Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
The Journal of Systems and Software 73 (2004) 455–466
The Journal of Systems and Software
www.elsevier.com/locate/jss

Retrieve images by understanding semantic links and clustering image fragments ☆

Hai Zhuge *

Knowledge Grid Research Group, Key Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, P.O. Box 2704, Beijing 100080, China

Received 16 February 2003; received in revised form 25 August 2003; accepted 27 August 2003
Available online 25 December 2003

Abstract

The main obstacle to realize real semantic-based image retrieval is that semantic description of versatile images is difficult. The basic ideas of this paper are that the semantics of an object can be refined through top–down orthogonal semantic classification and that the semantics of an object can be reflected by the semantics of relevant objects and the semantic relationships between them. To reflect the semantic relationships between images, we propose a set of primitive semantic links as the enhancement of the hyperlinks connecting Web pages. The semantic link network is the natural extension of the hyperlink network so it can inherit the existing theory and method on hyperlink network. Based on the single semantic image established upon the orthogonal semantic space and the semantic link space, the proposed image retrieval approach enables users to obtain the semantic clustering of relevant images rather than a list of isolated images as the output of the current search engine and to browse images along semantic paths with the support of semantic link reasoning. Semantic matching and reasoning for realizing intelligent semantic image retrieval is based on graph operation and matrix operation.

Keywords: Information Grid; Image retrieval; Knowledge Grid; Semantic link; Semantic Web; Web

1. Introduction

Finding effective approaches to retrieve images distributed on the Web has attracted many researchers (Gevers and Smeulders, 1999; Gupta and Jain, 1997; Harmandas et al., 1997; Lempel and Soffer, 2001). There are three main categories of Web image retrieval approaches: the text-based approaches, the content-based approaches, and the hyperlink-based approaches.

The text-based approaches apply the text-based information retrieval algorithms to the annotated images including keywords, caption of image, text surrounding image, entire text of containing page, and filenames. The text-based retrieval approaches support certain natural language or topic-descriptive queries. The content-based image retrieval approaches apply image analysis techniques to extract visual features from images (Cascia et al., 1998; Kanth et al., 1998). The features are extracted at the preprocessing stage and stored in the retrieval system's database. The extracted features are usually of high dimensionality and need some dimension reduction to allow scalability of these systems. As a kind of content-based approach, the relevance feedback approaches collect users' interests and preferences to refine retrieval result (Zhang et al., 2003). The hyperlink-based approaches make use of the link structure to retrieve relevant images (Lempel and Moran, 2000). Their common basic premise is that a Web page P displays or links to an image when the author of P considers the image to be of value to the viewers of the page.

Although these previous image retrieval approaches have achieved certain progress through assistant approaches like establishing user models, to find effective

☆ This work was supported by the National Science Foundation and the National Basic Research Plan of China.
* Tel.: +86-1062565533; fax: +86-1062567724.
E-mail address: zhuge@ict.ac.cn (H. Zhuge).

0164-1212/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.jss.2003.08.243

approach to describe the semantics of images is still a key issue that remains unsolved in realizing real semantic-based image retrieval. The semantic description of image itself is difficult, but the semantics of an image can be reflected by relevant images and the semantic relationships between them.

This paper proposes a set of primitive semantic links and relevant rules as the basis of describing and deriving semantic relationships. The semantic links can construct a kind of semantic link network that is the semantic extension of the current hyperlink network. Different from the approaches in Semantic Web area (http://www.semanticweb.org), the semantic links are the natural extension of the hyperlink, so the semantic link network can inherit the existing theories and approaches of the hyperlink network (Adamic and Huberman, 2000; Henzinger et al., 1999; Kleinberg, 1999). The semantic reasoning and semantic matching used in reasoning and retrieval can be realized through graph operations and matrix representation and operations, which provide the semantic basis for realizing intelligent semantic image retrieval. The proposed approach enables users to obtain the retrieval result that is a semantic clustering of relevant images rather than a list of isolated images as the output of the current search engine and also enables users to browse images by wandering along semantic paths with the support of semantic link reasoning.

The establishment of a connective semantic link network of images is the basis of the proposed approach just as the hyperlink network to the hyperlink-based approaches. There are two ways to generate semantic links: (1) create by human with the help of assistant tools; and, (2) create by automatic mechanism, which is based on analyzing the semantic relationships between relevant descriptive terms with the help of data mining algorithms, domain ontology and other information processing technologies like the co-occurrence analysis. We will briefly introduce our research progress on how to generate semantic links in the ongoing work portion of the conclusion section.

We can imagine the scenario of the growth of future semantic link Web: it is constructed by some people at the initial stage just as the initial Web, then its resources grow rapidly with exponential expansion of users who publish resources and establish relevant semantic links. People would further seek automatic approaches for synthesizing large-scale semantic links to support intelligent behavior and investigate their distribution and rules.

We have proposed the method, model and tool for making semantic links in text resources (Zhuge, 2003), and implemented the proposed approach in the Information and Knowledge Grid platform to enhance efficiency and intelligence (Zhuge, 2002). The bottom-level representation and mutual understanding basis of the proposed approach are the well-studied ontology and the standard markup languages like XML and RDF (Berners-Lee et al., 2001; Fensel et al., 2001; Fikes and Farquhar, 1999; Hendler, 2001; Klein, 2001; Mack et al., 2001).

2. Related work

Many Web image search engines integrate the text-based retrieval approaches and the content-based image retrieval approaches. Some approaches extract visual features and associate with text in the containing pages. The images are classified into topics according to a specialized taxonomy. The others use associated text and feature extraction to support queries that include both the search topic and certain visual properties of the target image. A text-based image retrieval approach that uses the connectivity information to induce textual annotations of images has been suggested (Harmandas et al., 1997). In pure content-based image retrieval systems (Gevers and Smeulders, 1999), images are classified into photographed and synthetic images, and photographs are further classified into portraits and indoor/outdoor scenes. The approach extracts many visual features from images, and supports queries by image examples and features.

Link structure analysis is a technique for finding authoritative Web pages and it has improved the ability of search engines to rank quality pages (Dean and Henzinger, 1999; Rafiei and Mendelzon, 2000). Link structure analysis is based on the notion that a link directed from page P to page Q can be viewed as an endorsement of Q by P, and as some form of positive judgment by P of Q's content. Two important types of techniques in link structure analysis are co-citation based schemes and random walk based schemes (Henzinger et al., 1999). The main idea behind the co-citation based schemes is that P_1 and P_2 are regarded as sharing a mutual topic of interest when both P_1 and P_2 point to some page Q. Similarly, it is probable that Q_1 and Q_2 share some mutual topic when P links to both Q_1 and Q_2. Some image retrieval approaches based on the hyperlink analysis have been proposed (Chakrabarti et al., 1998; Lempel and Soffer, 2001).

Different from hyperlinks, a semantic link directed from page P to page Q is a kind of semantic judgment by P of Q's content, which can be positive or negative to some extent. For example, "P—*not-harmonious* → Q" means that P's content is not harmonious with Q's content. Furthermore, two pages P_1 and P_2 may belong to different topics of interest when P_1 and P_2 both point to page Q with different semantic links. Similarly, pages Q_1 and Q_2 may have different topics of interest when page P connects both Q_1 and Q_2 with different semantic links.

3. Semantic space of resources

3.1. Construct single semantic image for versatile resources based on orthogonal semantic space and semantic link space

An orthogonal semantic space is an n-dimensional semantic space that organizes resources according to the orthogonal classification semantics of resources. Each point in the space determines a set of resources belonging to the same semantic category. We have proposed three normal forms to normalize the orthogonal semantic space so as to efficiently locate resources (Zhuge, 2002; Zhuge, in press).

An orthogonal semantic space can be generally defined as an n-dimensional distance space denoted as $(\{X_1, X_2, \ldots, X_n\}, \rho)$, where X_i is axis and ρ is distance. A coordinate represents a semantic category on axis, represented as a graph $\langle G, E\rangle$, and for any two nodes n, $n' \in G$, if $\langle n, n'\rangle \in E$ then n is a sub-category of n'. If n and n' belong to the same category, the distance between them, $d = (n, n')$, is the shortest path between n and n' on G. If n and n' do not belong to the same category then $d(n, n') = \infty$. For any two points $m(x_1, \ldots, x_n)$ and $m'(x'_1, \ldots, x'_n)$ in the orthogonal semantic spaces, the distance between them is defined as follows:

$$\rho(m, m') = \left(\sum_{i=1}^{n} d^2(x_i, x'_i)\right)^{1/2},$$

where $d(x_i, x'_i)$ is the projection of ρ on one dimension. We can prove that the n-dimensional semantic space satisfies: (1) $\rho(m_1, m_2) = 0$ if and only if $m_1 = m_2$; (2) $\rho(m_1, m_2) = \rho(m_2, m_1)$; and, (3) $\rho(m_1, m_3) \leqslant \rho(m_1, m_2) + \rho(m_2, m_3)$, where m_i is a point in the orthogonal semantic space.

A semantic link space consists of a set of resources and the semantic relationships defined on the resource set. The semantic relationship is represented as directed links associated with semantic factors. The semantic links connect resources to form semantic link networks. Overlooking the semantic link space, it consists of high-level *semantic clusters*, within which the semantic links are much more denser than those between them. Different view angles may see different semantic clusters. People's retrieval destination usually focuses on one semantic cluster. The semantic link space is also a kind of distance space represented as ⟨Resources, SemanticLinks⟩. The distance between any two nodes is a kind of semantic distance determined by the semantic factors and the shortest semantic path between two nodes (i.e., the number of semantic links between them). A semantic path is an interconnected semantic link chain in form of P_1—$\alpha_1 \rightarrow P_2, P_2$—$\alpha_2 \rightarrow P_3, \ldots, P_{n-1}$—$\alpha_n \rightarrow P_n$, where $\alpha_i = \alpha_{i+1}$ or $\alpha_i \Rightarrow \alpha_{i+1}$ (i.e., α_i implies α_{i+1}).

Definition 1 (*Semantic connectivity*). A set of resources is called *weak semantic connective* if these resources share at least one semantic dimension in orthogonal semantic space. A set of resources is called *strong semantic connective* if (1) they belong to the same semantic category (i.e., there exists a coordinate that semantically includes these resources), or (2) for any two resources, there exists a semantic path in the semantic link space.

A *single semantic image* can be constructed for versatile resource entities by establishing semantic mapping between resources specified in the orthogonal semantic space and the semantic link space (Zhuge, 2002). Resources can be uniformly understood and operated at this single semantic image level. For example, we can use $P(C_1, C_2, C_3)$—$\alpha_1 \rightarrow Q(C'_1, C'_2, C'_3)$ to express both orthogonal classification semantics and semantic link semantics, where coordinates (C_1, C_2, C_3) and (C'_1, C'_2, C'_3) accurately locate the categories of P and Q in a three-dimensional orthogonal semantic space. We can design an SQL-like language to support resource operation in the single semantic image. For example, the following two statements express the retrieval requirements using both the orthogonal semantic space and the semantic link space.

SELECT R FROM RS(X_1, X_2, X_3) WHERE R—$\alpha \rightarrow P(C_1, C_2, C_3)$.
SELECT R FROM RS(X_1, X_2, X_3) WHERE $R(C'_1, C'_2, C'_3)$—$\alpha \rightarrow P(C_1, C_2, C_3)$.

The "FROM" portion "RS(X_1, X_2, X_3)" specifies the orthogonal semantic space that the resource R belongs to is a three-dimensional semantic space (X_1, X_2 and X_3 are its three axes). The "WHERE" portion refines the search scope by specifying the semantic relationship between R and resource P belonging to the point (C_1, C_2, C_3). The second statement further refines the search scope by specifying the category that R belongs to.

The general architecture of the proposed approach is described in Fig. 1. The semantic Web including

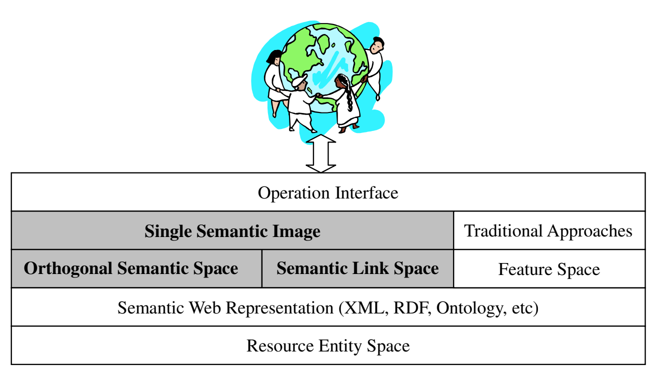


Fig. 1. General architecture of the proposed approach.

ontology is the representation and understanding basis of the orthogonal semantic space, semantic link space, and feature space. The following part of this paper mainly discusses the semantic link space.

3.2. Semantic description of image

The semantics of an image can be reflected by three facets as described in the following frame: (1) *feature* and *content*, where "Lattice-Meridian" is used for accurate map positioning, "Layer" is used to differentiate multiple feature layers on the same image because an image can be the overlay of multiple additive feature layers. "Topic" reflects general content of the image and limits its semantic region using a topic hierarchy. "File-address" records the address of the image file, and "Data-description" describes the data that explains the image content such as the geographical information (we do not discuss in detail herein). (2) *Semantic category*, which is reflected by its semantic coordinates in the orthogonal semantic space that organizes images in normal forms. (3) *Inter-image-semantics*, which is reflected by the semantic relationships between images such as the meaning-semantic-links, position-semantic-links, and existence-semantic-links. These links are represented as a pointer with a type directed from one resource or fragment (predecessor) to the other (successor). These semantic links connect relevant images to construct semantic link image networks. We will introduce them in detail in the following parts of this paper.

Image description:
  Feature and content:
    Lattice-Meridian: $((a_1, b_2), (c_1, d_2))$
    Layer: Layer-type
    Topic: Topic-hierarchy or Topic-description-text
    File-address: URL
    Data-description: XML
  Semantic-coordinate: two-dimensional coordinate (x, y) or three-dimensional coordinate (x, y, z), etc.
  Inter-image-semantics:
    Meaning-semantic-links
    Position-semantic-links
    Existence-semantic-links

3.3. Meaning semantic link

A *meaning semantic link* reflects the meaning relationship between two resources in the form of $P—\alpha \rightarrow Q$, where α is a semantic factor (Zhuge, 2003). The meaning semantic link includes the following types:

(1) *Cause-effect link*, denoted as $P—ce \rightarrow Q$, which defines that the predecessor P is the cause of its successor Q, and Q is the effect of P. The cause-effect link has transitive characteristic, i.e., $O—ce \rightarrow P$, $P—ce \rightarrow Q \Rightarrow O—ce \rightarrow Q$ holds. According to the transitive characteristic, the cause-effect reasoning can be formed through chaining cause-effect links.
(2) *Implication link*, denoted as $P—imp \rightarrow Q$, which defines that the semantics of the predecessor P implies that of its successor Q. The implication link has the transitive characteristic, i.e., $O—imp \rightarrow P$, $P—imp \rightarrow Q \Rightarrow O—imp \rightarrow Q$ holds. The implication link can help the reasoning mechanism find the semantic implication relationship between images.
(3) *Subtype link*, denoted as $O—st \rightarrow P$, which defines that the successor P is a part of its predecessor O. The subtype link has the transitive characteristic, i.e., $O—st \rightarrow P$, $P—st \rightarrow Q \Rightarrow O—st \rightarrow Q$ holds.
(4) *Similar-to link*, denoted as $O—(sim, sd) \rightarrow P$, which defines that the semantics of the successor P is similar to that of its predecessor O, where sd is the similarity degree between O and P. The similar-to link is not transitive because different similar-to relationship may concern different facets.
(5) *Instance link*, denoted as $O—ins \rightarrow P$, which defines that the successor is the instance of the predecessor.
(6) *Sequential link*, denoted as $O—seq \rightarrow P$, which defines that the predecessor O should be browsed or visited before its successor P, i.e., the content of P is the successor of the content of O. The sequential link has the transitive characteristic, i.e., $O—seq \rightarrow P$, $P—seq \rightarrow Q \Rightarrow O—seq \rightarrow Q$ holds. The transitive characteristic enables relevant sequential links to be connected to form a sequential chain.
(7) *Reference link*, denoted as $O—ref \rightarrow P$, which defines that the successor P is the further explanation of its predecessor O. The reference link has transitive characteristic, i.e., $O—ref \rightarrow P$, $P—ref \rightarrow Q \Rightarrow O—ref \rightarrow Q$ holds.
(8) *Equal-to link*, denoted as $P—equ \rightarrow Q$, which means that P and Q are the same. The equal-to link can be regarded as the special case of the similar-to link.
(9) *Undefined link*, denoted as $P—und \rightarrow Q$, which means that the relationship between P and Q is unidentified, but it may be identified during reasoning.

For each semantic link, we can define a reverse semantic link to represent the reverse semantic relationship.

3.4. Position semantic link

Position semantic relationship exists between image pieces that can constitute a complete image. They can be expressed by the following 10 types of semantic links:

(1) O—*above* $\rightarrow P$ represents the position semantic relationship that O is above P;
(2) O—*below* $\rightarrow P$ represents the position semantic relationship that O is below P;
(3) O—*left* $\rightarrow P$ represents the position semantic relationship that O is on the left of P;
(4) O—*right* $\rightarrow P$ represents the position semantic relationship that O is on the right of P;
(5) O—*east* $\rightarrow P$ represents the position semantic relationship that O is to the east of P;
(6) O—*south* $\rightarrow P$ represents the position semantic relationship that O is to the south of P;
(7) O—*west* $\rightarrow P$ represents the position semantic relationship that O is to the west of P;
(8) O—*north* $\rightarrow P$ represents the position semantic relationship that O is to the north of P;
(9) O—*front* $\rightarrow P$ represents the position semantic relationship that O is in front of P;
(10) O—*behind* $\rightarrow P$ represents the position semantic relationship that O is behind P.

The position semantic links can be abstracted as O—$\beta \rightarrow P$, where semantic factor $\beta \in \{above, under, left, right, north, south, east, west, front, behind\}$. The following semantic factor pairs: "*up*" and "*below*", "*left*" and "*right*", "*south*" and "*north*", "*east*" and "*west*", "*front*" and "*behind*" are mutual symmetric semantic factors. Position semantic links satisfy the following transitive characteristic.

Characteristic: O—$\beta \rightarrow P$ and P—$\beta \rightarrow Q \Rightarrow O$—$\beta \rightarrow Q$.

More meaningful position semantic links can be formed by composing different semantic factors. For example, we can obtain O—*north-west* $\rightarrow P$ according to O—*north* $\rightarrow P$ and O—*west* $\rightarrow P$. Similarly we can form the following meaningful position semantic links through composition: O—*south-east* $\rightarrow P$; O—*south-west* $\rightarrow P$; and, O—*north-east* $\rightarrow P$. Generally, the composition of two semantic links O—$\alpha \rightarrow P$ and O—$\beta \rightarrow P$ generates a new semantic link: O—α-$\beta \rightarrow P$ according to suitable values of α and β.

4. Rules and reasoning

4.1. Rules for meaning semantic links

Rules can be derived for chaining relevant meaning semantic links and obtaining meaningful reasoning result. For example, if we have two links: O—$ce \rightarrow P$ and P—$ce \rightarrow Q$, then we can obtain the reasoning result: O—$ce \rightarrow Q$ according to the transitive characteristic of the cause-effect link. The reasoning process can be represented as a reasoning rule: O—$ce \rightarrow P$, P—$ce \rightarrow Q \Rightarrow O$—$ce \rightarrow Q$. A reasoning rule can also be simply represented as $\alpha \cdot \beta \Rightarrow \gamma$, where α, β, $\gamma \in \{ce, imp, ins, st, sim, ref, seq, equ\}$, for example, above reasoning rule can be represented as $ce \cdot ce \Rightarrow ce$.

A simple case of the reasoning is that all the semantic links have the same type (called single-type reasoning). According to the transitive characteristic of the semantic links, we have the following reasoning rule: P_1—$\alpha \rightarrow P_2, P_2$—$\alpha \rightarrow P_3, \ldots, P_{n-1}$—$\alpha \rightarrow P_n \Rightarrow P_1$—$\alpha \rightarrow P_n$, where $\alpha \in \{ce, imp, st, seq, ref, equ\}$.

Heuristic rules for connecting different types of links are shown in Table 1. Rules 1–5 are for the connection between the cause-effect link and the other links. Rules 6–9 are for the connection between the implication link and the other links. Rules 10–13 are for the connection between the sub-type link and the other links. Rules 14–16 are for the connection between the instance link and the other links. Rules 17–24 show the fact that the sequential connection satisfies additivity. These rules can be proved after formally defining the semantic links. To avoid complex formal statements, the reasoning rules in Table 1 are used as heuristic rules.

An order relationship exists between these semantic links: $equ \leqslant ref \leqslant ins \leqslant st \leqslant imp \leqslant ce$, where the right one reflects stronger semantic relationship between two documents than the left one does. In order to obtain a good reasoning result, the reasoning mechanism should find the strongest link between the candidate links. Summarizing the rules in Table 1, we have the following characteristic.

Characteristic 1. For a given connection of two semantic links: $\alpha \cdot \beta$ if $\beta \leqslant \alpha$ then we have $\alpha \cdot \beta \Rightarrow \alpha$ hold.

The meaning semantic links can also be inexact. An inexact semantic link reflects the possibility of its existence. We herein use a certainty degree cd to reflect such a possibility. Therefore, an inexact semantic link can be represented as: P—$(\alpha, cd) \rightarrow P'$, where $\alpha \in \{ce, imp, st, sim, ins, seq, ref, equ\}$. An inexact single-type semantic link reasoning takes the following form: P_1—$(\alpha, cd_1) \rightarrow P_2, P_2$—$(\alpha, cd_2) \rightarrow P_3, \ldots, P_n$—$(\alpha, cd_n) \rightarrow P_{n+1} \Rightarrow P_1$—$(\alpha, cd) \rightarrow P_{n+1}$, where $cd = f(cd_1, \ldots, cd_n)$, and minimum "*Min*" is one choice of function f. Different semantic factors determine different computing function.

Different types of inexact semantic links can also be chained for reasoning according to the rules in Table 1. For example, rule 1 can be extended as the following inexact rule: P—$(ce, cd_1) \rightarrow P'$, P'—$(imp, cd_2) \rightarrow P'' \Rightarrow P$—$(ce, f(cd_1, cd_2)) \rightarrow P''$. More inexact rules can be similarly formed. Another kind of inexactness is caused by the "similar-to" link, for example, connecting the cause-effect link with the similar-to link can obtain the following inexact reasoning rules: P—$ce \rightarrow P'$, P'—$(sim, cd) \rightarrow P'' \Rightarrow P$—$(ce, cd') \rightarrow P''$, where cd' depends on cd ($cd' = cd$ is a simple choice).

Table 1
Reasoning rules for meaning semantic links

No.	Rules	Summarization
Rule 1	P—$ce \to P'$, P'—$imp \to P'' \Rightarrow P$—$ce \to P''$	$ce \cdot \beta \Rightarrow ce$
Rule 2	P—$ce \to P'$, P'—$st \to P'' \Rightarrow P$—$ce \to P''$	$ce \cdot \beta \Rightarrow ce$
Rule 3	P—$ce \to P'$, P'—$sim \to P'' \Rightarrow P$—$ce \to P''$	$ce \cdot \beta \Rightarrow ce$
Rule 4	P—$ce \to P'$, P'—$equ \to P'' \Rightarrow P$—$ce \to P''$	$ce \cdot \beta \Rightarrow ce$
Rule 5	P—$ce \to P'$, P'—$ins \to P'' \Rightarrow P$—$ce \to P''$	$ce \cdot ins \Rightarrow ce$
Rule 6	P—$imp \to P'$, P'—$st \to P'' \Rightarrow P$—$imp \to P''$	$imp \cdot st \Rightarrow imp$
Rule 7	P—$imp \to P'$, P'—$ins \to P'' \Rightarrow P$—$ins \to P''$	$imp \cdot ins \Rightarrow ins$
Rule 8	P—$imp \to P'$, P'—$ce \to P'' \Rightarrow P$—$ce \to P''$	$imp \cdot ce \Rightarrow ce$
Rule 9	P—$imp \to P'$, P'—$ref \to P'' \Rightarrow P$—$ref \to P''$	$imp \cdot ref \Rightarrow ref$
Rule 10	P—$st \to P'$, P'—$ce \to P'' \Rightarrow P$—$ce \to P''$	$st \cdot ce \Rightarrow ce$
Rule 11	P—$st \to P'$, P'—$imp \to P'' \Rightarrow P$—$imp \to P''$	$st \cdot imp \Rightarrow imp$
Rule 12	P—$st \to P'$, P'—$ref \to P'' \Rightarrow P$—$ref \to P''$	$st \cdot ref \Rightarrow ref$
Rule 13	P—$st \to P'$, P'—$ins \to P'' \Rightarrow P$—$ins \to P''$	$st \cdot ins \Rightarrow ins$
Rule 14	P—$ins \to P'$, P'—$ce \to P'' \Rightarrow P$—$ce \to P''$	$ins \cdot ce \Rightarrow ce$
Rule 15	P—$ins \to P'$, P'—$imp \to P'' \Rightarrow P$—$imp \to P''$	$ins \cdot imp \Rightarrow imp$
Rule 16	P—$ins \to P'$, P'—$ref \to P'' \Rightarrow P$—$ins \to P''$	$ins \cdot ref \Rightarrow ref$
Rule 17	P—$ins \to P'$, P_1—$ins \to P_1' \Rightarrow (P$—$seq \to P_1)$—$ins \to (P$—$seq \to P_1')$	P—$\beta \to P'$, P_1—$\beta \to P_1' \Rightarrow$ $(P$—$seq \to P_1)$—$\beta \to (P$—$seq \to P_1')$
Rule 18	P—$ref \to P'$, P_1—$ref \to P_1' \Rightarrow (P$—$seq \to P_1)$—$ref \to (P$—$seq \to P_1')$	Same as above
Rule 19	P—$seq \to P'$, P_1—$seq \to P_1' \Rightarrow (P$—$seq \to P_1)$—$seq \to (P$—$seq \to P_1')$	Same as above
Rule 20	P—$ce \to P'$, P_1—$ce \to P_1' \Rightarrow (P$—$seq \to P_1)$—$ce \to (P$—$seq \to P_1')$	Same as above
Rule 21	P—$imp \to P'$, P_1—$imp \to P_1' \Rightarrow (P$—$seq \to P_1)$—$imp \to (P$—$seq \to P_1')$	Same as above
Rule 22	P—$st \to P'$, P_1—$st \to P_1' \Rightarrow (P$—$seq \to P_1)st \to (P$—$seq \to P_1')$	Same as above
Rule 23	P—$sim \to P'$, P_1—$sim \to P_1' \Rightarrow (P$—$seq \to P_1)$—$sim \to (P$—$seq \to P_1')$	Same as above
Rule 24	P—$equ \to P'$, P_1—$equ \to P_1' \Rightarrow (P$—$seq \to P_1)$—$equ \to (P$—$seq \to P_1')$	Same as above

4.2. Rules for position-semantic links

Rules can be derived from the position semantic links according to the implication or symmetric semantic relationship between the semantic links. For example, O—$above \to P \Rightarrow P$—$below \to O$ means that if the position of O is above P then the position of P is below O. Table 2 shows a set of rules on position semantic links. The following two rules can be generalized from the rules in Table 2:

Generalization Rule 1: O—α_1-$\alpha_2 \to P \Rightarrow P$—$\beta_1$-$\beta_2 \to O$ holds if and only if: (1) "α_1-α_2" and "β_1-β_2" are both meaningful, (2) O—$\alpha_1 \to P \Rightarrow P$—$\beta_1 \to O$, and (3) O—$\alpha_2 \to P \Rightarrow P$—$\beta_2 \to O$ hold.

Generalization Rule 2: If α and β are mutual symmetric semantic factors, then O—$\alpha \to P \Rightarrow P$—$\beta \to O$ holds.

4.3. Semantic link network reasoning

A semantic link network (SLN) reflects a kind of semantic environment where the nodes co-exist. A node can be an image or an image fragment, a Web page, and another SLN. Image pieces can be reused in different semantic environment to compose images of different content. An SLN with n nodes can be represented as a matrix, $\mathrm{SLN} = [S_{ij}]_{n\times n}$, where S_{ij} is a set of semantic factors representing the semantic links directed from node i to node j.

Table 2
Rules for position-semantic links

No.	Rules	Category
1	O—$above \to P \Rightarrow P$—$below \to O$	O—$\alpha \to P \Rightarrow P$—$\beta \to O$
2	O—$below \to P \Rightarrow P$—$above \to O$	Same as above
3	O—$left \to P \Rightarrow P$—$right \to O$	Same as above
4	O—$right \to P \Rightarrow P$—$left \to O$	Same as above
5	O—$north \to P \Rightarrow P$—$south \to O$	Same as above
6	O—$south \to P \Rightarrow P$—$north \to O$	Same as above
7	O—$east \to P \Rightarrow P$—$west \to O$	Same as above
8	O—$west \to P \Rightarrow P$—$east \to O$	Same as above
9	O—$front \to P \Rightarrow P$—$behind \to O$	Same as above
10	O—$behind \to P \Rightarrow P$—$front \to O$	Same as above
11	O—$north\text{-}west \to P \Rightarrow P$—$south\text{-}east \to O$	O—α_1-$\alpha_2 \to P \Rightarrow P$—$\beta_1$-$\beta_2 \to O$
12	O—$south\text{-}east \to P \Rightarrow P$—$north\text{-}west \to O$	Same as above
13	O—$south\text{-}west \to P \Rightarrow P$—$north\text{-}east \to O$	Same as above
14	O—$north\text{-}east \to P \Rightarrow P$—$south\text{-}west \to O$	Same as above

SLN-based reasoning includes the following two types. (1) *Semantic link reasoning*, which carries out by chaining semantic links. If we can define the chaining of semantic links as a kind of multiply operation on semantic factors and define the addition of a semantic link between a pair of nodes as a kind of addition operation on semantic factors, then the semantic link reasoning can be realized by matrix multiply operation. (2) *Semantic graph reasoning*, which carries out according to the semantic inclusion relationship between different SLN sub-graphs. Judging the semantic inclusion relationship between two SLNs plays the key role in this type of reasoning.

5. Image retrieval and clustering

5.1. Orthogonal semantic factors

Orthogonal semantics exists between position semantic factors. We use $\alpha_1 \perp \alpha_2$ to denote "α_1 is the orthogonal semantic factor of α_2". The following are six sets of orthogonal relationships:

(1) "*below*" ⊥ "*left*", "*below*" ⊥ "*right*", "*up*" ⊥ "*left*", and "*up*" ⊥ "*right*";
(2) "*south*" ⊥ "*west*", "*south*" ⊥ "*east*", "*north*" ⊥ "*west*", and "*north*" ⊥ "*east*";
(3) "*behind*" ⊥ "*left*", "*behind*" ⊥ "*right*", "*behind*" ⊥ "*up*", "*behind*" ⊥ "*below*";
(4) "*below*" ⊥ "*behind*" ⊥ "*south*", "*behind*" ⊥ "*east*", "*behind*" ⊥ "*west*", "*behind*" ⊥ "*north*";
(5) "*below*" ⊥ "*front*" ⊥ "*south*", "*front*" ⊥ "*east*", "*front*" ⊥ "*west*", "*front*" ⊥ "*north*";
(6) "*below*" ⊥ "*front*" ⊥ "*left*", "*front*" ⊥ "*right*", "*front*" ⊥ "*up*", "*front*" ⊥ "*below*".

The following two rules can be derived from above orthogonal relationships.

Symmetry rule: If $\alpha_1 \perp \alpha_2$ then we have $\alpha_2 \perp \alpha_1$ holds.

Orthogonal rule: In a two-dimensional space, if $\alpha_1 \perp \alpha_2$, $\alpha_3 \perp \alpha_1$ and $\alpha_4 \perp \alpha_2$ then we have $\alpha_3 \perp \alpha_4$ holds.

An image P cannot be uniquely determined by just knowing a relevant image Q and a semantic link between P and Q because there may exist multiple semantic links between P and Q. The following rule is for determining the position of an image.

Position determination rule: The position of an image X can be determined in a two-dimensional space by two position semantic links $X—\alpha_1 \to O$ and $X—\alpha_2 \to P$ if and only if α_1 and α_2 have orthogonal semantics.

The orthogonal semantic factors and rules should be normally stored and managed as knowledge so as to effectively support intelligent image retrieval.

5.2. Retrieving and clustering meaningful images

Besides the meaning semantic links and position semantic links, images can also be clustered according to the following *existence semantic relationships* so as to satisfy users' retrieval requirements.

(1) O—*co-occurrence* → P, which means that O is co-occurrence with P;
(2) O—*not-co-occurrence* → P, which means that O is not co-occurrence with P;
(3) O—*harmonious* → P, which means that O is harmonious with P;
(4) O—*not-harmonious* → P, which means that O is not harmonious with P;
(5) O—*complement* → P, which means that O is the complement of P.

Definition 2. A given semantic link image network is called a *minimum SLN* if and only if: (1) there does not exist such links $O—\alpha \to P$ and $O'—\alpha \to P'$ that O is the synonym of or implied by O' and P is the synonym of or implied by P'; and, (2) there does not exist such a semantic link that is implied by or can be derived from existing links through reasoning rules.

Definition 3. A semantic link network is called a *minimum connective SLN* if it is a minimum SLN and there exists a path between any two nodes.

The minimum connective SLN enables the retrieval mechanism to consider the most important relationship during search and display. The main process of the proposed image retrieval approach includes the following seven steps:

Step 1. The user inputs requirements through retrieval interface. The retrieval mechanism obtains the topic, keywords and coordinates that describe the category of the required image.
Step 2. The retrieval mechanism locates a set of images K according to the given coordinates.
Step 3. The retrieval mechanism refines K according to the topic and keywords.
Step 4. The retrieval mechanism minimizes K by removing redundant links.
Step 5. The retrieval mechanism clusters image fragments in K according to the meaning semantic links, position semantic links, existence semantic links, content description, and features.
Step 6. The retrieval mechanism arranges relevant images considering the characteristics of Web page so that the current Web browser can display the retrieval result.
Step 7. The retrieval mechanism displays the Web page in the Web browser as the retrieval result.

The existing approaches like relevance feedback and learning approach can be incorporated into above process to achieve better retrieval effect (Zhang et al., 2003).

6. Implement Image retrieval in information grid

Our implemented Information Grid provides the following image operations: (1) *put*, to upload the image and describe its semantics and relevant semantic links, and then put it into the information space of the Information Grid by selecting proper coordinates; (2) *get*, to retrieve images from the information space of the Information Grid by selecting proper coordinates; (3) *browse*, to browse images in form of list or along semantic links; (4) *delete*, to delete an image and all the related semantic links; (5) *update*, to update images and relevant links; (6) *create*, to create an image space in the Information Grid; (7) *open*, to open the personal image space to the peers so that they can share images with each other; and, (8) *recommend*, to recommend images to the users who need them.

Fig. 2 shows the operation interface of the Information Grid. The up-portion displays the operation buttons, the left column lists the topic tree, the middle portion shows a two-dimensional semantic space where each rectangle represents a point of the space determined by respective coordinates, and the low-portion is the interface for describing image. Fig. 3 shows the retrieval interface where the low-portion expresses users' retrieval requirements such as keywords, time limitation, title, and author. The user should locate a point in the two-dimensional space to reduce the search space.

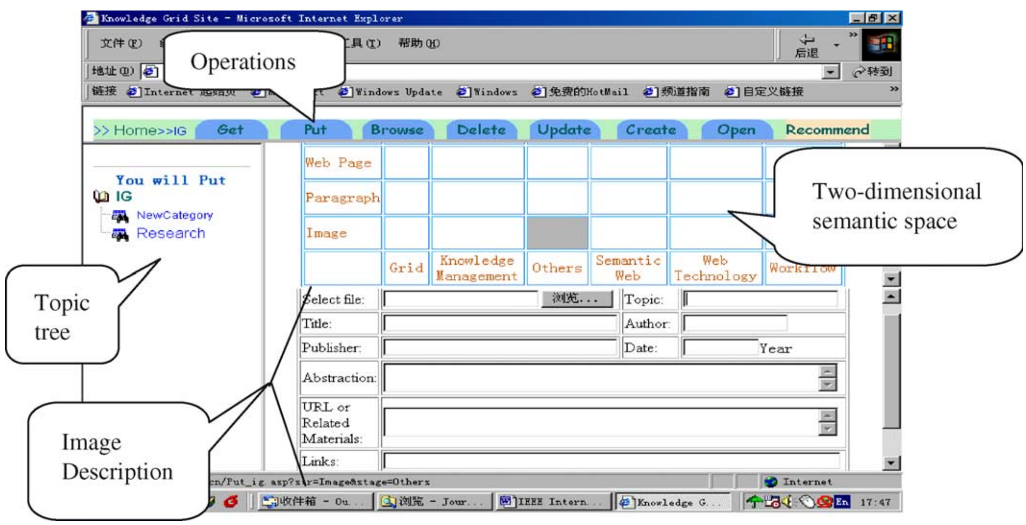


Fig. 2. Information description interface of Information Grid.

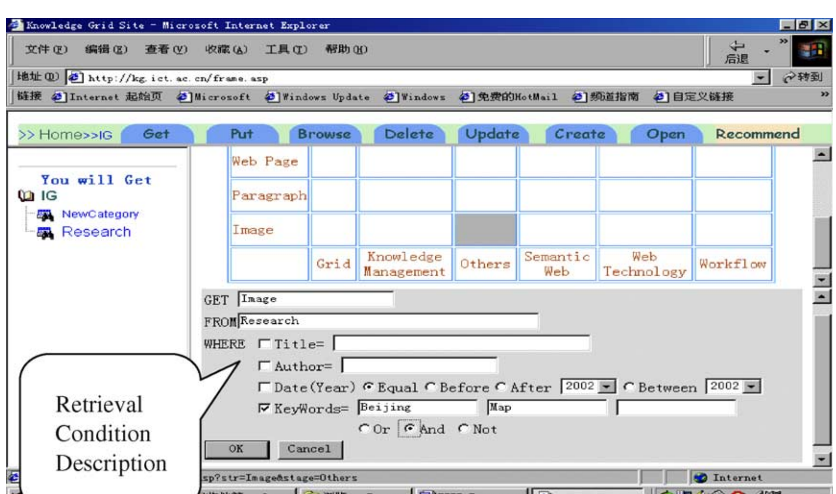
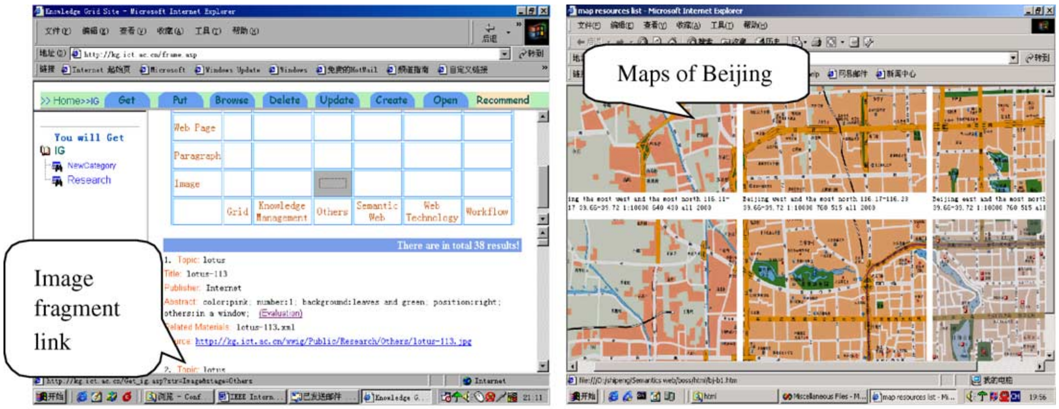


Fig. 3. Retrieval interface of Information Grid.

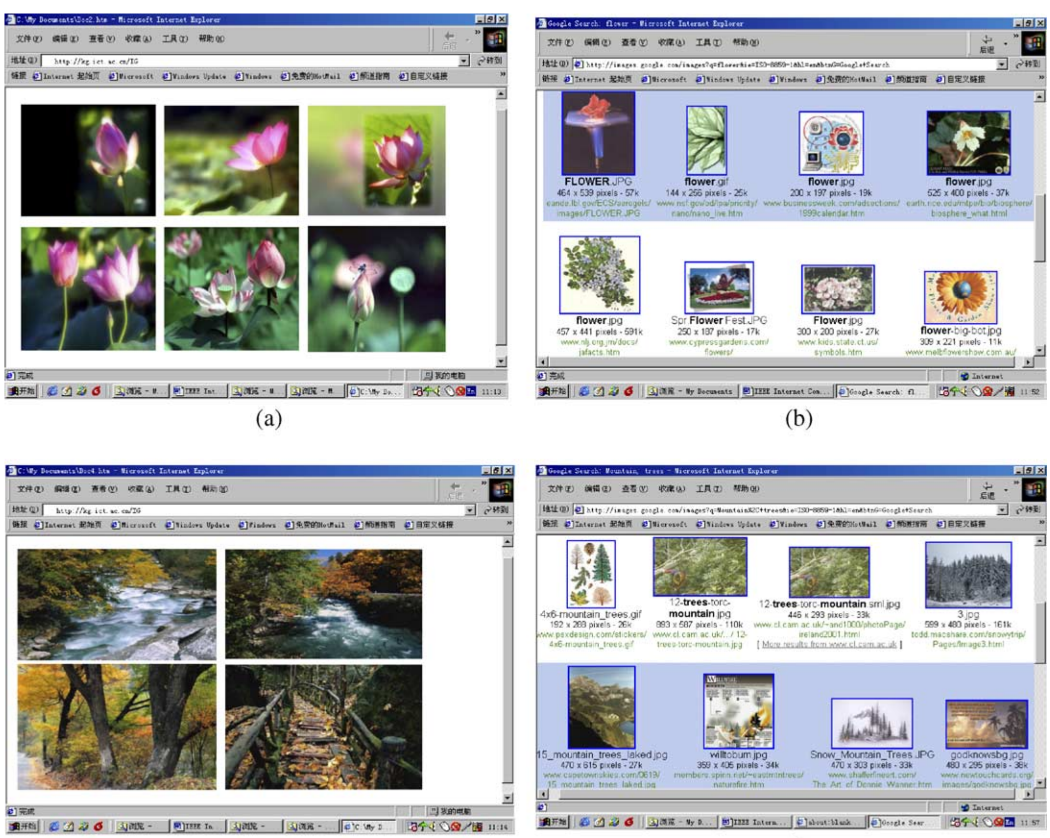


(a) (b)

Fig. 4. Two different interfaces for displaying retrieval result. (a) Retrieval result without clustering, (b) retrieval result with clustering.

Fig. 4(a) and (b) compare the result of retrieval in case of without clustering and with clustering respectively. The retrieved image items are listed separately in the left figure. Users could click the related URL or point to the link to display semantic links and then to choose to view relevant images page by page. In the right figure, the retrieved image fragments are clustered according to their semantic relationships. These image fragments may be created by different authors and stored in different resource (image) spaces.

Semantic factors can be used as keywords when retrieving images. The retrieval mechanism will analyze the relationships between input keywords, semantic links and topic hierarchies when clustering images. Fig. 5(a)



(a) (b) (c) (d)

Fig. 5. Examples of retrieval by using semantic factors as keywords. (a) Result of retrieval with keyword "flower" and the "co-occurrence" semantic factor. (b) Result of Google's retrieval with the same set of keywords as (a). (c) Result of retrieval with keyword "tree, mountain" and the "harmonious" semantic factor. (d) Result of Google's retrieval with the same set of keywords as (c).

shows the retrieval result by providing the "*co-occurrence*" semantic factor as keyword. Fig. 5(b) shows the retrieval result of Google by providing the same set of keywords as Fig. 5(a), and (c) shows the retrieval result of providing the "*harmonious*" semantic factor as keyword. Fig. 5(d) shows the retrieval result of Google by providing the same set of keywords as Fig. 5(c). Although the effectiveness of the proposed approach depends on the completeness of the description of semantic links, the above retrieval examples intuitively demonstrate the semantic link network's potentiality that the current hyperlink network does not have.

The retrieval mechanism can display relevant semantic links at the bottom of every page just as hyperlinks. Hyperlinks are also displayed as the special case of the semantic link. Users can preview the relevant next-step semantic links by pointing to the link mark and then decide which to click to browse relevant images. Users can also know the reasoning result corresponding to the link mark when pointing to it. The reasoning is based on the proposed semantic link rules.

7. Comparisons and discussion

The closely related work that has been done is the hyper-link-based approaches, which are based on the following notions: (1) images which are co-contained in pages are likely to be related to the same topic; (2) images which are contained in pages that are co-cited by a certain page are likely related to the same topic; and (3) images which are contained in authoritative pages on one topic are good candidates to be quality images on that topic. As we have mentioned in the introduction portion, these notions are no longer right in the semantic link network.

We have carried out experiments to compare the recall and precision for retrieving a given set of semantic link image networks under the same set of query conditions. Each network contains 30 image nodes. Fig. 6 shows the experimental result of comparing the change of precision and recall with the change of the number of semantic links. Fig. 7 shows that the precision and recall change with the number of types of semantic links. This experiment shows that the *retrieval efficacy depends not only on the number of semantic links but also on the types of the semantic links included in a semantic link network*. This phenomenon can be more precisely verified through carrying out large-scale experiments and adding random samples.

The underline premise of the proposed approach implies that the image retrieval efficacy depends on the providers' semantic description on given images. This is a limitation, but the proposed approach is still significant because the hyperlink-based retrieval approaches also depend on the pre-established hyperlinks, and this

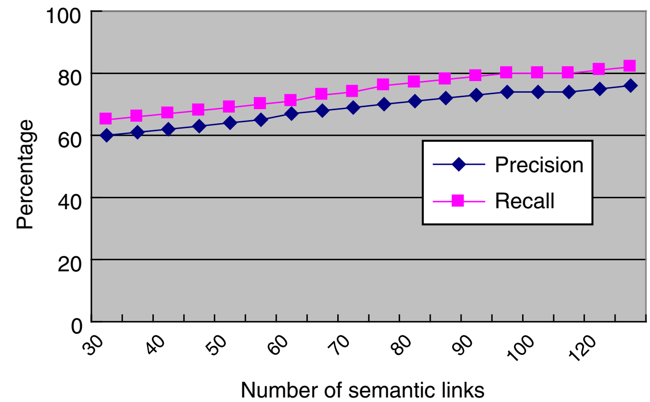


Fig. 6. Recall and precision change with the number of semantic links.

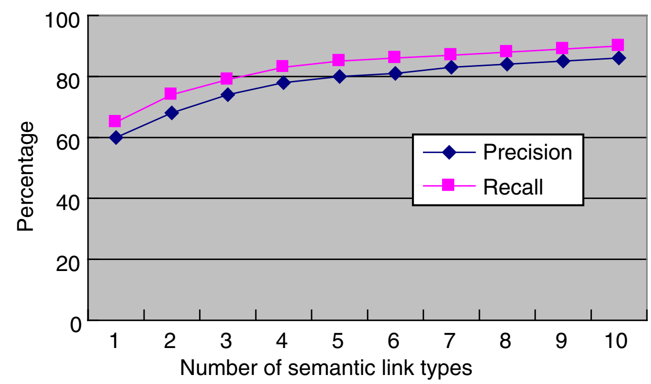


Fig. 7. Recall and precision change with the number of types of the semantic links.

limitation can be overcome to some extent by our ongoing work introduced in the conclusion section. If no semantic links are established, the proposed approach becomes traditional text-based or content-based approaches. The proposed approach does not exclude the existing approaches. Incorporating the methods used in the text-based approaches, the content-based approaches, and the hyperlink-based retrieval approaches into the proposed approach can help refine the retrieval result.

Scientists have proposed Semantic Web to overcome the shortcomings of the hyperlink-based Web (Berners-Lee et al., 2001). In the proposed approach, the semantic link space and the orthogonal semantic space constitute a single semantic image mechanism for versatile resources. The image retrieval obtains support from the orthogonal semantic space retrieval and the semantic link reasoning. The semantic link network is the natural extension of the traditional hyperlink network, so it can inherit the existing theory and method of the hyperlink network such as the link analysis approach when carrying out further research (Adamic and Huberman, 2000; Kleinberg and Lawrence, 2001; Kleinberg, 2000).

Human recognition of objective world concerns low-level semantics and high-level semantics. Low-level semantics refers to surface features like keywords, structure, color and texture, etc. High-level semantics is the meaning including content, metaphor, analogy, etc. The two semantic levels concern two different semantic spaces. Mapping and transition between them exist in human understanding and thinking process, which concerns complex psychological process and cognitive process. Current Web applications and software tools are mainly based on low-level semantics. Some researchers expect to bridge the gap between the low-level semantic space and the high-level semantic space through establishing domain ontology. However, the ontological approach neglects the epistemological aspect of human cognitive process. The approach proposed in this paper is an effort to establish the high-level semantic space and relevant computing model.

8. Conclusion

This paper proposes a new image retrieval approach based on the orthogonal semantic space and the semantic link space. The basic ideas are that the semantics of an image can be located and refined by orthogonal semantics and reflected by the semantics of relevant images and the semantic relationships between them. Multiple types of semantic links such as meaning semantic links and position semantic links are proposed to describe such inter-image semantic relationships. Semantic links connect relevant images to form semantic link network. Rules derived from these semantic links support the semantic-link-based reasoning. The retrieval result of the proposed approach can be the clustering of relevant images rather than a simple list of isolated images. Users can also browse image along semantic paths with the support of the semantic link reasoning. Previous approaches do not have these features. The semantic link network not only reflects a kind of abstract semantic relationship between entity resources but also establishes a kind of typed link relationship between entities. This is the natural extension of the hyperlink network so it can inherit the existing theory and method on hyperlink network. The efficiency of image retrieval can be improved by organizing orthogonal semantic space. Experiments show that the retrieval efficacy depends not only on the number of semantic links but also on the types of the semantic links included in a semantic link network. The orthogonal semantic space and the semantic link space establish a single semantic image mechanism for versatile resources, which supports uniform retrieval and processing in future Web environment.

Ongoing work includes the following aspects: (1) Develop approaches and tools for deriving new semantic links according to existing semantic link networks because we cannot expect all the semantic link networks made by different people to be all inter-connected. We have made some important progress on the automatic mechanism through incorporating data mining techniques to automatically discover some semantic links in text resources, adopting analogical reasoning to propose new semantic links, and using semantic link reasoning to derive, maintain semantic links and verify the soundness before normally adding new semantic links. (2) Investigate dynamic characteristic and evolution rules in the case of semantic change, since a semantic link network should be able to evolve with time to keep up-to-date content. (3) Apply the proposed approach to realize dynamic clustering of geographical images and integrate their containing data information. (4) Apply the proposed approach to organize document resources to realize an intelligent online education system. (5) Investigate the ranking algorithm that can differentiate the importance among nodes and among links, the growth model of the semantic link network and the analysis method of resource distribution in large-scale semantic link network, and the computing model for the single semantic image mechanism.

The proposed approach can be applied to semantic-based retrieval of other types of resources after making some revision on the resource description as well as adding more types of semantic links and rules.

Acknowledgements

The author thanks all members of China Knowledge Grid Research Group, especially P. Shi, Y. Li, R. Jia, N. Zhang, J. Liu and L. Zheng for their diligent work on implementing relevant software prototypes. Thanks also go to the anonymous reviewers who provide helpful comments on earlier version of this paper.

References

Adamic, L.A., Huberman, B.A., 2000. Power-law distribution of the World Wide Web. Science 287 (24), 2115.

Berners-Lee, T., Hendler, J., Lassila, O., 2001. Semantic Web. Scientific American 284 (5), 34–43.

Cascia, M.L., Sethi, S., Sclaroff, S., 1998. Combining textual and visual cues for content-based image retrieval on the worldwide web. In: Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries (Santa Barbara, CA). IEEE Computer Society, Los Alamitos, CA.

Chakrabarti, S., Dom, B., Gibson, D., Kleinberg, J., Raghavan, P., Rajagopalan, S., 1998. Automatic resource list compilation by analyzing hyperlink structure and associated text. In: Proceedings of the 7th International WWW Conference, Brisbane, Australia.

Dean, J., Henzinger, M.R., 1999. Finding related pages in the world wide web. In: Proceedings of the 8th International World Wide Web Conference, Toronto, Canada.

Fensel, D. et al., 2001. OIL: An ontology infrastructure for the Semantic Web. IEEE Intelligent Systems 16 (2), 38–45.

Fikes, R., Farquhar, A., 1999. Distributed repositories of highly expressive reusable ontologies. IEEE Intelligent Systems 14 (2), 73–79.

Gevers, T., Smeulders, A.W.M., 1999. The PicToSeek www image search system. In: Proceedings of the IEEE International Conference on Multimedia Computing and Systems, Florence, Italy.

Gupta, A., Jain, R., 1997. Visual information retrieval. Communications of the ACM 40 (5), 71–79.

Harmandas, V., Sanderson, M., Dunlop, M.D., 1997. Image retrieval by hypertext links. In: Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Philadelphia, PA, USA.

Hendler, J., 2001. Agents and the Semantic Web. IEEE Intelligent Systems 16 (2), 30–37.

Henzinger, M.R., Heydon, A., Mitzenmacher, M., Najork, M., 1999. Measuring index quality using random walks on the Web. In: The Eighth International World Wide Web Conference, Toronto, Canada.

Kanth, K., Agrawal, D., Singh, A., 1998. Dimensionality reduction for similarity searching in dynamic databases. In: Proceedings of the ACM SIGMOD International Conference on Management of Data, Seattle, Washington, USA.

Klein, M., 2001. XML, RDF, and Relatives. IEEE Internet Computing. 5 (2), 26–28.

Kleinberg, J.M., 1999. Authoritative sources in a hyperlinked environment. Journal of ACM 46 (5), 604–632.

Kleinberg, J., 2000. The small-world phenomenon: an algorithmic perspective. In: Proceedings of the 32nd ACM Symposium on Theory of Computing, Portland.

Kleinberg, J., Lawrence, S., 2001. The structure of the Web. Science 294 (30), 1849–1850.

Lempel, R., Moran, S., 2000. The stochastic approach for link-structure analysis (SALSA) and the TKC effect. In: Proceedings of the 9th International WWW Conference, Amsterdam, The Netherlands.

Lempel, R., Soffer, A., 2001. PicASHOW: Pictorial authority search by hyperlinks on the web. In: Proceedings of the 10th International WWW Conference, Hong Kong, China.

Mack, R., Ravin, Y., Byrd, R.J., 2001. Knowledge portals and the emerging knowledge workplace. IBM Systems Journal 40 (4), 925–955.

Rafiei, D., Mendelzon, A.O., 2000. What is this page known for? Computing web page reputations. In: Proceedings of the 9th International WWW Conference, Amsterdam, The Netherlands.

Zhang, H., Chen, Z., Li, M., Su, Z., 2003. Relevance feedback and learning in content-based image search. World Wide Web 6 (2), 131–155.

Zhuge, H., 2002. Clustering soft-devices in the semantic grid. IEEE Computing in Science and Engineering. 4 (6), 60–63.

Zhuge, H., 2003. Active document framework (ADF): Concept and method. In: Proceedings of the 5th Asia Pacific Web Conference, Xian, China Springer LNCS, vol. 2642, pp. 341–346.

Zhuge, H., in press. Resource space model, its design method and tool. Journal of Systems and Software.

Hai Zhuge is a professor at the Institute of Computing Technology, Chinese Academy of Sciences. He is serving on the editorial boards of the *Journal of Systems and Software*, *Information and Management*, *Journal of Computer Science and Technology*, and *Future Generation Computer Systems*. His current research interests include: future Web model, Semantic Grid and Knowledge Grid, Web-based workflow model and knowledge flow model. He is now the leader of the China Knowledge Grid Research Group (Vega-KG, http://kg.ict.ac.cn). He is the author of two books and over 40 papers appeared mainly in leading international conferences and the following international journals: *Communications of the ACM*; *IEEE Computer*; *IEEE Intelligent Systems*; *IEEE Computing in Science and Engineering*; *IEEE Transactions on Systems*, *Man*, *and Cybernetics*; *Information and Management*; *Decision Support Systems*; and *Journal of Systems and Software*.