Visual Media Retrieval Framework Using Web Service

Yunmook Nah[1], Bogju Lee[1], and Jungsun Kim[2]

[1] School of Electrical, Electronics, and Computer Engineering, Dankook University, San 8 Hannam-dong, Yongsan-gu, Seoul 140-714, Korea
{ymnah, blee}@dku.edu
[2] School of Electrical Engineering and Computer Science, Hanyang University, 1271 Sa-dong, Ansan-si, Kyungki-do 425-791, Korea
jskim@cse.hanyang.ac.kr

Abstract. The need for content-based image retrieval from image databases is ever increasing rapidly in many applications, such as electronic art museums, internet shopping malls, internet search engines, and medical information systems. Many such image resources have been previously developed and widely spread over the internet. In this paper, we propose a Web Service-driven architecture, named the HERMES(tHE Retrieval framework for visual MEdia Service), to support effective retrieval on large volumes of visual media resources. We explain how semantic metadata and ontology can be utilized to realize more intelligent content-based retrieval on visual media data.

1 Introduction

Searching image data is one of the essential functions for image database systems (or video database systems) and multimedia database systems, which take important role in majority of emerging IT(information technology) applications, such as electronic art museums, GIS(Geographical Information Systems), digital library, internet e-commerce, EDMS(Electronic Document Management Systems), and medical information systems. The need for content-based image retrieval(CBIR) from image databases is ever increasing rapidly in these applications.

The pioneering work has been done by IBM's QBIC(Query By Image Content) system, which supports queries using color, shape, sketch, and texture features on images, such as post stamps, art pictures, and trademark drawings [1]. The Chabot system was another interesting CBIR system, with high level concepts, such as "light yellow" and "sunset," as well as low level features based on color [2]. One of the most recent research work has been done by the SIMPLIcity system which supports CBIR based on the color, texture, shape features, while increasing matching correctness by utilizing local features on regions [3]. In the medical domain, the KMeD(Knowledge-Based Medical Database) system utilizes semantic modeling focusing on object shapes and spatial relationships between them [4,5]. We developed NERD-IDB, which supports the meaning-based retrieval on neuroscience image databases [6]. We also proposed web catalog image retrieval system, which support intelligent retrieval using keyword, color, texture features and high-level concepts [7,8].

S. Shimojo et al. (Eds.): HSI 2005, LNCS 3597, pp. 104 – 113, 2005.

Web Services are a standardized way of integrating Web-based applications using open standards including XML, the SOAP(Simple Object Access Protocol), the WSDL(Web Service Description Language), and the UDDI(Universal Description, Discovery, and Integration) specification. XML structures the message, SOAP transfers the message, WSDL describes the available services, and UDDI list them. XML describes both the nature of each self-contained function and the data that flows among systems [9]. The CPXe, proposed by Eastman Kodak, is a Web Service driven photo marketing system, but this system supports simple keyword-based searching only and does not utilize semantics [10]. The SOTA is an ontology-mediated Web Service system for smart office task automation [11]. The previously developed image databases, such as QBIC, Chabot, NERD-IDB, etc, can be considered as image service providers from the viewpoint of Web Services. The main idea of this paper is to exploit the Web Service-based approach to support intelligent retrieval on large volumes of visual media resources, widely distributed over the web.

This paper is an effort to make visual media, such as image and video data, better utilized by visual media consumers. The major purpose of this paper is: 1) to show how visual media, widely spread over the internet, can be effectively searched using Web Service technology and 2) to present how semantic metadata and ontology can be utilized to realize more practical and intelligent content-based retrieval on visual media. The architecture named the HERMES(tHE Retrieval framework for visual MEdia Service) is a Web Service-enabled visual media retrieval framework which consists of HERMES/B node (Web Service broker) and multiple HERMES/P nodes (Web Service providers), each servicing their own visual media resources.

The remainder of this paper is organized as follows. Section 2 describes overview of visual media metadata. Section 3 explains the Web Service-based visual media retrieval framework architecture. In section 4, we show how ontology and metadata are utilized, to allow more intelligent retrieval on visual media data. Section 5 describes Web Service workflow for service registration and query processing. Finally, section 6 summarizes the paper.

2 Visual Media Metadata

In database terms, *formatted data* means traditional data, such as numeric data and character data, and *unformatted data* means new kinds of multimedia data, such as text, image, graphics, audio, video, spatial data, time series data, and document data. Metadata means *data for data*. Let's assume image data is bitmap(or pixelmap) itself, which is usually called *raw data*. Image metadata is *data for image data*. There are two types of basic metadata for images.

- **registration metadata:** Image resolution(width, height), color map, compression ratio, etc are typical examples of registration metadata. This metadata is required to display and manipulate images. In image files, this information is usually hidden within image headers.
- **description metadata:** Image title, caption, keywords, natural language descriptions, and image file names are typical examples of descriptive metadata. This metadata is used to search images, when content-based retrieval operations are not supported.

For content-based retrieval, we can define multiple layers of image metadata on top of raw image data.

- **global feature metadata layer:** Average values or multi-dimensional vectors representing color, texture, and shape of images are examples of global feature metadata. This metadata is heavily used in current content-based retrieval systems.
- **local feature metadata layer:** Average values or multi-dimensional vectors representing color, texture, and shape of each objects or regions belonging to a given image are examples of local feature metadata.
- **semantic contents metadata layer** (or **knowledge layer**)**:** Subjective feelings and knowledges on images, such as concepts, meaning, category, spatial relationships, or other useful interpretations are examples of semantic metadata.

(a) Sample image data


```
<?xml version="1.0" encoding="UTF-8"?>
<dc:COMMON xmlns:dc = "http://purl.org/dc/elements/1.1/"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:xsi= "http://www.w3.org/2001/XMLSchema-instance" xsi:SchemaLocation=" ./dc.xsd">
    <dc:TITLE> Flowers </dc:TITLE>
    <dc:CREATOR> Van Gogh's </dc:CREATOR>
    <dc:RIGHTS> Neue Pinakothek </dc:RIGHTS>
    <dc:DATE> Year 1998 </dc:DATE>
</dc:COMMON>
```


(b) Image metadata expressed in DC standard

Fig. 1. Sample image and its metadata

There are many metadata standards, such as DC(Dublin Core) [13] for electronic cataloging, VRA(Visual Resources Association) [14] for visual art images, and MPEG-7 MDS [15] for video metadata description. Each of these metadata is proposed by different organization with different format. Also, these metadata standards usually put focus on description metadata only. MPEG-7 MDS partly covers global feature, local feature and semantic contents metadata. Figure 1 shows a sample image data and its metadata represented in DC standard.

3 Visual Media Service Architecture

The architecture named the HERMES is a Web Service-enabled visual media retrieval framework architecture which consists of HERMES/B node (Web Service broker) and multiple HERMES/P nodes (Web Service providers), each servicing visual media resources using their own metadata standard format or customized metadata format. The overall HERMES architecture is depicted in Figure 2.

Figure 3 and Figure 4 show the detail architecture of HERMES/B and HERMES/P, respectively. Albeit we did not show in Figure 2, besides the broker and service providers, we assume that there are other entities that provide local and/or global feature extraction services. They are called F.E.S (Feature Extraction Service) Providers. But, HERMES does not exclude the possibility of internal feature extraction services.

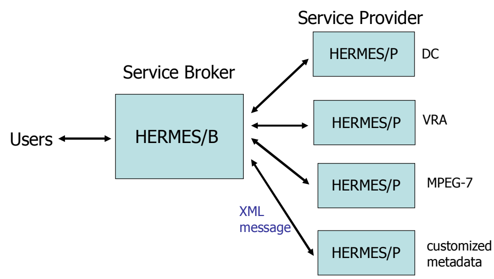

Fig. 2. The HERMES architecture

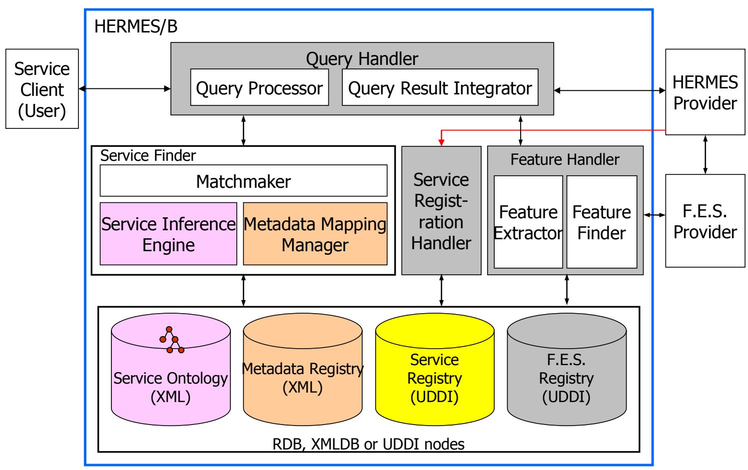

Fig. 3. HERMES Broker architecture

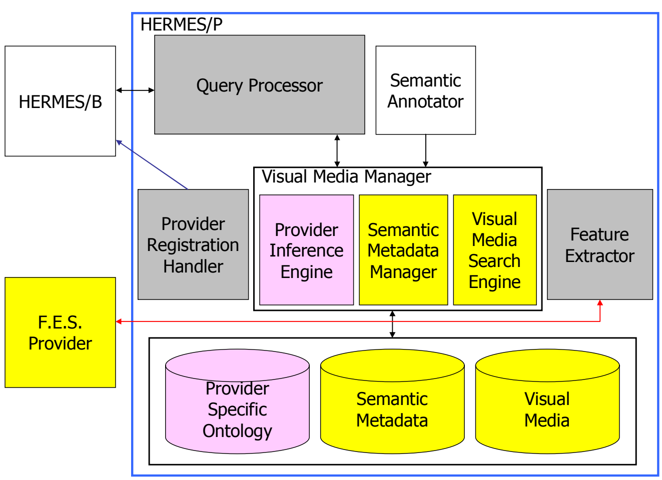

Fig. 4. HERMES Provider architecture

The major functions of HERMES/B modules are as follows. *Query Handler* receives user queries, reformulates queries using metadata and service ontology, and sends the reformulated queries to HERMES/Ps. *Query Processor* transforms query string into provider-specific XML queries. *Query Result Integrator* combines query results and sends them back to users. *Matchmaker* finds best service provider list for the given query. *Service Inference Engine* determines service provider list by using Service Ontology and Service Registry. *Metadata Mapping Manager* transforms query string into provider-specific format by using Metadata Registry. *Service Registration Handler* registers services provided by HERMES/Ps. Provider-specific service and metadata information are stored in Service Registry and Metadata Registry. *Feature Handler* selects suitable F.E.S. Providers and/or obtains features using them.

The major functions of HERMES/P modules are as follows. *Query Processor* receives and extends reformulated queries from HERMES/B. *Visual Media Manager* coordinates image searching processes. *Provider Inference Engine* reformulates user queries by using Provider Specific Ontology. *Semantic Metadata Manager* manages semantic metadata for each image using Semantic Annotator. Semantic metadata consists of description metadata, global feature, local feature and semantic contents metadata. Concepts in query term are transformed into corresponding color values. *Visual Media Search Engine* searches provider images by utilizing description metadata, global feature, local feature and semantic contents metadata. This module performs similarity-based retrieval using multi-dimensional index structures, such as B-trees and R-trees. *Provider Registration Handler* receives image category and common metadata information from HERMES/B. It also reports provider service types and metadata standard types (or its own metadata schema) to HERMES/B. *Semantic Annotator* is a provider-side tool to annotate metadata for each image.

4 Ontology Representation and Metadata Mapping

The representative sample queries are as follows: find 'modern painting' images whose creator is 'Albers' (Q1), find 'passionate' images whose creator is 'Van Gogh' (Q2), find photos of 'Californian' nature (Q3) and find pictures similar with the given image (Q4). Q1-Q3 are example queries, which require ontology-based service site matchmaking and metadata mapping. For Q4, we need to extract global and local features of the sample image by using F.E.S. Providers.

For Q1, we need service ontology to find provider sites related with image, art, painting and modern painting. For Q2, we need definition of concept 'passionate' with related emotion ontology to find 'passionate' images. For Q3, we need service ontology to find sites related with image, photo and scenery. Figure 5 shows Service Ontology, which is used by HERMES/B. The Service Ontology can be viewed as the superset of all the provider-specific ontologies from the HERMES/P. In other words, the provider-specific ontologies of the HERMES/P take the part of the Service Ontology. In this paper, the ontologies are assumed to have the tree form.

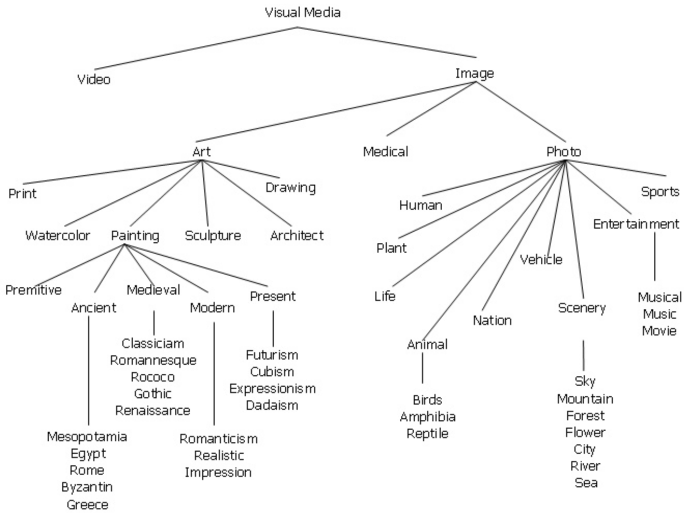

Fig. 5. The ontologies

The Service Ontology is stored in RDF format, as shown in Figure 6. The RDF has the form of subject-property-object which can easily represent various relationships such as ‘is-a’.

```
<owl:Ontology rdf:about="">
   <owl:imports rdf:resource="http://purl.org/dc/elements/1.1/"/>
   <owl:imports rdf:resource="http://www.w3.org/2001/XMLSchema"/>
   <owl:imports rdf:resource="http://www.w3.org/2000/01/rdf-schema"/>
   <owl:imports rdf:resource="http://www.w3.org/2002/07/owl"/>
   <owl:imports rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns"/>
 </owl:Ontology>
 <owl:Class rdf:ID="VisualMedia"> </owl:Class>
 <owl:Class rdf:ID="Video"> <rdfs:subClassOf rdf:resource="#VisualMedia"/>
 </owl:Class>
 <owl:Class rdf:ID="Image"> <rdfs:subClassOf rdf:resource="#VisualMedia"/>
 </owl:Class>
 <owl:Class rdf:ID="Art"> <rdfs:subClassOf rdf:resource="#Image/>
 </owl:Class>
 <owl:Class rdf:ID="Medical"> <rdfs:subClassOf rdf:resource="#Image"/>
 </owl:Class>
```

Fig. 6. Service Ontology in RDF format

Each metadata standard and customized metadata can use different element name for the exactly same element. For example, an element ‘Creator’ can be tagged by different element name ‘Creator Name’ in another metadata. Metadata Mapping Manager is used to handle these problems. Metadata Registry of HERMES/B consists of ProviderElement table, Mapping table and MDR(MetaData Registry) table. ProviderElement table stores provider-specific metadata information, MDR table stores standard metadata information and Mapping table stores mapping information between metadata formats. During the service registration process (see Section 5.2), when the given metadata schema type of current provider is not yet registered, the corresponding metadata schema information is processed and stored into ProviderElement table, Mapping table and MDR table of Metadata Registry.

```
<Map>
<MDRelement>
<DataElementName> <element>Creator</element> </DataElementName>
<ElementID>DC001</ElementID>
<DataType> <DataTypeName>string</DataTypeName> </DataType>
</MDRelement>
<LocalElement>
<ElementName>CreatorName</ElementName>
<ElementPath> <parent> Image </parent> </ElementPath>
<DataType> <DataTypeName>string</DataTypeName> </DataType>
</LocalElement>
<MappingRule>
<MappingType> <SemanticType>Direct_substitution</SemanticType>
</MappingType>
</MappingRule>
</Map>
```

Fig. 7. Metadata mapping example

Figure 7 is a mapping instance example. This XML instance shows that the 'Creator' element of DC standard is equivalent with 'Creator Name' of MDR element. By using this mapping information, a query with term Creator = 'Albers' can be reformulated as Creator Name = 'Albers.'

5 Web Service Interfaces and Workflow

In this section, we will first describe the Web Service interfaces for HERMES/B, HERMES/P and F.E.S. Providers, and then show the basic service workflow of the system in terms of two important scenarios: service registration and query processing scenarios.

5.1 Web Service Interfaces

For the sake of brevity, we only show externally significant operations in each of the Web interfaces for HERMES/B, HERMES/P and F.E.S. Providers. HERMES/B interface consists of the following operations: *getAllDomains* returns the list of all domains derived from the service ontology; *registerMe* registers a named service provider to HERMES/B together with information about metadata schema; *getFESInfo* returns the list of Feature Extraction Services for a specific domain images; and *acceptQuery* accepts client's query and returns a set of images collected from one or more service providers. HERMES/P interface includes *acceptQuery*, which analyzes reformulated query from HERMES/B and returns appropriate images. Feature Extraction Engine interface includes *GetFeatureInformation*, which receives a list of images and returns the list of extracted features.

5.2 Service Registration Scenario

In this scenario, service providers register themselves on the HERMES/B together with the information about their metadata schema. The service registration occurs in two phases: Service domain search and Service provider registration.

In first phase, the Provider Registration Handler of HERMES/P first calls getAllDomains() service of HERMES/B and receives visual media service domain information and service ontology from Service Registration Handler of HERMES/B. It then determines its corresponding service domain from the list.

In second phase, the Provider Registration Handler registers itself by calling registerMe() service of HERMES/B, while providing information, such as provider ID, domain list, metadata schema type, metadata schema version and metadata schema. This service call is received by the Service Registration Handler of HERMES/B. The Service Registration Handler then calls registerProvider() service of the Service Finder and stores UDDI-related information, such as ProviderID, Service Key, Service URL and tModel, into Service Registry and service domain list and metadata-related information into Metadata Registry. The service domain list information is stored into Matchmaker Cache table of Metadtata Registry by referencing appropriate Service Ontology and Service Registry.

Service providers can optionally use the Semantic Annotator for their registration for convenience purpose. Also, the Semantic Annotator of HERMES/P provides metadata fill-in form to insert description metadata and semantic contents metadata for each image. The input fields of this semantic annotation form are determined according to the metadata standard used by each HERMES/P. The whole metadata for each image, annotated manually or generated automatically, are stored as the Semantic Metadata of corresponding HERMES/P.

5.3 Query Processing Scenarios

When user(Service Client) fills query form, the query string is received by acceptQuery() service of the Query Handler of HERMES/B. The Query Handler then calls customizedQueryWithProviderList() service of the Service Finder. The Service Finder returns service provider list with provider-specific queries, which are reformulated by using service ontology and metadata mapping information of HERMES/B. For Q1, the appropriate service providers for ‘modern painting’ is determined by considering Service Ontology, Service Registry and Matchmaker Cache table of Metadata Registry. To find which providers service the ‘modern painting’, the Service Ontology is traversed from the root to the target node (Visual Meida - Image - Art - Painting - Modern, for Q1). All the nodes under the target node (Romanticism, Realistic, and Impression, for modern painting) are traversed too. Any provider that services any of the resulted node set should be included in the target provider list. The provider-service information is of course located in the Service Registry and Metadata Registry. For example, a provider that services ‘Art’ is included in the target provider list while the provider that services ‘Photo’ is not. Q3 is similarly processed: ‘Scenery’ under ‘Photo’ is the target node in this case. The query like Q2 is not easy to determine the provider list since there is no provider-specific information in the query. The query is simply transferred to all the providers.

The reformulated queries by HERMES/B are transferred to HERMES/P by calling acceptQuery() service of HERMES/P wrapper interface. A query received by HERMES/P is sent to executeQuery() service of the Visual Media Manager and then reformulated again by considering provider-specific ontology and semantic metadata of HERMES/P. In case of Q1, the ‘modern painting’ is expanded to {‘Romanticism’, 'Realistic', and 'Impression'} using the provider-specific ontology in HERMES/P so that the search engine can easily match the target images. In case of Q2, the emotional term 'passionate' is changed to corresponding color values by using Semantic Metadata [16]. The final formulated query is then transferred to the Visual Media Search Engine of HERMES/P. Query results from the participating HERMES/Ps are finally combined, ranked, and returned to the Service Client by the Query Result Integrator of HERMES/B. In case of Q4, the internal workflow of both HERMES/B and HERMES/P are still similar to that of previous query examples, except that interaction with external (possible internal) F.E.S. Providers are required to extract local and global features from the submitted image.

6 Conclusion

In this paper, we proposed HERMES architecture, which is a Web Service-enabled visual media retrieval framework architecture which consists of HERMES/B (broker) and multiple HERMES/P (provider), each servicing visual media resources using their own metadata standard format or customized metadata format. The HERMES architecture intensively uses semantic metadata and ontology to realize more practical and intelligent content-based retrieval on visual media. The proposed architecture can be utilized to effectively search visual media, which are widely spread over the internet, using leading-edge Web Service technology.

Our effort for the experimental implementation of the HERMES architecture is at an early stage. The fuller exposition of implementation details are deferred to subsequent papers. We believe that experimental studies of the performance aspects are highly meaningful subjects for future research.

Acknowledgements

This work was supported by grant No. R01-2003-000-10133-0 from the Basic Research Program of the Korea Science and Engineering Foundation. We would like to thank Chulbum Ahn, Eunyoung Kwon, Sungwoo Lee, Woosang Cho, Sangjin Han, Myoung Mi Yang and Byung Hun Jeong, who are currently developing the prototype of HERMES architecture.

References

1. Flickner, M. et al.: Query by Image and Video Content: The QBIC System. Computer. IEEE CS Press (Sept. 1995) 23-32
2. Ogle, V.E. and Stonebraker, M.: Chabot: Retrieval from a Relational Database of Images. Computer. IEEE CS Press (Sept. 1995) 40-48
3. Wang, J.Z., Li, J., and Wiederhold, G.: SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries. Transactions on Knowledge and Data Engineering, 23(9). IEEE CS Press (2001)
4. Chu, W.W., leong, I.T., and Taira, R.K.: A Semantic Modeling Approach for Image Retrieval by Content. VLDB J., 3. (1994) 445-477

5. Chu, W.W., Hsu, C.-C., Cardenas, A. F., and Taira, R. K.: Knowledge-Based Image Retrieval with Spatial and Temporal Constructs. Transactions on Knowledge and Data Engineering, 10(6). IEEE CS Press (1998) 872-888
6. Nah, Y. and Sheu, P. C.-y.: Image Content Modeling for Neuroscience Databases. Proc. Int'l Conf. on Software Eng. And Knowledge Eng.(SEKE). ACM Press (July 2002) 91-98
7. Lee, B. and Nah, Y.: A Color Ratio based Image Retrieval for e-Catalog Image Databases. Proceedings of SPIE: Internet Multimedia Management Systems II, Vol. 4519. SPIE (August 2001) 97-105
8. Hong, S., Lee, C., and Nah, Y.: An Intelligent Web Image Retrieval System. Proceedings of SPIE: Internet Multimedia Management Systems II, Vol. 4519. SPIE (August 2001) 106-115
9. Chung, J.-Y., Lin, K.-J. and Mathieu, R. G.: Web Services Computing: Advancing Software Interoperability. Computer, 36(10). IEEE CS Press (Oct. 2003) 35-37
10. Thompson, T., Weil, R. and Wood M. D.: CPXe: Web Services for Internet Imaging. Computer, 36(10). IEEE CS Press (Oct. 2003) 54-62
11. Tsai, T.M., Yu, H.-K., et al.: Ontology-Mediated Integration of Intranet Web Services. Computer, 36(10). IEEE CS Press (Oct. 2003) 63-71
12. Nah, Y. and Sheu, P. C.-y.: Searching Image Databases by Content. Proc. KSEA-SC Symposium. CSU, Fullerton (March 2002)
13. Dublin Core Organization: Dublin Core Metadata Initiative. http://dublincore.org /index.shtml/
14. Visual Resources Association Data Standards Committee: VRA Core Categories Ver. 3.0. http://www.vraweb.org/vracore3.htm
15. ISO IEC JTC1/SC29/WG11: Overview of the MPEG-7 Standard. MPEG2001/N4031, Singapore (March 2001)
16. Hong, S. and Nah, Y.: A Design and Implementation of Intelligent Color Image Retrieval System based on Emotion Information. Proc. Korean Database Conference. (May 2004) 243-250
17. Eakins, J.P. and Graham, M.E.: Content-based Image Retrieval A report to the JISC Technology Applications Programme. (Jan 1999) http://www.unn.ac.uk/iidr/research/cbir/report.html
18. Lew, M.S.: Next Generation Web Searches for Visual Content. Computer. IEEE CS Press (Nov. 2000) 46-53