient calculation of final activation

As mentioned earlier, the final activation values can be obtained immediately from (10) by computing a matrix exponential. However, a direct approach could be inappropriate due to efficiency reasons. Indeed, the exponential power series expansion can be extremely costly to compute, the principal matrix M for thousands of documents and terms extended via Kronecker products will not likely fit in computer memory even for document collections of moderate size, and, most importantly, the result of matrix exponentiation is actually not needed in the explicit form, since solution (10) that represents the final activation requires only the action of the matrix exponential on a vector of initial activation.

Taking into account the above considerations, we have chosen a method for computing matrix exponentials proposed in [22], with subsequent adaptation to the problem at hand. The main underlying principle of this method is to approximate

$$\boldsymbol{a}(t) = e^{tM}\boldsymbol{a}_0 = \boldsymbol{a}_0 + \frac{(tM)}{1!}\boldsymbol{a}_0 + \frac{(tM)^2}{2!}\boldsymbol{a}_0 + \cdots \tag{12}$$

by an element of the Krylov subspace:

$$\mathcal{K}_p(tM, \boldsymbol{a}_0) = Span\{\boldsymbol{a}_0, (tM)\boldsymbol{a}_0, \ldots, (tM)^{p-1}\boldsymbol{a}_0)\}, \tag{13}$$

Table 1: Regular keyword terms and their statistics of occurrence in Carole's data set.

ocean	282	trees	347	crown	47
balloons	29	balloon	44	carousel	42
car	54	kitty	15	classroom	130
drawing	15	team	10	buildings	29
street	71	bike	19	billboard	21
sailboat	78	rocks	84	dolphins	22
bed	27	parrot	31	leaves	24
flowers	146	beach	99	deck	31
pool	158	flag	21	sunset	15
sunrise	26	ball	52	poles	87
tracks	59	snow	82	fields	57
trail	14	rainbow	20	church	26
doll	12	birthday	87	lagoon	32
pond	27	kitchen	54	umbrella	10
cow	13	table	18	tricycle	11
crib	11	river	19	peaks	13
forest	33	home	24	driveway	16
waterfall	14	vineyards	13	presents	14
orca	24	crowd	13	lake	13

where p, the dimension of the Krylov subspace, is far smaller than mn, the order of the principal matrix M. Another valuable property of this computational technique is the ability to attain the final result by using only the action of the principal matrix on a vector throughout the intermediate calculations. Thus, in order to achieve better efficiency, we take advantage of this quality by adapting the algorithm so that principal matrix M never needs to be formed explicitly:

$$M\boldsymbol{a} = vec(K_B A) + vec(A S_B) \tag{14}$$

With the above modifications in place, the matrix exponentiation algorithm achieves near-realtime performance whereby the prototype implementation computes the final activation for the collection of over five thousand images in less than a minute's time.

3 Experimental results

For an experimental evaluation of the proposed method on the task of content-based image retrieval we have selected a subset of Carole's annotated digital image collection [3]. This data set consists of $m = 5222$ images each annotated with regular keywords that form a vocabulary of $n = 57$ unique terms appearing in the annotation corpus at least ten times. The overview of the term content of this data set together with the occurrence statistics is summarized in Table 1. For this data set, the two similarity domains are represented by their respective similarity matrices: feature-based and semantic. The former matrix, K, has been constructed by computing all pairwise similarities among images based on their tiled HSV color correlogram content [11], while the latter, S, has been computed with the use of pairwise relatedness measure among image annotation keywords using the method developed by P. Resnik [18]. The actual similarity values in both S and K were mapped into the target range of $[0, 1]$ by applying a generic Gaussian distance substitution kernel [10]. Once computed, S and K were kept fixed throughout all of the experiments.

The data were subdivided into the training and testing parts in such a way that for every unique query term $\{T_q | q \in 1 \dots n\}$ there were up to ten training images, while the remaining ones were used as a test set. The choice to use a relatively small number of training images was motivated by the intention to make the experiment configuration resemble a typical query-by-example scenario, where the user selects only a few relevant images to formulate a query. The experiments were to be set up so as to make it possible to evaluate the proposed dual diffusion spreading activation method together with a features-only baseline technique, and compare their ability to retrieve relevant images for each individual query term T_q. In order to do so, each single-word query was represented by a set of images Q from the training set that had been

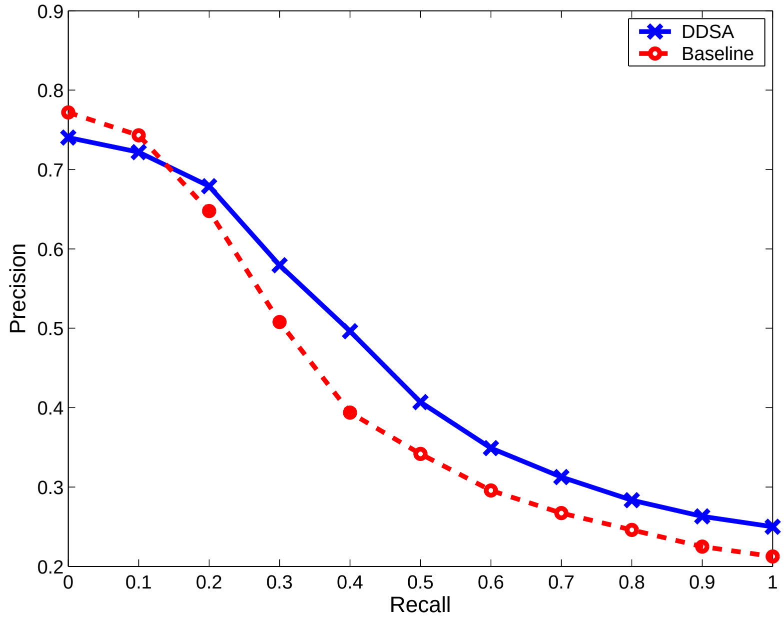

Figure 3: An 11-point average precision diagram of retrieval performance of the proposed method (solid line) and the baseline approach (dashed line).

annotated with T_q. For both methods, the performance was measured in terms of non-interpolated average precision [2], which is the average of precision values P_i for each of the testing set images relevant for a given query. In the case of the proposed method, the P_i quantities were calculated from the ranking produced by sorting the test images by the final activation values a_{iq} corresponding to the query term T_q. These activation values, in turn, were obtained by setting up an initial activation matrix as described in Section 2.2, and running the diffusion process until completion. As for the baseline method, the P_i were derived from the ranking determined by sorting the test images according to the their best feature-based similarity to any image from query set Q, i.e. $\max_{z \in Q} k(D_z, D_i^{test})$ for every test image D_i^{test}, without taking into account the semantic information.

The results of these experiments, shown graphically in terms of 11-point average precision in Figure 3, render the non-interpolated average precision of 21.25% for the baseline and 24.99% for the proposed method after averaging over all of the query terms. While this improvement is an appreciable outcome in itself, the real differences between the two compared methods may only be seen clearly by examining side by side query terms and the actual terms from the annotation of the top ranked test images, as shown in Table 2. In other words, this table gives examples of query terms together with the terms a particular query managed to elicit as the annotation associated with the top ranked retrieved images of the test data set. As Table 2 demonstrates, the proposed DDSA method retrieves images that are more semantically coherent with the query. For instance, query term *bed* retrieves images annotated with *crib* and *doll* in the proposed DDSA method, whereas the baseline approach that ignores the high-level semantic information places images annotated with *rocks* and *street* onto the top ranked list.

Because of the significant differences in the number and quality of related terms retrieved by the proposed and baseline approaches, we performed additional experiments with slight modifications of the assessment criteria. Namely, in order to account for the contribution of the related terms, we introduced a non-interpolated average semantic precision, that would average the precision values of the top fifty images weigthed by their semantic similarity to the query term. Similarly, we define semantic recall as the total weight of semantic similarity of test images to the query terms, which accounts for the test images that had been annotated with terms either exactly the same as, or semantically similar to that of the query. The results of these experiments render the non-interpolated average semantic precision of 29.66% for the baseline and 37.62% for the proposed DDSA method after averaging over all of the query terms. As one may expect, there is a more significant overall improvement, since the contribution of the sematically re-

Table 2: Examples of the true annotation terms of the top ranked test images for the proposed method (DDSA) and the baseline

Query term	*Baseline*	*DDSA*
bed	bed, rocks, street	bed, crib, doll
balloons	balloons	balloons, balloon, doll
bike	bike, rocks, waterfall, trees, ocean, street	bike, tricycle, car, sailboat
flowers	trees, deck	flowers, trees, vineyards, fields
pond	pond, trees, deck, rocks	pond, lake, lagoon, ocean, river

Table 3: Summary of the experimental evaluation

Performance measure	*1-class SVM*	*SVDD*	*Baseline*	*DDSA*
Precision	15.44%	15.50%	29.66%	37.62%
Recall	10.77%	10.93%	18.53%	24.19%

lated terms has now been taken into account. In addition to that, we observe that the overall semantic recall measure increases as well. For the sake of providing a more complete evaluation, we also tested the methods of one-class SVM [4, 21] and support vector data description [24, 25] (SVDD) applied directly on matrix K, even though these techniques could be considered to be beyond the scope of this paper. Indeed, these methods are primarily designed for classification tasks instead of ranking[1] , and make no attempt to take advantage of the semantic information. This fact helps explain their unremarkable performance on the prblem at hand, which appears to be in accordance with previously reported results on the use of one-class approaches in content-based image processing [9]. The summary of these experiments is given in Table 3.

In addition to the above evaluation, we have also investigated the performance of the proposed method with respect to its dependence on the choice of the parameter τ that specifies the termination time of the spreading activation process. The results of these experiments, shown in Figure 4, have confirmed the validity of our conjecture to estimate τ as the shortest time to near-depletion of a single activation source, as per (11). As expected, Figure 4 shows that setting τ either too low or too high reduces the performance gain of the proposed method in comparison with baseline. However, we also observe that the range of the target residual activation level γ suggested in section 2.3 does indeed correspond to the region of the best performance improvement. Finally, we also have found that the proposed method outperforms the baseline even when a diffusion process operates only on the feature-based similarities, albeit by a small margin, which positions this contribution as a comparable alternative to the traditional feature-based retrieval.

4 Conclusion and future perspectives

We have introduced a content-based information retrieval method inspired by the ideas of spreading activation models. In response to a given query, the proposed approach computes document ranks as their final activation values obtained upon completion of a diffusion process. This diffusion process is dual in the sense that it models the spreading of the query's initial activation simultaneously in two similarity domains: low-level feature-based and high-level semantic. The formulation of the diffusion process relies on an approximation that makes it possible to compute the final activation efficiently as a solution to a linear sys-

[1]We have adapted the implementation of one-class SVM and SVDD classifiers for the ranking task by sorting the documents according to the raw output of the classifier decision function trained with the default parameters. Other popular classification techniques, such as two-class SVM, were deemed not applicable due to their requirement to have both positive and negative data instances for training.

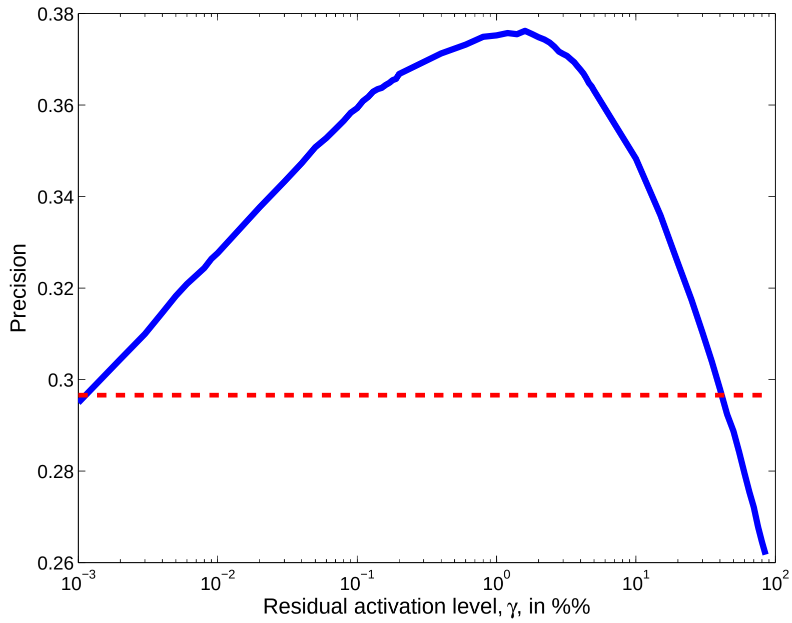

Figure 4: Solid line shows the dependence of the non-interpolated average semantic precision of the proposed method on the residual percentage level γ used to compute the termination time τ of the spreading activation. The dashed line shows the performance of the baseline feature-based retrieval for comparison.

tem of differential equations via a matrix exponential, without the need to resort to an iterative simulation. The empirical performance of the described dual diffusion model has been evaluated in terms of standard non-interpolated average precision and recall measures, together with their semantics-aware versions, on the task of content-based digital image retrieval. The obtained experimental results demonstrate that the proposed method achieves better overall performance compared to traditional feature-based approaches.

As a part of the further development of the proposed method, we have already begun investigating its extensions to multi-ple-stage usage scenarios with relevance feedback, modeling of the activation to include both positive and negative examples, combinations of several terms in a query, equivalent kernel formulations, and several other closely related aspects of the method described here.

References

[1] ASWATH, D., AHMED, S. T., D'CUNHA, J., AND DAVULCU, H. Boosting item keyword search with spreading activation. In *Web Intelligence* (2005), pp. 704–707.

[2] BAEZA-YATES, R., AND RIBEIRO-NETO, B. *Modern Information Retrieval*. Addison Wesley, 1999.

[3] BOUGUET, J.-Y., DULONG, C., KOZINTSEV, I., AND WU, Y. Requirements for benchmarking personal image retrieval systems. In *Proceedings of SPIE Photonics West, Electronic Imaging 2006* (2006), S. Santini, R. Schettini, and T. Gevers, Eds., vol. 6061, SPIE.

[4] CHANG, C.-C., AND LIN, C.-J. *LIBSVM: a library for support vector machines*, 2001. Software available at http://www.csie.ntu.edu.tw/˜ cjlin/libsvm.

[5] CRESTANI, F. Application of spreading activation techniques in information retrieval. *Artificial Intelligence Review 11*, 6 (1997), 453–582.

[6] CRESTANI, F., AND LEE, P. L. Searching the web by constrained spreading activation. *Information Processing and Management 36*, 4 (2000), 585–605.

[7] CRISTIANINI, N., SHAWE-TAYLOR, J., AND LODHI, H. Latent semantic kernels. *J. Intell. Inf. Syst. 18*, 2-3 (2002), 127–152.

[8] DEERWESTER, S., DUMAIS, S., FURNAS, G., LANDAUER, T., AND HARSHMAN, R. Indexing by latent semantic analisys. *Journal of the American Society of Information Science 41*, 6 (1990), 391–407.

[9] GOH, K.-S., LI, B., AND CHANG, E. Y. Using one-class and two-class SVMs for multiclass image annotation. *IEEE Transactions on Knowledge and Data Engineering 17*, 10 (2005), 1333–1346.

[10] HAASDONK, B., AND BAHLMANN, C. Learning with distance substitution kernels. In *26th Pattern Recognition Symposium of the German Association for Pattern Recognition (DAGM 2004)* (Tübingen, Germany, 2004), Springer Verlag.

[11] HUANG, J., KUMAR, S. R., MITRA, M., ZHU, W.-J., AND ZABIH, R. Image indexing using color correlograms. In *CVPR '97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)* (Washington, DC, USA, 1997), IEEE Computer Society, p. 762.

[12] KANDOLA, J., SHAWE-TAYLOR, J., AND CRISTIANINI, N. Learning semantic similarity. In *Advances in Neural Information Processing Systems 15*, S. T. S. Becker and K. Obermayer, Eds. MIT Press, Cambridge, MA, 2003, pp. 657–664.

[13] KONDOR, R. I., AND LAFFERTY, J. D. Diffusion kernels on graphs and other discrete input spaces. In *ICML '02: Proceedings of the Nineteenth International Conference on Machine Learning* (San Francisco, CA, USA, 2002), Morgan Kaufmann Publishers Inc., pp. 315–322.

[14] MAGENNIS, M., AND VAN RIJSBERGEN, C. J. The potential and actual effectiveness of interactive query expansion. In *SIGIR '97: Proceedings of the 20th annual international ACM SIGIR conference on Research and development in information retrieval* (New York, NY, USA, 1997), ACM Press, pp. 324–332.

[15] MILLER, G. A. Wordnet: a lexical database for English. *Commun. ACM 38*, 11 (1995), 39–41.

[16] MUNSON, B. R., YOUNG, D. F., AND OKIISHI, T. H. *Fundamentals of Fluid Mechanics*. John Wiley & Sons, 1990.

[17] PREECE, S. E. *A Spreading Activation Network Model for Information Retrieval*. PhD thesis, University of Illinois at Urbana-Champaign, 1981.

[18] RESNIK, P. Using information content to evaluate semantic similarity in a taxonomy. In *International Joint Conference for Artificial Intelligence (IJCAI-95)* (1995), pp. 448–453.

[19] RUMELHART, D. E., AND NORMAN, D. A. Representation in memory. In *Steven's handbook of experimental psychology*, vol. 2. Wiley, 1988, pp. 511–587.

[20] SALTON, G. *Automatic Information Organization and Retrieval*. McGraw Hill Text, 1968.

[21] SCHÖLKOPF, B., PLATT, J. C., SHAWE-TAYLOR, J. C., SMOLA, A. J., AND WILLIAMSON, R. C. Estimating the support of a high-dimensional distribution. *Neural Comput. 13*, 7 (2001), 1443–1471.

[22] SIDJE, R. B. Expokit: a software package for computing matrix exponentials. *ACM Transactions on Mathematical Software 24*, 1 (1998), 130–156.

[23] SIOLAS, G., AND D'ALCHÉ-BUC, F. Support vector machines based on a semantic kernel for text categorization. *IJCNN 05* (2000), 5205.

[24] TAX, D. M. J. Ddtools, the data description toolbox for matlab, June 2006. version 1.5.3.

[25] TAX, D. M. J., AND DUIN, R. P. W. Support vector data description. *Mach. Learn. 54*, 1 (2004), 45–66.