

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
6 December 2007 (06.12.2007)

PCT

(10) International Publication Number
WO 2007/137352 A1

(51) International Patent Classification:
G06F 17/30 (2006.01)

(21) International Application Number:
PCT/AU2007/000746

(22) International Filing Date: 29 May 2007 (29.05.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2006902880 29 May 2006 (29.05.2006) AU

(71) Applicant (for all designated States except US): **UNIVERSITY OF WOLLONGONG** [AU/AU]; Northfields Avenue, Wollongong, New South Wales 2522 (AU).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **OGUNBONA, Philip** [AU/AU]; 68 Walang Avenue, Figtree, New South Wales 2525 (AU). **YE, Lei** [AU/AU]; 3 Sophia Court, North Rocks, New South Wales 2151 (AU).

(74) Agent: **FISHER ADAMS KELLY**; Level 29 Comalco Place, 12 Creek Street, Brisbane, Queensland 4000 (AU).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

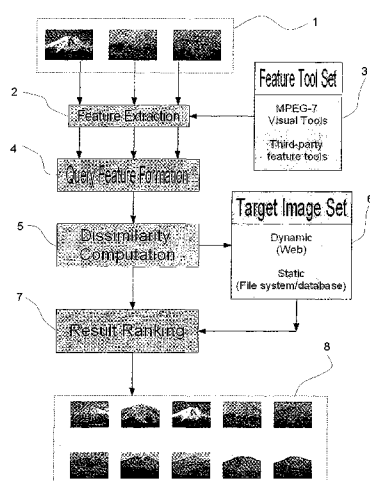
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CONTENT BASED IMAGE RETRIEVAL



(57) Abstract: A content based image retrieval system that extracts images from a database of images by constructing a query set of features and displaying images that have a minimum dissimilarity metric from images in the database. The dissimilarity metric is a weighted summation of distances between features in the query set and features of the images in the database. The method is useful for image searching such as web-based image retrieval and facial recognition.

WO 2007/137352 A1

CONTENT BASED IMAGE RETRIEVAL

This invention relates to a search tool for retrieval of images. In particular, it relates to a method of retrieving images based on the content
5 of the images.

BACKGROUND TO THE INVENTION

One of the most significant challenges faced in the information age is the problem of identifying required information from the vast quantity of
10 information that is accessible, particularly via the world wide web. Numerous text-based search engines have been developed and deployed. The best known of these are popular search engines that use keyword searching to retrieve pages from the world wide web. These engines include Google®, and Yahoo®.

15 Although it has been said that a picture is worth a thousand words, it cannot be said that image retrieval technology is as developed as text-based retrieval technology. Retrieval of images from a large collection of images remains a significant problem. It is no longer practical for a user to browse a collection of thumbnails to select a desired image. For instance,
20 a search as simple as "Sydney Opera House" results in 26000 hits in a Google® Images search at the time of writing.

Existing solutions to retrieving a particular image from a large corpus of images involves three related problems. Firstly, the images must be indexed in some way, secondly a query must be constructed and thirdly
25 the results of the query must be presented in a relevant away. Traditionally the images have been indexed and searched using keywords with the results being presented using some form of relevancy metric. Such an approach is fraught with difficulties since keyword allocation generally requires human tagging, which is a time-intensive process, and many
30 images can be described by multiple keywords.

An alternate approach is to use semantics classification methods as

described by Wang et. al. in "SIMPLcity: Semantics-Sensitive Integrated Matching for Picture Libraries" published in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 23, No 9, September 2001. The paper describes a region-based retrieval system that characterizes
5 regions by colour, texture, shape and location. The system classifies images into semantic categories, such as textured-nontextured, graph-photograph. Images are then retrieved by constructing a similarity measure based on a region-matching scheme that integrates properties of all the regions in the images. The Wang paper also includes a useful
10 summary of known content based image retrieval technologies.

Another approach is described by Jacobs et. al. in "Fast Multiresolution Image Querying" published in Proceedings of SIGGRAPH 95, In Computer Graphics Proceedings, Annual Conference Series, 1995, ACM SIGGRAPH, New York, 1995. Jacobs et. al. describe a pre-
15 processing approach that constructs signatures for each image in a database using wavelet decomposition. A signature for a query image is obtained using the same process. The query signature is then used to access the signatures of the database of images and a metric constructed to select images with similar signatures. The problem with this approach is
20 the necessity to pre-process all searchable images in order to derive a signature.

Iqbal and Aggarwal investigate the impact of feature integration on retrieval accuracy in their paper, "Feature Integration, Multi-image Queries and Relevance Feedback in Image Retrieval" presented at the 6th
25 International Conference on Visual Information Systems, Miami, Florida, 24-26 Sep 2003, pp 467-474. They extracted features of structure, color and texture from images in a database of 10221 images. They then measured retrieval performance using structure alone, color alone, texture alone, color and texture, and structure, color and texture. For image
30 retrieval they used CIRES (Content-based Image REtrieval System) developed by the University of Texas – Austin. Perhaps unsurprisingly they found that image retrieval was most effective when structure, color and texture were used. They also found that using multiple query images

resulted in more effective image retrieval.

Furthermore, Iqbal and Aggarwal investigated the benefit of user interaction via relevance feedback. Relevance feedback allows a user to indicate positive, negative and unsure images from the collection if images returned by an initial query. The query is modified by the user feedback and re-run. They found significant improvement in image retrieval with user feedback.

Although the recent prior art for image retrieval has a bias towards the problem of retrieving images from the world wide web it will be appreciated by persons skilled in the art that the problem is not dependent on the nature of the data store. The same prior art is relevant to selecting an image from a local store of images on a personal computer.

OBJECT OF THE INVENTION

It is an object of the present invention to provide a search method for content based image retrieval.

Further objects will be evident from the following description.

DISCLOSURE OF THE INVENTION

In broad terms the invention resides in a method of extracting images from a set of images including the steps of:
constructing a query set by extracting a set of features from one or more selected images;
constructing a dissimilarity metric as the weighted summation of distances between the features in the query set and features of images in the set of images; and
displaying the images having a minimum dissimilarity metric.

Preferably the weighted summation uses weights derived from the query set.

Suitably the invention further includes the step of ranking the order

of display of the displayed images. The images may be displayed in order from least dissimilar by increasing dissimilarity although other ranking schemes such as size, age, filename would also be possible.

5 BRIEF DETAILS OF THE DRAWINGS

To assist in understanding the invention preferred embodiments will now be described with reference to the following figures in which:

FIG 1 is a flowchart displaying the main steps in a method of content based image retrieval;

10 FIG 2 displays a screenshot exemplifying an initial search as a starting point for a first application of the invention;

FIG 3 displays a screenshot exemplifying a set of images from the initial search;

15 FIG 4 displays the screenshot of FIG 3 with three images selected to form the query set;

FIG 5 displays a screenshot of the results of content based image retrieval according to the invention;

FIG 6 displays a screenshot of image thumbnails in a directory; and

20 FIG 7 displays the screenshot of FIG 6 with three images selected to form a query set.

DETAILED DESCRIPTION OF THE DRAWINGS

In describing different embodiments of the present invention common reference numerals are used to describe like features.

25 The goal of the method is to retrieve images based on the feature content of images and a user's query concept. The user's query concept is automatically derived from image examples supplied or selected by the user. It achieves the goal with an innovative method to extract perceptual importance of visual features of images and a computationally efficient
30 weighted linear dissimilarity metric that delivers fast and accurate retrieval

results.

In multi-image query systems, a query is a set of example images $Q = \{I_{q1}, I_{q2}, \dots, I_{qQ}\}$. The set of example images may be any number of images including one. Much of the prior art constructs a query based upon
5 a single query image but the preferred approach of this invention is for a user to provide at least two and preferably three images. The user supplied images may be selected directly from a database or may be identified through a conventional image search, such as that mentioned above using Google[®] Images.

10 For the following description the target image set, sometimes called the image database, is defined as $T = \{I_m : m = 1, 2, \dots, M\}$. The query criteria is expressed as a similarity measure $S(Q, I_j)$ between the query set Q and an image I_j in the target image set. A query system $Q(Q, S, T)$ is a mapping of the query set Q to a permutation T_p of the target image set T , according
15 to the similarity $S(Q, I_j)$, where $T_p = \{I_m \in T : m = 1, 2, \dots, M\}$ is a partially ordered set such that $S(Q, I_m) > S(Q, I_{m+1})$. In principle, the permutations are that of the whole database, in practice only the top ranked output images are evaluated.

The method of content based image retrieval is summarised in FIG
20 1 and explained in greater detail below. The method commences with the query set 1. The feature extraction process 2 extracts a set of features using a feature tool set 3, which may be any of a range of third party feature tools, including those mentioned above. A query is then formed 4 from the extracted features.

25 The query can be thought of as an idealized image constructed to be representative of the images in the query set.

A key aspect of the invention is calculation of a dissimilarity metric 5 which is applied to the target image set 6 to identify images that are similar to the set of features forming the query. The images are then ranked 7 and
30 presented to the user 8.

Feature Extraction

The feature extraction process bases the query on low level structural descriptions of images. An image object I can be described by a set of features $X = \{x_n : n = 1, 2, \dots, N\}$. Each feature is represented by a k_n -
 5 dimensional vector $x_n = \{x_1, x_2, \dots, x_{k_n}\}$ where $x_{n,i} \in [0, b_{n,i}] \subset R$, R is the real number. The n^{th} feature extraction is a mapping from image I to the feature vector as:

$$x_n = f_n(I) \quad (1)$$

The invention is not limited to extraction of any particular set of
 10 features. A variety of visual features, such as color, texture or facial features, can be used. Third party visual feature extraction tools can be plugged into the system.

For example, the popular MPEG-7 visual tools is suitable, the MPEG-7 Color Layout Descriptor (CLD) is a very compact and resolution-
 15 invariant representation of color which is suitable for high-speed image retrieval. It uses only 12 coefficients of 8x8 DCT to describe the content from three sets (six for luminance and three for each chrominance), as expressed as follows.

$$x_{CLD} = (Y_1, \dots, Y_6, Cb_1, Cb_2, Cb_3, Cr_1, Cr_2, Cr_3) \quad (2)$$

20 The MPEG-7 Edge Histogram Descriptor (EHD) uses 80 histogram bins to describe the content from 16 sub-images, as expressed as follows.

$$x_{EHD} = (h_1, h_2, \dots, h_{80}) \quad (3)$$

While the MPEG-7 set of tools is useful, the invention is not limited to this set of feature extraction tools. As is evident from the prior art there
 25 are a range of feature extraction tools that characterize images according to such features as colour, hue, luminance, structure, texture, location, etc.

As mentioned above, the invention may be applied to a set of facial features to identify a face from a database of faces. The feature extraction process may extract facial features such as distance between the eyes,
 30 colour of eyes, width of nose, size of mouth, etc.

Query Feature Formation

The query concept of the user is implied by the example images selected by the user. The query feature formation module generates a
5 virtual query image feature set that is derived from the example images.

The fusion of features forming one image may be represented by

$$x^i = (x_1^i \oplus x_2^i \oplus \dots \oplus x_n^i) \quad (4)$$

For a set of query images the fusion of features is

$$X = (x^1 \oplus x^2 \oplus \dots \oplus x^m) \quad (5)$$

10 The query feature formation implies an idealized image which is constructed by weighting each feature in the feature set used in the feature extraction step. The weight applied to the i^{th} feature x_i is:

$$w_i = f_w^i(x_1^1, x_2^1, \dots, x_n^1; x_1^2, x_2^2, \dots, x_n^2; \dots; x_1^m, x_2^m, \dots, x_n^m) \quad (6)$$

The idealized image I_Q constructed from the set of query images Q
15 could then be considered to be the weighted sum of features x_i in the feature set:

$$I_Q = \sum_i w_i x_i \quad (7)$$

Dissimilarity Computation

20 The feature metric space X_n is a bounded closed convex subset of the k_n -dimensional vector space R^{kn} . Therefore, an average, or interval, of feature vectors is a feature vector in the feature set. This is the base for query point movement and query prototype algorithms. However, the average feature vector may not be a good representative of other feature
25 vectors. For instance, the colour *grey* may not be a good representative of colours *white* and *black*.

In the case of a multi-image query, the distance is measured between the query image set $\{I_{q1}, I_{q2}, \dots, I_{qQ}\}$ and an image $I_j \in T$, as

$$D(Q, I_j) = D(\{I_{q1}, I_{q2}, \dots, I_{qQ}\}, I_j) \quad (8)$$

The invention uses a distance function expressed as a weighted summation of individual feature distances, as follows

$$D(I_q, I_m) = \sum_{i=1}^N w_i \cdot d_i(x_{qi}, x_{mi}) \quad (9)$$

5 This equation calculates a measure which is the weighted summation of a distance metric d between query feature x_q and queried feature x_n .

The weights w_i are updated according to the query set using equation (6). For instance, the user may be seeking to find images of
 10 bright coloured cars. Conventional text based searches cannot assist since the query 'car' will retrieve all cars of any colour and a search on 'bright cars' will only retrieve images which have been described with these words, which is unlikely. However, an initial text search on cars will retrieve a range of cars of various types and colours. When the user
 15 selects a query set of images that are bright the query feature formation will give greater weight to the luminance feature than, say, colour or texture. On the other hand if the user is looking for blue cars the query set will be selected from only blue cars. The query feature formation will give greater weight to the feature colour and to the hue blue than to luminance
 20 or texture.

In each case the dissimilarity computation is determining a similarity value that is based in the features of the query set selected by the user without the user being required to define the particular set of features being sought. It will be appreciated that this is a far more intuitive image
 25 searching approach than is available in the prior art.

Result Ranking

The images extracted from the image set using the query set are conveniently displayed according to a relevancy ranking. There are
 30 several ways to rank the output images and the invention is not limited to

any specific process. One convenient way is to use the dissimilarity measure described above. That is, the least dissimilar (most similar) images are displayed first followed by more dissimilar images up to some number of images. Typically the twenty least dissimilar images might be
 5 displayed.

So, the distance between the query image set and a target image in the database is defined as follows, as is usually defined in a metric space.

$$d(Q, I_j) = \min_{I_q \in Q} \{d(X_q, X_j)\} \quad (10)$$

The measure of (10) has the advantage that the top ranked images
 10 will be similar to one of the example images, which is highly expected in a retrieval system, while in the case of the prototype query, the top ranked images will be similar to an image of average features, which is not very similar to any of the example images. The former will give better experience to the user in most applications.

15

Example 1

A demonstration implementation of the invention has been implemented using Java Servlet and JavaServer pages technologies supported by Apache Tomcat[®] web application server. It searches the
 20 images based on image content on the Internet via keyword based commercial image search services like Google[®] or Yahoo[®]. The current implementation may be accessed using any web browsers, such as Internet Explorer or Mozilla/Firefox, and consists of a 3-step process to search images from the Internet.

25 In order to demonstrate the operation of the invention it has been applied to the example of finding an image of the Sydney Opera House using Google[®] Images, which was mentioned above.

1) First Step: Keyword based search as shown in FIG 2. Use keywords to retrieve images from the Internet via a text based image search services to
 30 form an initial image set as shown in FIG 3.

2) Second Step: Select example images from the initial search results as

shown in FIG 4. Select image examples the user intends to search by clicking image checkboxes presented to the user from the keyword based search results.

- 3) Third Step: Conduct a search of all images using the query constructed from the sample images. The results are presented in a ranked sequence according to similarity metric as shown in FIG 5.

As can be seen from the example, the images of the result set shown in FIG 5 are all relevant whereas the images shown in FIG 3 include images of doubtful relevance.

10

Example 2

The invention can be integrated into desktop file managers such as Windows Explorer® or Mac OS X Finder®, both of which currently have the capability to browse image files and sort them according to image filenames and other file attributes such as size, file type etc. A typical folder of images is shown in FIG 6 as thumbnails. The user selects a number of images for constructing the query set by highlighting the images that are closest to the desired image. In the example of FIG 7 the user has selected images that have the Sydney Harbour Bridge as a background to the Sydney Opera House.

20

The user then runs the image retrieval program, which is conveniently implemented as a plug-in. In FIG 6 and FIG 7 the invention is activated by clicking the tick icon 9 on the tool bar.

25 Conclusion

The method of content based image retrieval described above has a number of advantages compared to the prior art systems including:

- Perceptual importance is derived automatically from user examples;
- The search process is intuitive;
- The user is not required to select features or weights for features;

30

- A weighted linear dissimilarity metric is generic, applicable to all features;
 - The weight generation and dissimilarity formula are computationally efficient and deliver very fast retrieval results;
- 5
- Feature extraction tools are pluggable – standard and third-party features can be integrated into the architecture;
 - Users need not supply negative examples.

Throughout the specification the aim has been to describe the invention without limiting the invention to any particular combination of

10 alternate features.

CLAIMS

1. A method of extracting images from a set of images including the steps of:
constructing a query set by extracting a set of features from one or more
5 selected images;
constructing a dissimilarity metric as the weighted summation of distances
between the features in the query set and features of images in the set of
images; and
displaying the images having a minimum dissimilarity metric.
- 10 2. The method of claim 1 wherein the query set is extracted from at
least two images.
3. The method of claim 1 wherein the query set is extracted using a
feature tool set.
4. The method of claim 1 wherein the query set is extracted using low
15 level structural descriptions of the images.
5. The method of claim 1 wherein the features are selected from one
or more of: colour; texture; hue; luminance; structure; location; facial
features.
6. The method of claim 1 wherein the query set is an idealized image
20 constructed as a weighted sum of the set of features.
7. The method of claim 6 wherein the idealized image is $I_Q = \sum_i w_i x_i$
where x_i is a feature and w_i is the weight applied to the feature.
8. The method of claim 1 wherein the weighted summation uses
weights derived from the query set.
- 25 9. The method of claim 1 wherein the dissimilarity metric is

$$D(I_q, I_m) = \sum_{i=1}^N w_i \cdot d_i(x_{qi}, x_{mi}).$$
10. The method of claim 1 further including the step of ranking the
order of display of the displayed images.

11. The method of claim 7 wherein the ranking is in order of similarity.
12. Software embedded in one or more computer-readable media and when executed operable to:
construct a query set by extracting a set of features from one or more
5 selected images;
construct a dissimilarity metric as the weighted summation of distances
between the features in the query set and features of images in the set of
images; and
display the images having a minimum dissimilarity metric.
- 10 13. The software of claim 12 further operable when executed to rank
the images having a minimum dissimilarity metric in order of similarity.

1/4

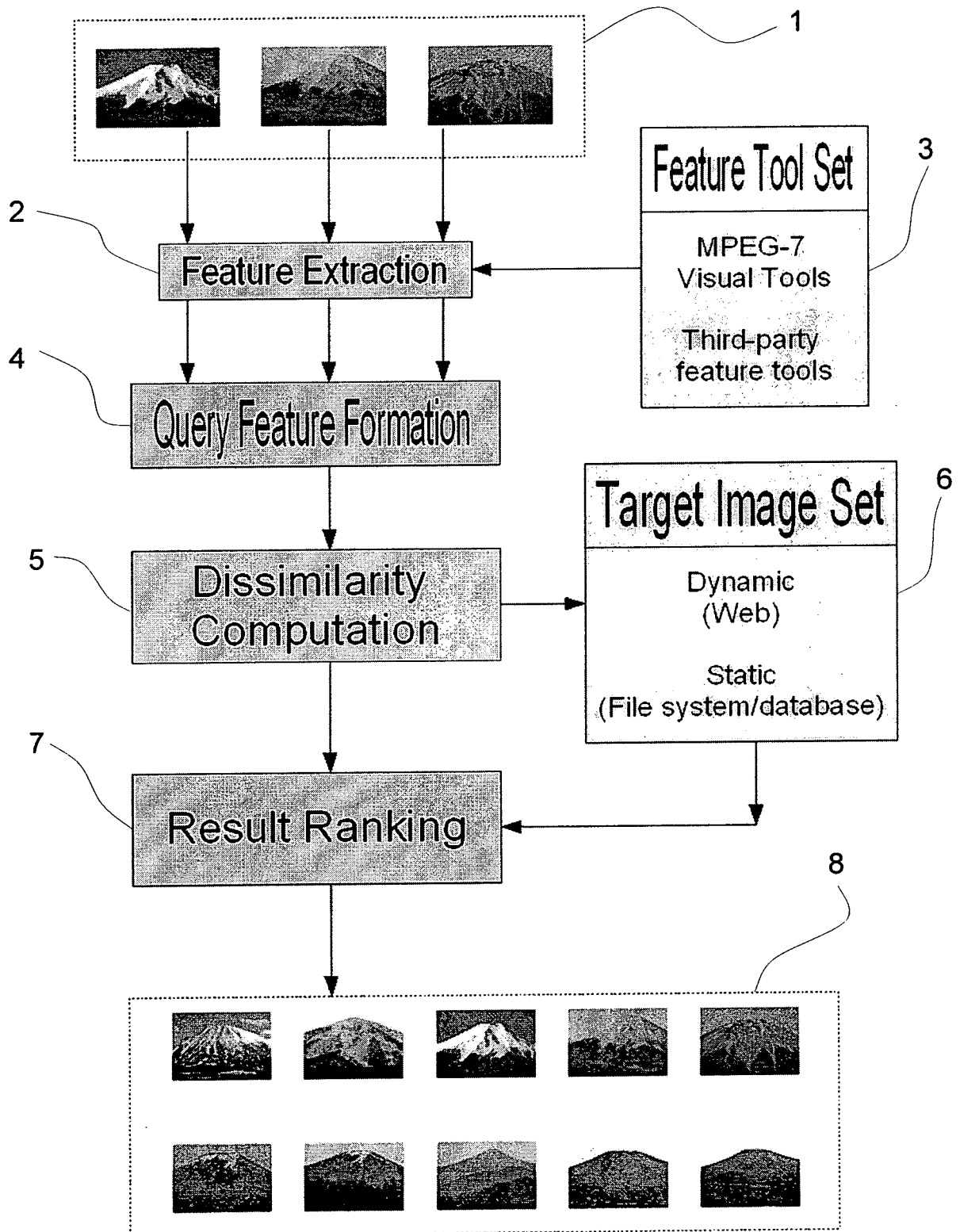


FIG 1

2/4

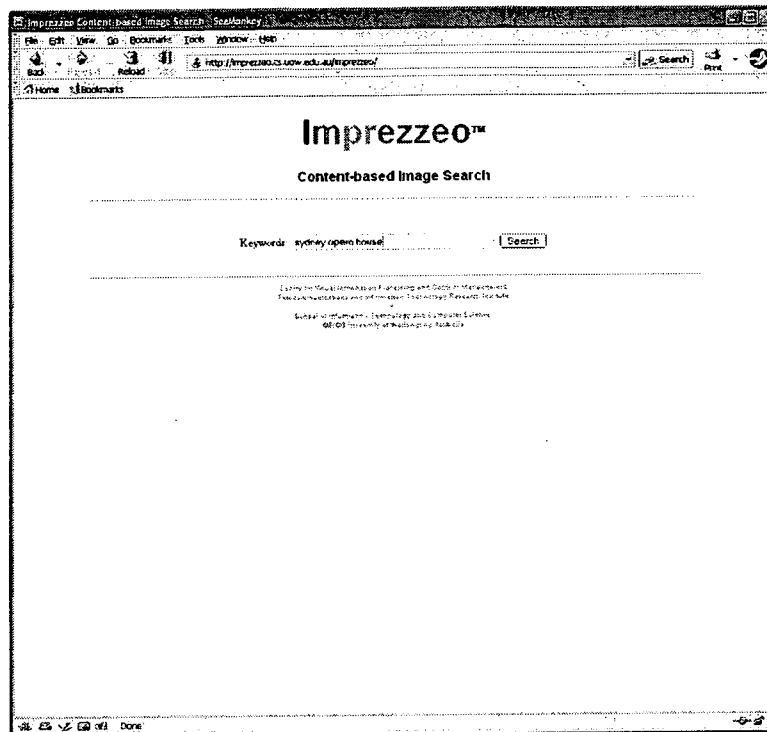


FIG 2

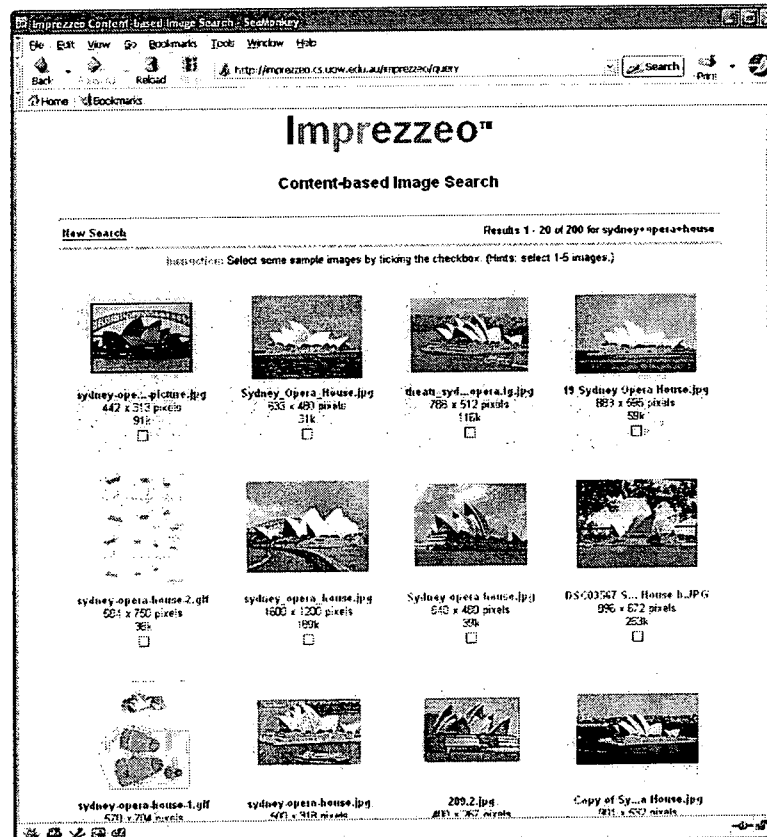


FIG 3

3/4

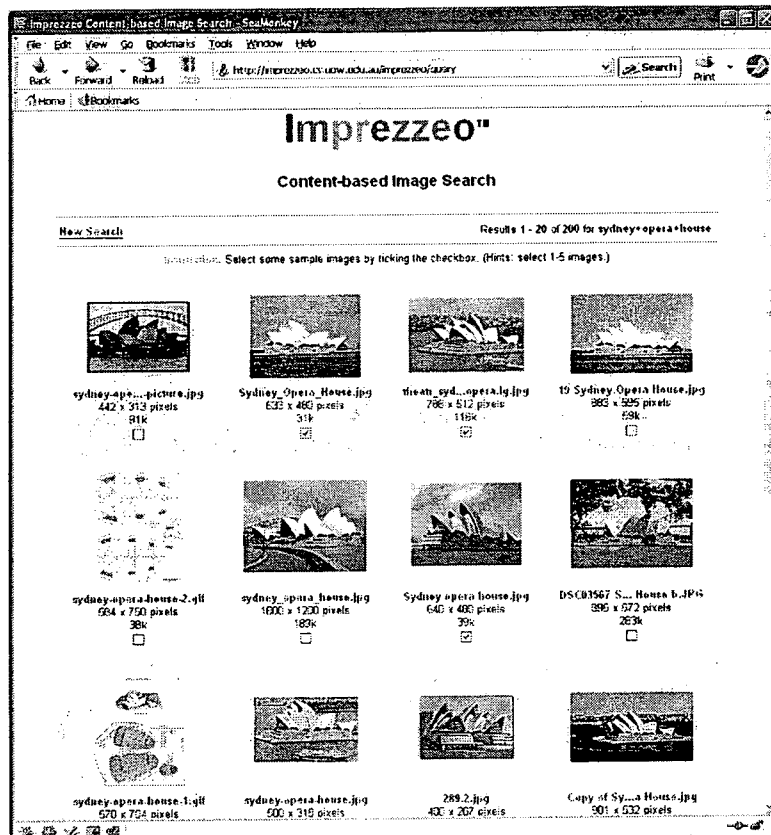


FIG 4



FIG 5

4/4

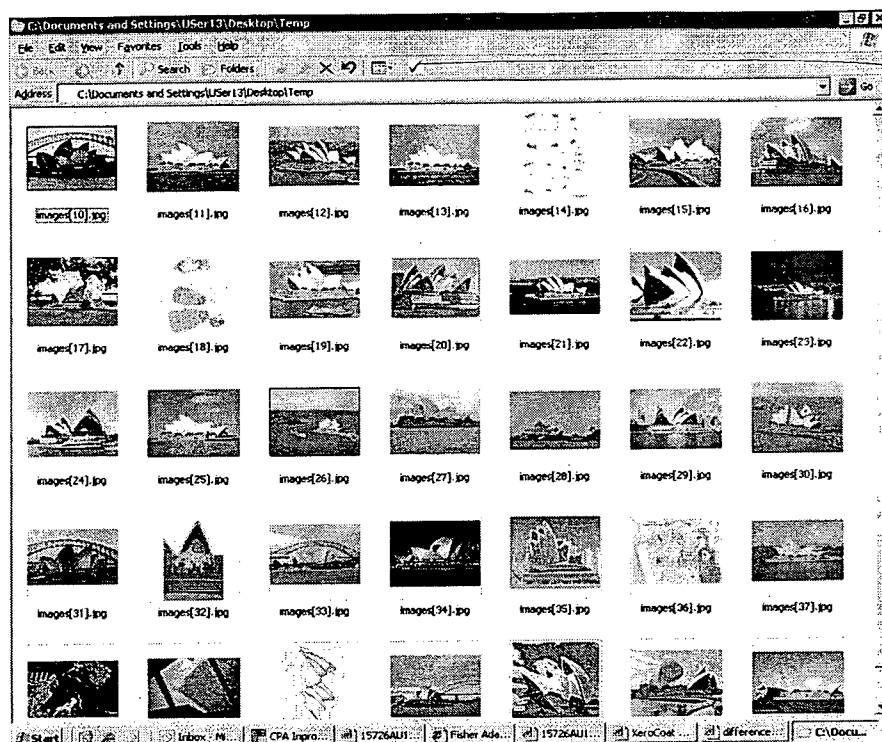


FIG 6

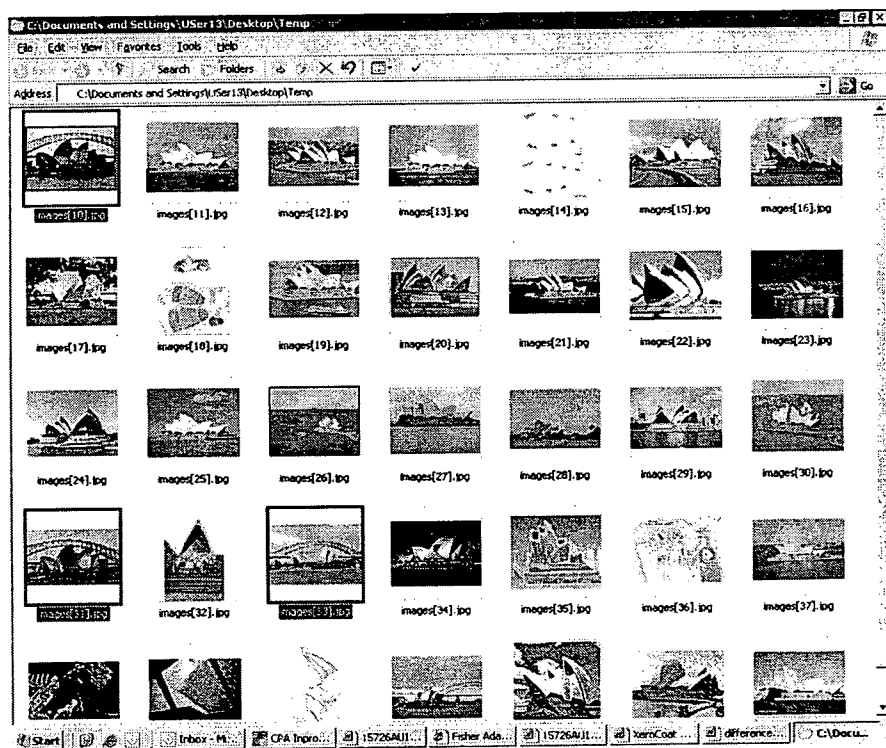


FIG 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2007/000746

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

G06F 17/30 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DWPI, USPTO, PCT Gazette, esp@cenet: image, feature, search, query, similar, weight

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US-2005/0131951 (ZHANG et al.) 16 June 2005 Paragraphs [0024], [0032], [0069], [0078]-[0079], [0115], [0121]-[0122], [0144]-[0145]	1-13
X	US-6859802-B1 (RUI) 22 February 2005 Column 5 line 58 to column 6 line 57, column 7 lines 5 to 21, column 11 line 51 to column 12 line 40.	1-13
X	US-2002/0178149-A1 (CHEN et al.) 28 November 2002 Paragraphs [0016]-[0036], especially paragraph [0018] and claim 2	1-13
X	US-5893095-A (JAIN et al.) 6 April 1999 Column 7 lines 9 to 22, column 11 lines 25 to 34, column 12 lines 7 to 22, column 12 line 62 to column 13 line 15	1, 3-5, 9-10, 12-13



Further documents are listed in the continuation of Box C



See patent family annex

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

18 July 2007

Date of mailing of the international search report

23 JUL 2007

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustalia.gov.au
Facsimile No. (02) 6285 3929

Authorized officer

MICHAEL HARDY

AUSTRALIAN PATENT OFFICE

(ISO 9001 Qualified Service)

Telephone No : (02) 6283 2547

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/AU2007/000746

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member	
US 2005131951	US 6748398	US 7111002	US 7113944
	US 2002174120	US 2004243541	US 2006248044
US 6859802	US 7028325	US 2005065929	US 2005086223
	US 2005159956	US 2005160457	
US 2002178149	US 6834288		
US 5893095	US 5911139	US 5913205	US 5915250
	US 5983237		
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			
END OF ANNEX			