An Integrated Content and Metadata Based Retrieval System for Art

Paul H. Lewis, Kirk Martinez, Fazly Salleh Abas, Mohammad Faizal Ahmad Fauzi, Stephen C. Y. Chan, Matthew J. Addis, Mike J. Boniface, Paul Grimwood, Alison Stevenson, Christian Lahanier, and James Stevenson

Abstract—A new approach to image retrieval is presented in the domain of museum and gallery image collections. Specialist algorithms, developed to address specific retrieval tasks, are combined with more conventional content and metadata retrieval approaches, and implemented within a distributed architecture to provide cross-collection searching and navigation in a seamless way. External systems can access the different collections using interoperability protocols and open standards, which were extended to accommodate content based as well as text based retrieval paradigms. After a brief overview of the complete system, we describe the novel design and evaluation of some of the specialist image analysis algorithms including a method for image retrieval based on sub-image queries, retrievals based on very low quality images and retrieval using canvas crack patterns. We show how effective retrieval results can be achieved by real end-users consisting of major museums and galleries, accessing the distributed but integrated digital collections.

Index Terms—Art images, content based image retrieval, crack analysis, wavelets.

I. INTRODUCTION

CONTENT based image retrieval from large digital collections has continued to challenge the image processing and computer vision communities in-spite of substantial research efforts in the area in the last ten to fifteen years [9], [23], [28]. Although content-based techniques offered, and continue to offer, the possibility of alleviating, if not overcoming, the problem of labour intensive keyword classification, their promise remains largely unrealized and commercial content based image retrieval systems have so far failed to achieve significant commercial success [24]. General approaches based mainly around color and texture feature extraction are still being refined [25], [26], while current shape based retrieval is only useful in rather limited domains [27].

The need for effective image retrieval facilities continues to increase as digital libraries, personal image capture systems and the applications of imaging in science and medicine all expand rapidly. The problem of image retrieval on the web is a related challenge and it is interesting to note that here too the approach has been to rely mainly on associated text and metadata. Most commercial users of image collections also continue to use metadata for general purpose retrieval tasks and the semantic gap is identified as the main culprit in the field [29]. Some of the most promising work attempting to overcome the problem is concerned with automating the building of associations between image features and semantic concepts [30], [31]. However, such research is in its infancy and does not offer viable solutions for real applications in the immediate future. Various commercial and close to commercial experimental research systems have appeared such as QBIC [37], Virage [32], Netra [33], VisualSEEK [34], Mars [35], Photobook [36] and many others offering a variety of techniques and user interfaces for content-based retrieval, but none of these is seen as the panacea for all image database problems.

This paper is concerned with the problem of delivering effective integrated content and metadata based image retrieval across disparate image databases for several major museums and galleries across Europe. Even when textual metadata is available, problems of incompatibility between metadata schema can make cross-collection searching problematic. Prior to this work, access from one gallery system to another was not common even when using only textual data and it was not undertaken at all for cross-collection content based retrieval. However, the value of cross-collection searching was seen as a highly desirable objective, not only between museums, but between multiple databases within a single museum. In this paper we show how several novel and specific content-based techniques combined with general content-based techniques and a novel approach to interoperability can make cross-collection content and metadata based searching an effective approach for real end-users. The work formed part of an EU funded project called ARTISTE.

Content-based retrieval (CBR) of art images has been a particular challenge [1]–[3] as the image content covers a vast range of subjects and requirements from end-users are often very loosely defined. "Traditional" content based image retrieval algorithms have typically been applied with only limited success. In our approach we have augmented these traditional content-based retrieval algorithms with algorithms targeted at specific retrieval tasks identified as a high priority for the museums. We have also addressed the need for seam-
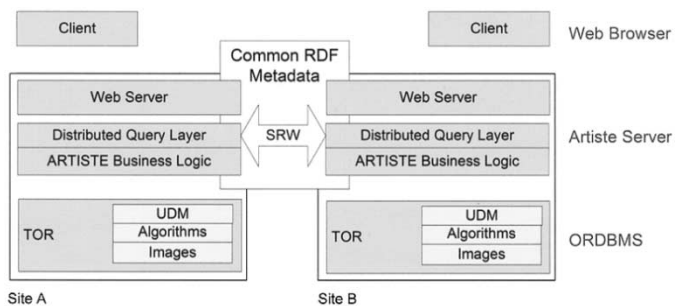
Fig. 1. ARTISTE system architecture.

less cross-collection metadata based searching which can be applied either separately or in combination with the content based searching. Individual database schemas of the galleries are preserved and mapped to a common schema using RDF (Resource Description Framework) [4] and the distributed architecture not only preserves each gallery's control over its own collection but accelerates and simplifies cross-collection searching in an integrated manner.

The next section gives a brief overview of the architecture developed to facilitate the distributed cross-collection searching and in the third section we describe the content based facilities implemented within the architecture. Particular attention is given to the novel special purpose algorithms introduced to provide sub-image matching and query by low quality image. The motivation for these algorithms and their evaluation are also presented in Section III. Real examples of combined and cross-collection searching are given in Section IV and conclusions and future work are presented in Section V.

II. ARCHITECTURE FOR CROSS COLLECTION SEARCHING

The ARTISTE architecture, developed to facilitate cross-collection content and metadata based searching, is illustrated schematically in Fig. 1. Each participating museum has its own instantiation of the system and its images were held as binary large objects (BLOBS) in an object relational database management system (ORDBMS) from NCR, although this has since been replaced with mySQL [8]. The image processing functions were implemented as user defined modules (UDMs), which are executed within the database, minimizing data transfers and facilitating use of the system's parallelism [7]. Metadata is also stored in the database and complex queries, involving both image processing functions and metadata can be called directly from SQL.

Access to the system is achieved through standard web browsers, the pages being generated in the presentation layer using JSP (Java Server Pages) and Java Servlet technology. The distributed query layer, which is implemented as a web service using XML messaging and SOAP (Simple Object Access Protocol), provides a single interface to the disparate image and metadata collections and enables end users to seamlessly search individual or combined collections. Interoperability is usually achieved by using open standards, but as retrieval standards have been traditionally concerned with text based searching, we participated in an initiative [39] to redesign the open standard for interoperability between digital libraries, z39.50, using web

technologies. The initiative proposed a Search and Retrieve Web Service (SRW) for searching databases containing metadata and objects and we extended the capabilities of SRW to enable both image content and metadata based search over multiple collections.

An image processing API was implemented to allow the Artiste Server to access the image processing functions within the ORDBMS. With over 2 Tbytes of images from the museums, scalability was a prime concern. Images accepted into the systems are analyzed by the UDMs to produce feature vectors which are also stored as BLOBS within the database. During a query, these feature vectors are compared by fast operators which return scores. All the image processing code was written in C/C++ using the VIPS/ip library [6]. An administration interface also allowed new algorithms to be added to the system.

The syntax and semantics of standard metadata terms are defined using resource description format (RDF) [4] and mappings are provided from each collection. The mappings refer to common metadata schema such as Dublin Core [5] but are not limited to this. RDF schema is also used to describe image content as a metadata item. The operators (implemented using image processing algorithms.) relating to image content are explicitly defined along with the constraints under which these operators can be applied. For example, the multi-scale color coherence vector algorithm, MCCV, (see following section for details) implements a "part of" operator which can be applied to an image to allow a sub-image to be found within it. The textual metadata equivalent would be to search within a string to find a key word. In this way it can be seen that a uniform approach is taken to the semantics of searching both textual content and image content.

The use of RDF provides a flexible solution to cross-collection searching. Mapping to common semantics requires no changes to local metadata and schemas. Multilingual translation of metadata attribute names allows the user to use their native language when specifying which attributes to search over for multiple collections. Finally, by publication of the RDF, a gallery can provide a machine readable description of any search and retrieval services it chooses to expose. This allows software clients to dynamically constrain the search facilities offered to users so that only the available query functionality is used.

III. CONTENT BASED IMAGE RETRIEVAL

Many content-based image retrieval (CBR) systems work with whole images and address questions of the form find me images similar to this query. General techniques based on such features as color distribution, texture, outline shape and spatial color distribution have been popular in the research literature and in content based retrieval systems. Several extensive general reviews of content-based image retrieval techniques have appeared in recent years [9], [23], [28]. CBR has been applied to art images since the 1980s with the Morelli project [1] and IBM's QBIC for example has been applied to such images in collaboration with UC-Davis [2] Previous approaches usually applied a generic CBR system to a group of Art images to see how useful it would be. In Artiste, by contrast, our aim has been

Fig. 2. The image of the baby represents a query, and the result of sub-image matching is shown as a rectangular boundary in the *Moise présenté à Pharaon* image on the right.

to augment these traditional CBR approaches with solutions to specific CBR problems by developing algorithms tailored to the tasks, rather than only producing generic solutions for all queries. Use of the color histogram [10] for comparing images has been popular, primarily because it is easy to compute, is fairly insensitive to small variations between images and is relatively efficient in terms of computation speed and storage requirements. It has the disadvantage that it does not capture any information about the distribution of the colors spatially within the image and various techniques to capture such information have been proposed including the color coherence vector approach [11] and the use of color pairs [12].

In the Artiste system we began by introducing standard CBR algorithms for color based retrieval based on histograms of various color spaces (RGB, HSV, LUV), for texture based retrieval using the pyramid wavelet transform [38] and for spatial color based retrieval using color coherence vectors. Feature vectors for all these algorithms were pre-computed when the images entered the system and facilitated very general content based searching on a full image basis which could be combined with metadata based searching to effect retrieval. However, many techniques like these only work on the complete image, will not allow the query image to be a sub-image of the image to be retrieved and often require similarity in image resolution between query and target.

A. Sub-Image Matching

Our collaborating galleries had a specific requirement to be able to retrieve the full image of a painting, given a query image which is of a fragment of the full painting (i.e., a sub-image). The query may have been captured under different conditions and possibly at a different stage in the restoration process from the parent. If the connection between the parent painting and the fragment has been lost, the need for sub-image matching in order to re-identify the parent becomes clear, particularly as the collections contain large numbers of very high resolution images. Our approach to the sub-image problem has been to divide the images into a pyramid of image patches, extracting feature vectors for each patch and matching the query feature vectors with each of the feature vectors from the database images to find the best match [13]. We have used a patch size of 64 by 64 pixels, a compromise reached by experimentation. The requirements were not for a match to the nearest pixel but to a general area. In this approach we also have patches offset horizontally by 32 pixels both horizontally and vertically to provide a hit location to within 32 pixels. The resolution of the image is then reduced by a factor of two and the process repeated recursively until the image is approximately covered by a single patch. In general, global techniques such as the color histogram are not effective for sub-image matching in their basic form and previous workers have also used the hierarchical approach.

For matching based on color combined with its layout, we used the color coherence vector (CCV) [11] as the feature vector of each tile, since it carries useful local spatial information. Our hierarchical version is named the multiscale CCV or MCCV algoritm. The color coherence vector records the numbers of coherent and incoherent pixels of each color in the patch, where a coherent pixel is one which belongs to a region of similar colored pixels whose relative size is above a certain threshold. The color space is quantised into 64 cells. The CCVs are coded for rapid matching and could be compared at an average rate of 265 000 per second on a 600 MHz Pentium III processor. In Fig. 2 we show a genuine museum query image which is a fragment of a painting captured before restoration work at the Louvre laboratory. The best match is shown in the same figure and it can be seen that this is the correct parent image but after restoration. (The before restoration image was not in the database). Note that the position of the match is highlighted, a useful facility when dealing with small fragments from ultra high resolution images. This particular parent image is 6328 by 4712 pixels. It should be stressed that the parent image was scanned at 1.47 pixels per mm of the original painting whereas the sub-image query was captured at a quite different resolution, 0.92 pixels per mm. In this test there were over 1000 images in

the database varying in size from 440 000 pixels to 30 000 000 pixels and the retrieval process took 45.8 seconds on a Pentium III 600 Mhz PC.

In order to evaluate the performance of the MCCV algorithm when sub-image quality is less than ideal we use the test database of 1000 images described above. Sets of query sub-images were selected at random from the test database and modified in a variety of ways and with varying strengths. Four strength levels are used for each modification type. Since the sub-images used in the tests were taken directly from the parent, it was important to try and simulate the wide range of possible sub-image queries which could be submitted as a result of different image capture conditions, resolutions and image sizes. The three types of modification chosen to achieve this were as follows.

- **Blurring**: which consisted of the application of a 3×3 averaging window, applied once for the first strength (S1) of blur, three times for the second (S2), five times for the third (S3) and seven times for the fourth strength (S4).
- **Noise addition**: which consisted of adding Gaussian noise to the sub-images. The strength of the noise was determined by the standard deviation, sigma, of the Gaussian noise distribution. A strength of one (S1) corresponds to a sigma of one, a strength of two (S2) corresponds to a sigma of three, S3 corresponds to a sigma of five and S4 to a sigma of seven.
- **Resizing**: which consisted of reducing the size of query sub-images by a specified percentage determined by the strength. A strength of one (S1) corresponds to 0% reduction in image size, S2 corresponds to 20% reduction, S3 to a 40% reduction and S4 to a 50% reduction. The resizing was achieved by averaging pixel values from the original which contribute to pixels in the reduced image.

The evaluation procedure was as follows.

```
Randomly select 100 images, I, from the
database.
For each image in I randomly select a
random sized sub image area and store in
K
For all image operations, O (Blur, noise,
sub-sample, identity)
For each image I_n in K
Apply image processing operation O to
image I_n to produce image J.
Create MCCV feature vector F from image
J.
Compare F with all feature vectors in
database.
Order similarities and record rank infor-
mation.
```

The graphs in Fig. 3 are cumulative graphs of the retrieval results. They show the percentage of correct matches, across the tests, found within the top X hits where X is given by the within rank value on the x coordinate. The legends in the graphs show the strengths at which the algorithms were applied, where S1 is the weakest through to S4 the strongest. Note that for each
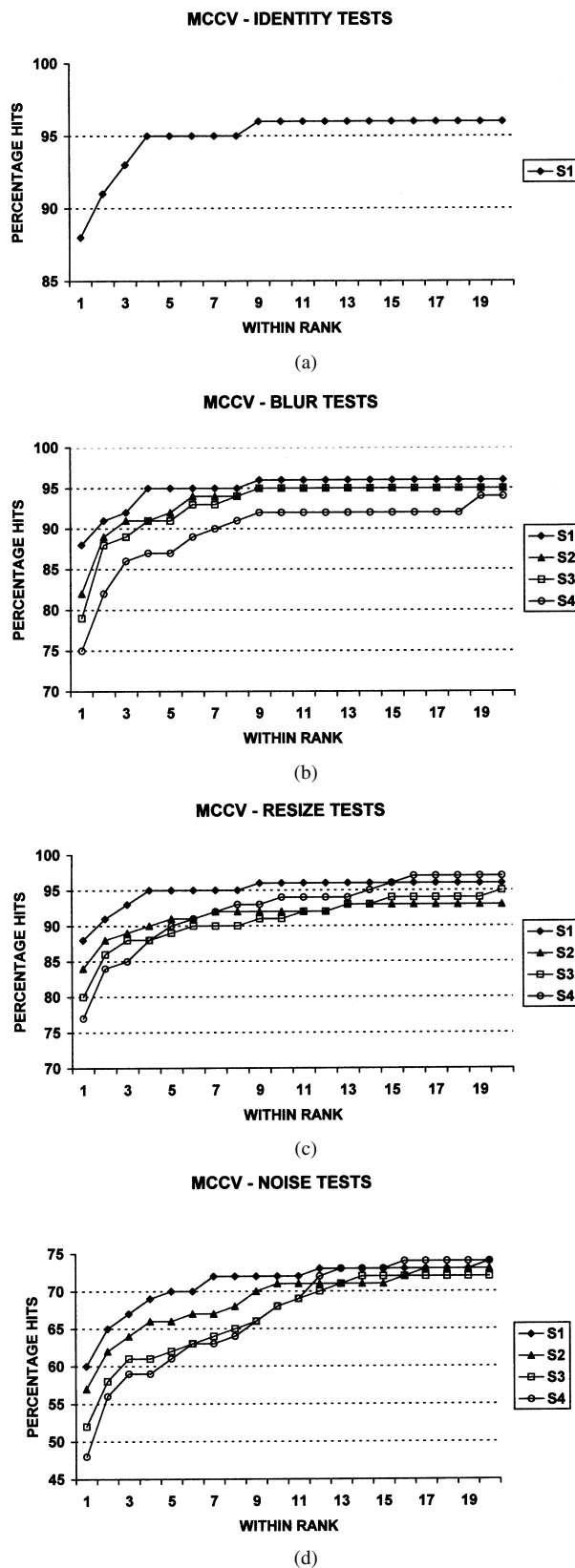

Fig. 3. Retrieval performance graphs for sub-image matching. S1 to S4 indicate the four strengths with which each of the distortions was applied (see text for definitions).

query in these tests there is only one correct match so precision and recall metrics would be less useful in this context.

The first graph (Identity Tests) shows the results when no corruption is applied to the sub-images. It can be seen that 88% of the tests gave the correct match as the best match found and 96% of the tests found the correct match in the top ten best matches. The Blur test shows that even with the highest strengths of blur, the correct matches are found above 90% of the time within the top 10 hits. The robustness to noise is less successful and the correct matches are found in the top 10 hits for only about 70% of the tests. Robustness to resizing is reasonable since, although only about 70% of the tests gave the correct match as the best match when applying the strongest resizing, it can be seen that at least 90% of the tests find the best match in the top ten for all strengths of resizing.

The multi-scale algorithm has been implemented in such a way that any appropriate algorithm applicable to whole images may be used in a multi-scale version without further implementation. Thus, for example, our texture matching algorithm, based on the wavelet transform, is also used for sub-image matching when the required basis for sub-image matching is texture.

B. Query by Low Quality Image

One of the problems which galleries sometimes encounter is being asked "have you got this particular painting or art work in your collection?" when the image submitted as the basis for the query is a very poor and anonymous representation of the original. The image is typically monochrome, perhaps a photocopy of a book image, a fax of a postcard or some other low quality representation. The query images are essentially binary and the problem for *query by example* is that, in order to calculate the similarity to a database image, it is necessary to degrade the database image in a way which depends on the query so that pre-computation of the feature vectors is not straightforward.

Two approaches to the problem were developed. Both involve resizing the images to some standard. The first approach uses a pixel by pixel comparison, resulting in a very slow but relatively high accuracy retrieval. This method is referred to as the slow query by low-quality image (QBLI) method. It is used as a yardstick for comparison to a faster, wavelet-based algorithm. We call this wavelet-based method the fast QBLI method. It is a modification of the slow QBLI method and has the advantage of high retrieval accuracy while maintaining a low computational load. Throughout this section genuine fax images, received by galleries, are used as our query examples.

C. Slow QBLI Method

The slow QBLI method is based on the fact that the low quality images are almost binary in nature. The query image is first converted to be totally binary, and the percentage of black (or white) pixels is computed. The image to be compared to the query is then converted to binary by thresholding the luminance in such a way that the percentage of its black (or white) pixels is similar to that in the binary query image. Both binary images are re-sampled to the same size 64×64 pixels) and a pixel by pixel comparison is made. The number of matching pixels, as a percentage of the total number of pixels, is used as the similarity measure and the database images are presented in decreasing order of similarity.

The slow QBLI algorithm is not ideal for interactive retrievals since it involves a high computational load. Using this algorithm, it is not possible to do any pre-computing without using large amounts of storage to hold multiple binary versions of each database image, corresponding to the range of possible thresholding values. For this reason, all calculations are performed during the retrieval process making this a slow but very accurate process for query by low quality images.

D. Fast QBLI Method

For this method the Pyramid Wavelet Transform (PWT) coefficients [14] are used to compute the feature vector of an image. The reason for using the PWT is that, for a nontextured image, the frequency content is concentrated in the low frequency region, thus image decomposition is needed just for the low frequency band. Moreover, the PWT has a compact feature vector and a low computational load, which is essential for the retrieval algorithm.

Since the quality of the query images is so low that they differ substantially from their originals, applying the PWT to the original image will not produce a feature vector close enough to the feature vector from the query. As in the slow QBLI method, the query image is first converted to a binary image, before the PWT is applied. A similar conversion to binary is applied to each of the database images, choosing a threshold which makes them as close as possible (in terms of the proportions of black and white pixels) to the binary fax image, before the PWT is applied.

This method may seem to be unsuitable for interactive retrieval applications since it requires the feature vectors of the images in the database to be computed in advance. However, due to the compact nature of the wavelet signatures, it is possible to implement the algorithm in an effective manner. The algorithm consists of two steps; binary image creation and feature vector computation and comparison.

As stated earlier, since the query images are almost binary, it is better to compute feature vectors in the binary domain. The query image can be converted to binary by thresholding at the centre of the grey scale range covered by the image. In order for the feature vector of a database image to be compared fairly with the feature vectors from the query, the database image must also be converted to binary. But the choice of threshold is not immediately obvious. For the original database image corresponding to the query, an appropriate threshold is again chosen as the one that produces a binary image with the same percentage of black (or white) pixels as the binary form of the query image. This percentage could be used for all the database images, but it varies from query to query.

How can the percentage be matched if the feature vectors from the binary versions of all the database images are to be pre-computed? Note that since the query image is the target and already effectively binary, it is the database images that must be made as close as possible to the binary query and not vice versa. A solution to this problem is to convert each database image into a set of different binary images corresponding to different percentages of black pixels between 0 and 100%. If sufficient binaries are created, the binary query image will then be very similar to one of these binaries for the original image. Ninety nine bi-

Fig. 4.   Fax images used in this experiment and their originals.

naries were created for each database image corresponding to percentages of black pixels from 1 to 99 in steps of 1%.

However, the binaries do not need to be stored. Calculating the feature vectors for the database involves calculating the PWT for each of the binary images for each image in the data-

base. This is implementable since the PWT is a fast algorithm and, more importantly, the feature vectors for each binary image have only a relatively small number of coefficients. During matching, for each database image, only one of the sets of wavelet coefficients will be used in the comparison,

namely the set associated with the same black pixel percentage as found in the binary query image.

Since the PWT can be applied only on dyadic square images, the binary images are all resized to 256×256. The resizing can also be done before the binary conversion. The PWT algorithm is applied and the image is decomposed into four sub-images (LL, LH, HL and HH). The LL band is decomposed further until the smallest sub-images are of size 4×4, i.e., six levels of decomposition. This results in 19 different sub-images or sub-bands.

Once the wavelet coefficients of a binary image are available, features are computed from each sub-band, resulting in 19 features for each binary image. The mean μ is the energy measure used to compute the features. Let the image sub-band be $W_{mn}(x, y)$ while mn denotes the specific sub-band, m is the decomposition level and $n = 1,2,3,4$ indicates the LL, LH, HL, HH bands respectively, then μ_{mn} is calculated by

$$\mu = \frac{1}{N_{mn}^2} \iint |W_{mn}(x, y)| dx dy$$

where N is the length of a particular sub-band mn. The feature vector f, for a particular binary image is therefore

$$f = [\mu_{mn}], n \neq 1 \text{ except for the coarsest level}, m = 6.$$
$$= [\mu_{1,2}, \mu_{1,3}, \mu_{1,4}, \mu_{2,2} \ldots, \mu_{6,1}, \mu_{6,2}, \mu_{6,3}, \mu_{6,4}].$$

The feature vectors for the database images will have $99 \times 19 = 1881$ coefficients, although only 19 will be used for comparison in each retrieval task. The distance classifier used is the Euclidean minimum distance. The distance between 2 features, i and j, is given by

$$d(i, j) = \sqrt{\sum \sum [\mu_{mn}^{(i)} - \mu_{mn}^{(j)}]^2}.$$

Once the distances are computed, the images will be retrieved in order of increasing distance from the query image.

Experiments were conducted using as the query, each of the 20 genuine fax images with a range of qualities as the query, and a database consisting of 1058 images of various types and sizes, including the original images of the 20 fax images. The fax images and their originals are shown in Fig. 4. The evaluation is based on the ability of the algorithm to retrieve the original image when the fax version of the original is used as the query. The results for the fast QBLI algorithm in Table I show the retrieval position of the original image among the 1058 database images, using Daubechies [15] family wavelets with 8 vanishing moments as the wavelet bases. Fig. 5 shows an example of three retrieval results using the fast QBLI algorithm.

Table I also shows the results obtained by using a basic *query by example* retrieval with the same PWT features but calculated from the raw query and database images without the initial binarization stage. These are the sort of results one might expect from a standard CBIR system without special algorithms for special cases. It can be seen that the basic *query by example* algorithm is particularly poor for these low quality queries, but the retrieval results obtained using the fast QBLI algorithm are very encouraging. All the original images are retrieved within the top 5. This is a good result considering the poor quality of

TABLE I
RETRIEVAL RESULTS USING 20 FAX IMAGES ON A DATABASE OF 1058 IMAGES

Query Image No.	Rank of Original		
	Basic PWT	Slow QBLI	Fast QBLI
1	104	1	1
2	369	1	1
3	15	1	1
4	21	1	3
5	272	1	1
6	130	1	1
7	258	1	1
8	2	1	3
9	502	1	1
10	302	20	2
11	603	1	1
12	299	1	1
13	60	1	1
14	495	1	4
15	500	1	2
16	339	1	1
17	15	1	2
18	264	1	4
19	1	1	1
20	1	1	1

TABLE II
COMPARISON OF COMPUTATION TIMES BETWEEN THE THREE ALGORITHMS

Query	Basic PWT	Slow QBLI	Fast QBLI
Time/s	1	130	1

some of the fax images. The results suggest that the distances between the fax images and their originals in our feature space are very close and should still produce good results for a larger image database. Different wavelet bases were also tested in this experiment, and it was found that the choice of wavelet base (Haar and Battle-Lemarie family [15]) has little effect on retrieval result. However the Daubechies wavelet gives a slightly better result, probably because it is compactly supported in both the time and frequency domain.

Table I also shows the results for the slow QBLI method. As expected, the slow QBLI algorithm gives very good retrieval results. All the originals were returned as the first match, except for one case only, which is because that particular fax image and its original were of different object-to-background ratio, hence resulting in rather dissimilar images after binarization. Table II compares the average time taken for retrieving images from the database of 1058 images with the basic PWT algorithm and the slow and fast QBLI algorithms. The times are for a 700 MHz Xeon processor. From Tables I and II it can be seen that the fast QBLI algorithm almost equals the slow QBLI method in terms of retrieval performance, but involves a much smaller computational load.

It is also important to note that the PWT algorithm applied on binary images helps to minimize computation time. To sum

Fig. 5. Fax images and their top six retrieved images.

up, it can be said that the fast QBLI method integrates the high accuracy of the slow QBLI method with the low computational load of the basic PWT method giving an effective approach to query with low quality example.

The two approaches to *query by low-quality image* give very good retrieval accuracy, although the speed of the fast QBLI method makes it much more suitable for interactive use. The fast QBLI method illustrates the importance of the wavelet-based
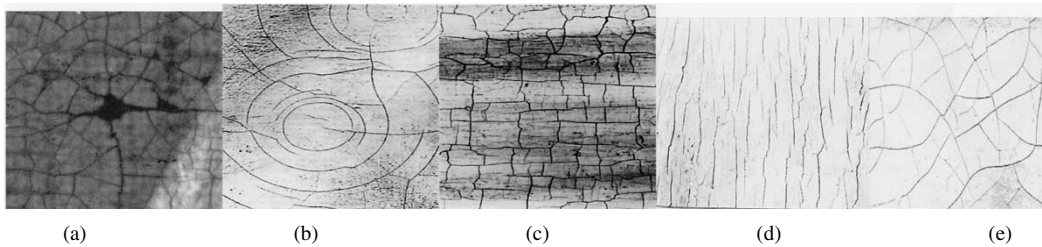
Fig. 6.   Crack classes: (a) spider-web, (b) circular, (c) rectangular, (d) unidirectional, and (e) random.

feature extractor in this application, where in PWT, we have a very fast algorithm and compact feature vectors. These two constraints are important in using the fast QBLI method where multiple feature extractions as well as multiple feature vectors are necessary for each database image. Other feature extraction techniques are either slow or have large feature vectors.

Another important observation from this part of the project is that although we used fax images in the evaluation, the method proposed is an excellent way to search using low quality images in general. This is shown by the fact that the basic PWT technique gives very poor retrieval accuracy when used in this kind of problem and suggests that other standard feature extraction methods will also fail to deliver a good performance if not accompanied by a tailored pre-processing stage.

E. Comments

These two retrieval algorithms are examples of the way in which the general purpose CBR facilities were augmented with tools for specific retrieval tasks. Other special purpose facilities are also being developed, some of which are concerned specifically with the paint crack patterns (craquelure) which can occur in paintings and which are of particular interest for conservation and monitoring. In this context, a new algorithm has been developed [16] in order to process two unusual query types: "find images with cracks like this image" and "find images with this class of cracks". The overall goal was to analyze sub-regions of the images in order to classify the cracks in the paint layer into one of a few classes [17]. Example classes are shown in Fig. 6.

We used a morphological top-hat operator [16] and adaptive grid-based thresholding [20] to detect the cracks. After thinning, the cracks are represented as chain code [21] and statistics gathered on the crack segments. These include lengths, straight line to actual length ratio (LTRL), directionality and orientation histogram [22].

Fig. 7 illustrates the result of a fuzzy k-means clustering of two of these features. Cluster FKM2 represents cracks which are very directional and straight for example. By combining cluster scores for all the features a crack type score can be obtained. The issue of scale is complex because an area of cracks can cover very large areas of an image. Further details of this aspect of the work can be found in [22].

IV. EXAMPLE OF CROSS COLLECTION SEARCHING

In Section II, we described how the integrated approach to cross-collection searching has been implemented and in the previous section we concentrated on a description and evaluation of two of the novel CBR algorithms. In this section we show by
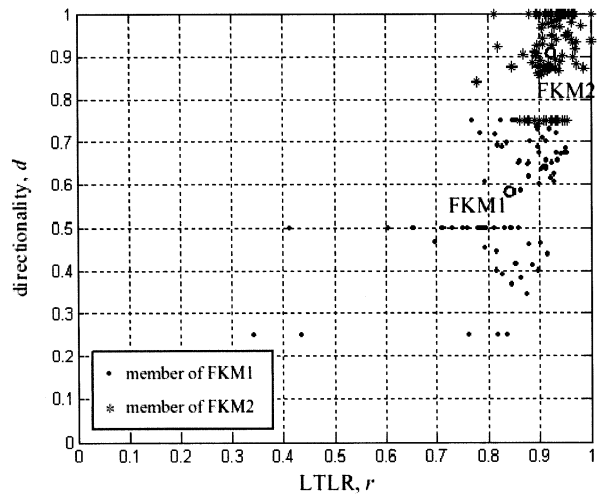

Fig. 7.   Results of fuzzy K-means clustering of two crack features. FKM1,2 show cluster centres.

example how seamless access to the multiple distributed collections is achieved. Fig. 8 shows a query combining image content and metadata searching across the Victoria and Albert Museum collection. The query uses the MCCV algorithm to find larger images, labeled as "Photograph" and containing the specified query as a sub-image. The results, in Fig. 9, show that the system finds the photograph containing the query image as well as other photographs containing similar features and colors.

In Fig. 10, the results of a second query are shown. Three collections were specified to be searched and metadata values were selected from a list showing only those terms which can be mapped to fields in each of the selected collections. The unified results contain images returned from each of the selected collections.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we have presented a system for seamless cross-collection content and metadata based searching of museum image collections. The development and use of new and emerging standards have provided an effective approach to interoperability. Novel special purpose algorithms developed for handling particular aspects of art image retrieval have been described and evaluated. It has been shown that the multi-scale color coherence vector (MCCV) technique can provide effective sub-image retrieval and is also applicable when the sub-image and target are captured at different resolutions. The wavelet transform has formed the basis for a facility for query by fax and by other low quality images, and work has been presented on query by crack type. These, and other

Fig. 8. Sample ARTISTE query.
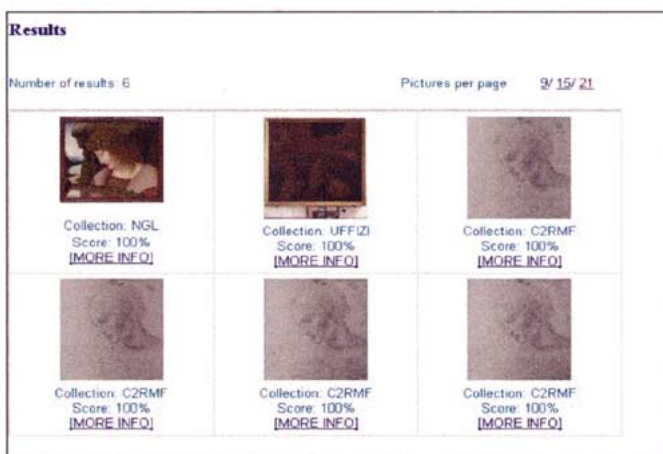

Fig. 9. Sample ARTISTE query results.

Fig. 10. Sample ARTISTE cross collection query results.

specialist algorithms, combined with more traditional content and metadata based facilities have been integrated into an effective framework for cross-collection searching.

The work of the Artiste project is being continued under a further European Project, Sculpteur, in which the database is being extended to hold 3-D models of museum artefacts and the system is being integrated with emerging semantic web tech-nology to enrich the knowledge base on the collections from the museums and galleries.

REFERENCES

[1] W. Vaughan, "Automated picture referencing: Another look at Morelli," in *Computers and the History of Art*. New York: Harwood, 1991, vol. 2, pp. 7–18.

[2] B. Holt and K. Weiss, "The QBIC project in the department of art and art history at UC Davis," in *The 60th ASIS Annu. Meet. 1997*, vol. 34, Digital Collections: Implications for users, funders, developers and maintainers, 1997, pp. 189–195.

[3] A. Del-Bimbo, M. Mugnaini, P. Pala, and F. Turco, "Visual querying by color perceptive regions," *Pattern Recognit.*, vol. 31, no. 9, pp. 1241–1253, 1998.

[4] RDF Resource Description Framework [Online]. Available: http://www.w3.org/RDF/

[5] Dublin Core Metadata Initiative [Online]. Available: http://www.dublin-core.org/

[6] J. Cupitt and K. Martinez, "VIPS: An image processing system for large images," *Proc. SPIE*, vol. 2663, pp. 19–28, 1996.

[7] P. Allen, M. Boniface, P. Lewis, and K. Martinez, "Interoperability between multimedia collections for content and metadata-based searching," in Online Proceedings of the World Wide Web Conference, Hawaii, 2002.

[8] MySQL: Open Source Database [Online]. Available: http://www.mysql.com

[9] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 1349–1380, Dec. 2000.

[10] M. J. Swain and D. H. Ballard, "Color indexing ," *Int. J. Comput. Vis.*, vol. 7, no. 1, pp. 11–32, 1991.

[11] G. Pass, R. Zabih, and J. Miller, "Comparing images using color coher-ence vectors," in *Proc. ACM Multimedia*, 1996, pp. 65–73.

[12] M. Westmacott, P. Lewis, and K. Martinez, "Using color pair patches for image retrieval," in *Proc. IS&T 1st Eur. Conf. Color in Graphics, Imaging and Vision*, April 2002, pp. 245–247.

[13] S. Chan, K. Martinez, P. Lewis, C. Lahanier, and J. Stevenson, "Handling sub-image queries in content-based retrieval of high resolution art im-ages," in *Proc. Int. Cultural Heritage Informatics Meeting*, vol. 2, Sept. 2001, pp. 157–163.

[14] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.

[15] Wavelab Toolbox ver 802 for Matlab [Online]. Available: http://www-stat.stanford.edu/wavelab/

[16] F. S. Abas and K. Martinez, "Classification of painting cracks for con-tent-based analysis," *Proc. SPIE*, pp. 149–160, Jan. 2003.

[17] S. Bucklow, "A stylometric analysis of craquelure," *Computers and Hu-manities*, vol. 31, pp. 503–521, 1989.

[18] F. Meyer, "Iterative image transformations for an automatic screening of cervical smears," *J. Histoch. Cytochem*, vol. 27, 1979.

[19] I. Giakoumis and I. Pitas, "Digital restoration of painting cracks," in *Proc. IEEE Int. Symp. Circuits and Signals*, May-June 31-3, 1998, pp. 269–272.

[20] N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–69, 1979.

[21] H. Freeman, "Boundary encoding and processing," in *Picture Pro-cessing and Psychopictories*. New York: Academic, 1970.

[22] F. S. Abas and K. Martinez, "Craquelure analysis for content-based re-trieval," in *Proc. IEEE 14th Int. Conf. Digital Signal Processing*, July 2002, pp. 111–114.

[23] Y. Rui, T. S. Huang, and S.-F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *J. Vis. Commun. Image Represent.*, vol. 10, pp. 39–62, 1999.

[24] J. P. Eakins and M. E. Graham. Content Based Image Retrieval: A Report to the JISC Technology Applications Programme. [Online]. Available: http://www.unn.ac.uk/iidr/report.html

[25] G. Qui, "Color image indexing using BTC," *IEEE Trans. Image Processing*, vol. 12, pp. 93–101, Jan. 2003.

[26] B. Zhu, M. Ramsey, and H. Chen, "Creating a large scale content-based airphoto image digital library," *IEEE Trans. Image Processing*, vol. 9, pp. 163–167, Jan. 2003.

[27] N. Arica and F. T. Yarman-Viral, "A compact shape descriptor based on the beam angle statistics," in *Int. Conf. on Image and Video Retrieval*, vol. 2728, Lecture Notes in Computer Science, 2003, pp. 152–162.

[28] N. Sebe, M. S. Lew, X. Zhou, T. S. Huang, and E. M. Bakker, "The state of the art in image and video retrieval," in *Int. Conf. on Image and Video Retrieval*, vol. 2728, Lecture Notes in Computer Science, 2003, pp. 1–8.

[29] P. Enser and C. Sandom, "Toward a comprehensive survey of the semantic gap in visual image retrieval," in *Int. Conf. Image and Video Retrieval*, vol. 2728, Lecture Notes in Computer Science, 2003, pp. 291–299.

[30] K. Barnard and D. Forsyth, "Learning the semantics of words and pictures," in *Proc. Int. Conf. Computer Vision*, vol. 2, 2001, pp. 408–415.

[31] R. Tansley, C. Bird, W. Hall, P. Lewis, and M. Weal, "Automating the linking of content and concept," in *Proc. ACM Multimedia Conference*, 2000, pp. 445–447.

[32] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Jain, and C. F. Shu, "The virage image search engine," *Proc. SPIE*, pp. 76–87, 1996.

[33] W. Y. Ma and B. S. Manjunath, "Netra: A toolbox for navigating large image databases," *Proc. IEEE Int. Conf. on Image Processing*, no. 1, pp. 568–571, 1997.

[34] J. R. Smith and S.-F. Chang, "Querying by color regions using the visualSEEK content-based visual query system," in *Intelligent Multimedia Information, Retrieval*, M.T Maybury, Ed. Menlo Park, CA: AAAI Press, 1996.

[35] Y. Rui, T. S. Huang, M. Ortega, and S. Mehotra, "A power tool in interactive content-based image retrieval," *IEEE Trans. Circuits, Syst., Video Technol.*, Sept. 1998.

[36] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Content based manipulation of image databases," *Int. J. Comput. Vis.*, vol. 18, no. 3, pp. 233–254, 1996.

[37] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "The QBIC project: Querying images by content, using color, texture, and shape," *Proc. SPIE*, pp. 173–187, 1993.

[38] M. F. A. Fauzi and P. H. Lewis, "Texture based image retrieval using multiscale sub-image matching," in *Proc. SPIE Conf. Image and Video Communications and Processing*, 2003, pp. 407–416.

[39] Zing: Z39.50 International: Next Generation [Online]. Available: http://www.loc.gov/z3950/agency/zing/zing-home.html

Kirk Martinez received the B.Sc. in physics (reading) and the Ph.D. degree in image processing from the University of Essex, U.K.

He is a Senior Lecturer in the School of Electronics and Computer Science, University of Southampton, Southampton, U.K. His previous research on high resolution imaging, Web access to images and content-based retrieval were all related to cultural heritage. His current research is semantic and content-based retrieval of art, augmented reality, and sensor networks.
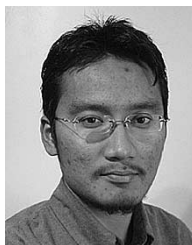
Fazly Salleh Abas received the B.S. degree in electronic and electrical engineering in 1999 from the University of Strathclyde, Glasgow, U.K. Currently with the School of Electronics and Computer Science, University of Southampton, where he is pursuing the Ph.D. degree in computer science.

His research interests include pattern recognition, computer vision, and content-based image retrieval.

Mohammad Faizal Ahmad Fauzi received the B.Eng. degree in electronic engineering from Imperial College of Science Technology and Medicine, London, U.K., in 1999. He is currently pursuing the Ph.D. degree in computer science at the School of Electronics and Computer Science at the University of Southampton, Southampton, U.K.

His current research interests include image processing and analysis, content-based image retrieval, wavelet theory and application, and pattern recognition.
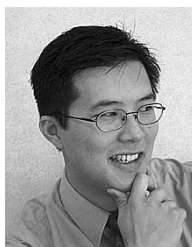
Stephen C. Y. Chan received the B.Sc and Ph.D. degrees from the University of Southampton, Southampton, U.K., in 1996 and 2001, respectively.

Currently, he is a Senior Design Engineer at Panasonic MCDE, U.K., developing embedded software for next generation mobile communication products. His main work at the Electronics and Computer Science Department at the University of Southampton was related to video analysis for semantic-level features to aid scene interpretation and understanding.

Paul H. Lewis received the B.Sc. degree in physics from Imperial College, London, U.K., and the Ph.D. in physics from London University in 1972.

He is a Senior Lecturer with the Intelligence, Agents, Multimedia Group in the School of Electronics and Computer Science at the University of Southampton, Southampton, U.K. His main research interests are in the area of image, video and semantic analysis, and their application to multimedia information handling and data mining.

Matthew J. Addis received the B.A. degree in physics from the University of Oxford, Oxford, U.K., and receive the Ph.D. in physics from the University of Southampton, Southampton, U.K.

He now works at IT Innovation where his roles include technical and project management in collaborative R&D and commercial consulting contexts. Interests include designing, developing and implementing e-Science, e-Business, and cultural heritage systems using Semantic Web, Agent, Grid, and workflow technologies.

Mike J. Boniface received the B.Eng. in multimedia communications.

He joined IT Innovation in 2000 after several years at Nortel working in an ISO9000 environment on internal services using Web access to database systems. He has experience of the full project lifecycle with expertise in requirements capture, software testing, and rapid development methods. He has interests in software engineering best practice and leads several major software development activities in IT innovation.

Alison Stevenson received the MA degree in English literature and film studies and the M.Sc. degree in information technology.

She is a Research Engineer at IT Innovation with a strong interest in cultural and heritage computing. She has research interests in the application of ontology languages, and other semantic Web technologies to promote interoperability and enable compliance with international standards of data retrieval. She is also responsible for the application of an ontological approach in several Grid projects.

Christian Lahanier received the Ph.D. degree in chemistry.

He was in charge of the Physics Department of the Laboratoire de recherche des musées de France for 15 years. He has developed X-ray technologies for the nondestructive analysis of art objects. In 1983, he became the Head of the Laboratory as Chief Engineer in charge of a European telematics project on Image Technology (NARCISSE). He set up a multilingual documentary information system and an ultra-high definition scanner.

Paul Grimwood received the M.Sc. degree in computer studies, the Ph.D. degree in neuroscience, and the B.Sc. degree in physiology.

He is a Research Engineer at the IT Innovation Centre in Southampton, Southampton, U.K. He is interested in the integration of computing and neuroscience, in particular for the advancement of understanding the mechanisms of learning and memory.

James Stevenson is the Photographic Manager of the Victoria and Albert Museum. He has worked at the V&A for 10 years and previously was Chief Photographer at the National Maritime Museum. His main area of interest is managing the large photographic production studio at the V&A from analogue to an entirely digital environment.