SMP and Cluster Architectures for Retrieval of Images in Digital Libraries

O. Kao
Department of Computer Science, TU Clausthal
Julius-Albert-Strasse 4, D-38678 Clausthal-Zellerfeld, Germany
e-mail: okao@informatik.tu-clausthal.de

Abstract: This paper presents an overview over parallel architectures for the efficient realisation of digital libraries by considering image databases as an example. The state of the art approach for image retrieval uses a priori extracted features and limits the applicability of the retrieval techniques, as a detail search for objects and for other important elements can't be performed. Well-suited algorithms for dynamic feature extraction and comparison are not often applied, as they require huge computational and memory resources. Integration of parallel methods and architectures enables the use of these alternative approaches for improved classification and retrieval of documents in digital libraries. Therefore implemented prototypes on a symmetric multiprocessor (SMP) and on cluster architecture are introduced in the paper. Performance measurements with a wavelet-based template matching method resulted into a reasonable speedup.

1. Introduction

Multimedia documents offer significant advantages in information representation and communication. Further the capacities and the performance of I/O devices, memories, networks, CPUs etc. are improving continuously and are nowadays widely available. Nevertheless 90% of all information is still on paper. One of the reasons is the lack of reliable methods for content analysis of the different media types, thus basic mechanisms and technologies for the management of multimedia data are not available yet. Many existing relational and object-oriented databases handle multimedia objects as BLOBs (*Binary Large Objects*) [1,6] and describe their content by a manually compiled and limited set of keywords. The retrieval is then realised by a full text search in the assigned set of keywords. A main disadvantage of this approach is given by the reduction of the complex media content on few keywords.

This paper focuses on image retrieval, as it is one of the important components in digital libraries. A complete image description requires the extraction and analysis of various information regarding the objects, persons, textures etc. in the scene. Subsequently a relationship between the image objects and the real world entities has to be established.

The possible image information can be classified into following groups:

- Raw image data is the matrix with the pixel colour values, which can be stored in various file formats.
- Technical information describes the image resolution, number of used colours, file format etc.
- Results of the image processing and analysis: this group encloses all extracted image features, like objects and regions, statistical characteristics and topological data, which define the spatial relationship between the image objects.
- Knowledge based information describes the relationship between the image elements and the real world entities, for example who or what is shown on the image etc.
- World oriented information concerns usually the acquisition time and date, photographer, etc.

The raw image data, the technical and world oriented information can be represented by well-known data structures and stored in existing databases. The user can either browse the database (similar to a printed image catalogue for example in an art gallery) or search the database by entering keywords like the name of the artist or the picture, date etc. These are compared with the technical and world oriented information. Examples for such *databases with images* are widely available on the Internet, for example as a part of e-Business and entertainment sites, web presentations of museums, art galleries, trademarks etc.

These primitive retrieval and annotation methods have many disadvantages in terms of time-intensive and difficult selection of appropriate keywords. Therefore an image database should support improved querying, retrieval and annotation methods, which enable a content-based similarity search in a general set of images.

Interfaces for the query of an image database include visual methods like query-by-pictorial-example (QBPE), query-by-painting/sketching or standard methods like browsing in a given set of sample images [5,7].

2. Image Retrieval

The state-of-the-art approach [1] for the creation and retrieval of image databases is based on the extraction and comparison of a priori defined features. These are directly derivable from the raw data and represent properties related to the dominant colours in the image and their distribution, important shapes, textures and consider the global layout. The extracted features can be combined and weighted in different ways resulting into logical (advanced) features, which represent the image content on a higher abstraction level. New works in this area consider the image semantics and try to define and integrate different kinds of emotions (e.g. [8]). The similarity degree of a query image and the target images is determined by calculation of a distance (for example Euclidian Distance L^2) between the corresponding features. The result is a sorted list with *n* distances, where *n* is the number of stored images. The first *k* elements correspond to the most similar images in the database, thus the raw data of those images is presented to the user as retrieval hits.

Satisfactorily recognition rates are reported for a number of implemented image retrieval systems and prototypes. The measurements are often performed on a small number of image classes. Thereby the intra-class distance is small (for example landscape images, faces, etc.) and the inter-class distance is large. Acceptable system response time are achieved, because no further processing of the image raw data is necessary during the retrieval process resulting into immense reduction of computing time. The easy integration in existing database systems is a further advantage of this approach.

Extraction of simple features results in disadvantageous reduction of the image content. Important details like objects, topological information etc. are not sufficiently considered in the retrieval process, thus a precise detail search is not possible. Furthermore, it is not clear, whether the known, relatively simple features can be combined in the right way for the retrieval of all kinds of images.

(a)

(b)

Figure 1. *Example for image retrieval with a priori extracted features: (a) Query image; (b) Retrieval results ordered by relevance*

An example for this kind of retrieval is given in Figure 1. The query image (Figure 1(a)) shows a person on a beach. The query starts with the application of the wavelet transformation on the sample image. The result is a set of wavelet coefficients (c_0, c_1, …, c_{63}) describing the image content (see [7] for a detailed description of this approach). Subsequently the Euclidian differences between the vector of the query image and the corresponding vectors of all images in the databases are calculated. Each of these results gives an indication about the similarity of the compared images. Thereby the background information predominates the image, thus a major part of the wavelet coefficients describes these parts of the scene. The person is not sufficiently considered, thus the best hits (Figure 1(b)) are similar beach scenes, but not images containing the person.

For the realisation of queries like "Show all images containing the marked object X" more powerful methods with dynamic feature extraction are necessary.

Definition: *Image Retrieval with Dynamic Extracted Features -- short dynamic retrieval -- is the process of analysis, extraction and description of any manually selected image elements, which are subsequently compared to all image sections in the database*.

An example for this operation is given by template matching. It is a basic image processing technique for the comparison between the selected region of interest and the corresponding image sections. The region of interest is represented by a minimal bounding rectangle (MBR) and subsequently correlated with all images in the database. Colour values, texture features, contours etc. build the fundaments for the similarity measurements. These attributes can be combined and weighted with heuristic determined coefficients. Thereby distortions caused by rotation and deviations regarding the size, colours, etc. have to be considered. In contrast to the state-of-the-art approach image features and attributes can't be a priori defined and extracted, thus all images in the database have to be processed. Figure 2 depicts this approach.

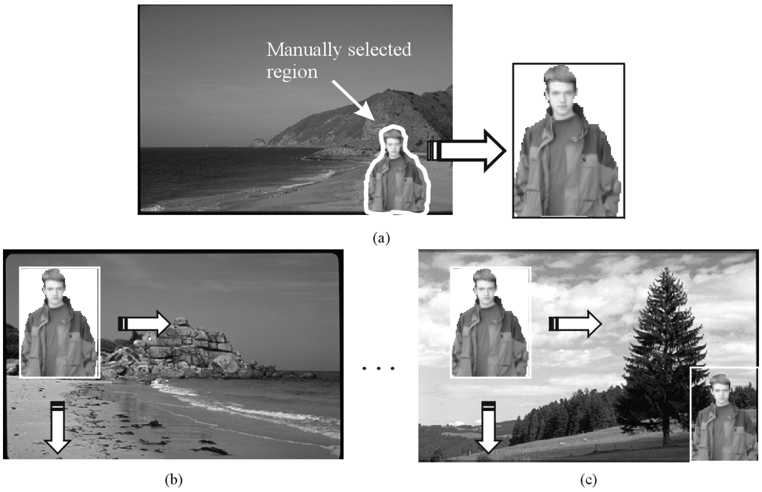

Figure 2. *Example for image retrieval with dynamic extracted features: (a) Creation of a region of interest; (b) Search for the region in an unsuitable image; (c) Person found in another image by template matching*

The query starts with a selection of a region of interest, which is subsequently represented by a minimal bounding rectangle. The background in the MBR is set with a neutral colour and not included in the image analysis. In the next step the wavelet transformation is applied on this search mask and a number of wavelet coefficients describes the mask content. These values are the fundaments for the object search in all image sections in the database. A number of weighted metrics are used for the comparison of the similarity between the search mask and the underlying image section. The global layout is not considered any more, thus the image in Figure 2(b) is not marked as retrieval hit, although the query and the target image represent beach scenes. In opposite thereof a forest image *with* the same person is correctly identified as a similar image.

This example shows, that dynamically extracted features improve and extend the image retrieval process. Specific image elements are sufficiently considered and may be found in different environments. But, more sophisticated -- especially scale and rotation invariant -- methods for the region description and comparison are necessary in order to enable a more efficient analysis and to increase the retrieval accuracy.

3. Parallel architectures for dynamic image retrieval

The cost for the improved flexibility and retrieval quality is the immense computation effort [2,3], which exceeds the capabilities of computer architectures with single processing element (PE). Furthermore we are already flooded with digital image material, which is getting bigger day for day. Digital photo and video technologies can be found in commerce, research and at home, a lot of pictorial information is available on the Internet. The scope of document management systems, digital libraries, photo archives in hospitals, authorities and companies, satellite images are rapidly growing. These and many other applications produce Petabytes of pictorial material per year. Therefore we examine the properties of different parallel architectures for the realisation of digital multimedia libraries in general and image databases in particular. From the viewpoint of the database community the parallel architectures are classified into [4]:

- Shared nothing architectures consist of replicated, interconnected computers. The communication and synchronisation is realised by message passing.
- Shared disk architecture: in this case all nodes of the parallel computer share the I/O subsystem.
- Shared everything architectures: all PEs have access to the memory and the I/O subsystem, thus shared variables can be used for the communication and synchronisation.

The important question is now, which computer architecture is suitable for the implementation of digital libraries. We analysed the basic retrieval operations and showed, that not the computing power but the huge data volume and their communication is the limiting factor. The applied operations may vary significantly regarding the processing time: some of them like edge detection are relatively simple and require only a split second computing time, whereas a template matching operation with many parameters necessities more than a minute processing time per image. A retrieval process ends with a creation of a hit ranking, thus a one-time unification of all sub results is necessary.

In the case of shared disk and shared everything architectures the I/O subsystem and the transfer of the huge image data to the main memory might be a bottleneck. On the other hand the workload distribution and the synchronisation can be realised easily and efficiently. Cluster architectures have an advantage that each node has an own I/O subsystem, thus the transfer effort is shared by a number of nodes. Moreover, the reasonable price per node enables the creation of systems with a large number of processing elements. Open problems concern workload balancing, synchronisation and data distribution as well as the general cluster problems like missing single system image and the large maintenance effort.

We implemented two prototypes for SMP and cluster architectures and performed extensive measurements with a large image library and different operators for image transformation, analysis and retrieval. Following subsection gives a very short overview over the basic principles of the prototypes.

4. Prototypes

Widespread examples for shared everything architectures are symmetrical multiprocessors (SMP). The SMP parallelisation of image processing operators is usually achieved by data partitioning: an image is separated into independent sections, which are subsequently distributed over the available PEs. This approach can't be used for the parallelisation of all retrieval methods, as some operators consider the dependencies between all pixels, thus an image separation could lead to falsification of the analysis results. Therefore our SMP prototype is based on task parallelisation. The efficiency and the speedup of the parallel application depend significantly on the characteristics of the I/O subsystem. Time intensive data transfers decrease the performance gain. In order to reduce the idle times of the PEs and minimise the conflicts between the PEs and the shared memory we developed a two-layer model, which is shown in Figure 3.

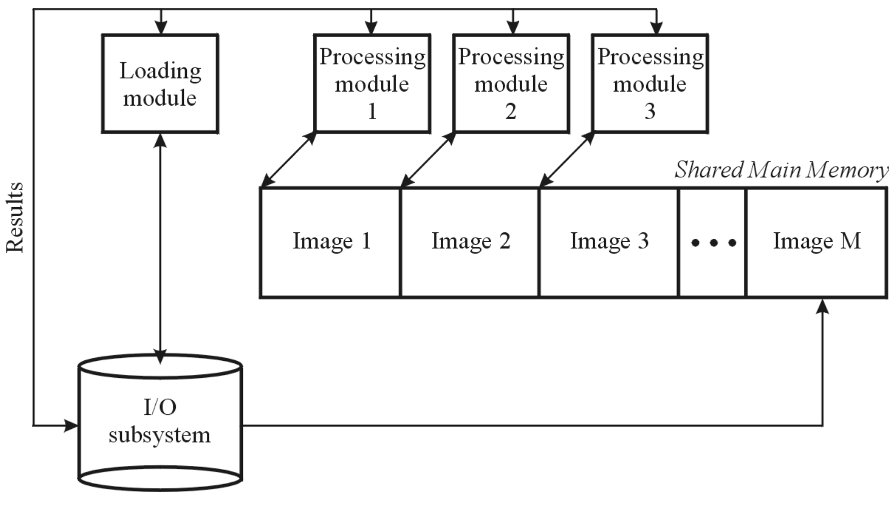

Figure 3. *Shared memory model for image retrieval*

The first layer contains only one loading module, which transfers the images from the file system into a particular area of the main memory. A dedicated PE executes this process permanently. The second layer consists of processing modules, which are assigned to the other available PEs in the system. Each processing module requests from the loading module a pointer on the next image in the memory and information about the image size. Subsequently the analysis operations are applied on this memory block and the results are stored in an index structure or in a file. While the PEs analyse the image the loading module fetches new images and stores them in a memory buffer. In the last phase the results of the image analysis are collected and evaluated, thus an image ranking can be generated.

A cluster is a parallel or distributed system consisting of collection of interconnected whole computers and used as a single, unified computing resource. Clusters of symmetric multiprocessors -- so called CLUMPs -- combine the advantages of two parallel paradigms: an easy programmable SMP model with the scalability and data distribution over many nodes. Therefore we constructed a CLUMPS-based prototype for efficient, affordable image retrieval with dynamic feature extraction. The nodes are subdivided in respect to their functionality into four groups:

- Query stations host the web-based user interfaces.
- Master node controls the cluster, receives the query requests and broadcast the algorithms with the search parameters as well as the example image and features to the computing nodes.
- Computing nodes: each of these nodes contains a part of the existing images and executes the feature extraction and the comparison with the local data. The results are sent to the media server.
- The media server has multiple functions. Firstly it is a redundant storage server and contains the whole image databases. Secondly it receives and compares the results of all computing nodes and sends the *k* best hits to the user on the query station.

A relational database stores static, a priori extracted features together with the corresponding index structures. A component called transaction manager analyses the queries. If only a priori extracted features are considered, the time intensive search of all images is replaced by a relatively simple next neighbours search on the index structures. Otherwise all images in the system must be processed. In this case the transaction manager directs the query to the distribution manager, which is installed on the master node and initiates the search on the computing nodes. A result manager combines all sub results, selects the best hits and passes finally the retrieved images to the user on the query station. A graphic representation of this architecture is shown in Figure 4.

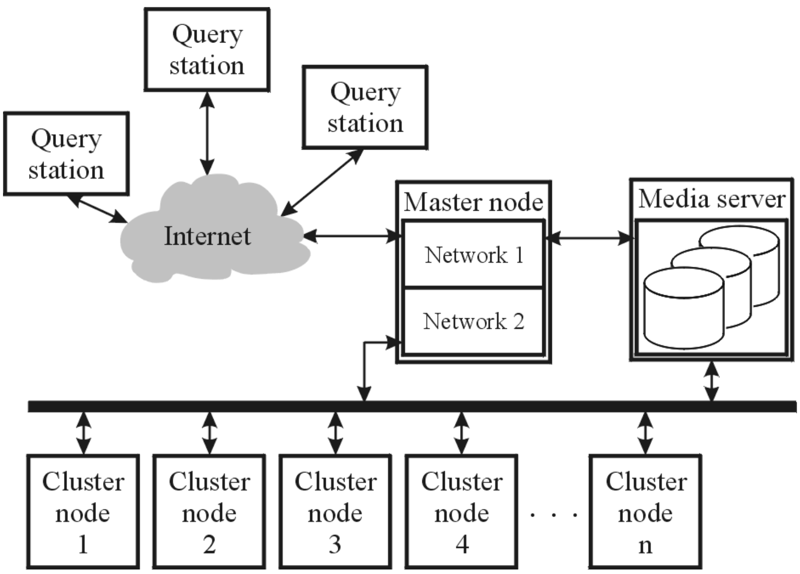

Figure 4. *Cluster Architecture for image retrieval*

5. Performance Measurements

Performance measurements are performed on a test database of with more than 100000 JPEG images. Due to the huge computation time the test are executed on a restricted set of 10000 images by using a wavelet-based template matching operator and a region of interest containing the person showed in Figure 2(a).

The SMP processing model is implemented on a parallel DEC Alpha 4100 Server with four PEs (Alpha 21164A, 600 MHz) and 1 GByte main memory, running Digital UNIX operating system. The largest speedup value of 3,56 is obtained with three processing threads and a single loading thread. The speedup is decreasing, if more threads are used. Further the saturation of the connection between the I/O sub system and the main memory reduces the performance gain, thus further modifications are necessary.

The hardware platform for the cluster-based image retrieval is a Beowulf cluster with six computing nodes, each of them with Dual Pentium III 667 MHz and 512 Mbytes main memory, all connected over a 100 MBit Ethernet. The reference value for the calculation of the speedup is obtained on a single cluster node. The images are equally distributed over all nodes, thus the workload is nearly equally shared over all nodes. However, small differences (less than 5%) between the processing times of the nodes are noted, which possibly result from caching effects and operating system activities.

The processing of 10000 images by using all twelve PEs necessities approximately 15071 seconds (251 minutes). This corresponds to a speedup of nearly 5. Thereby the object mask considers each 10th position in an image row. Much better retrieval results are achieved, when the mask slides pixel for pixel and examines all possible image sections. But this results in an immense compute time, thus much larger parallel systems are necessary in order to obtain acceptable system response times.

6. Conclusions

Parallel architectures can improve the organisation of digital libraries; enable the use of more powerful retrieval and analysis methods and significantly speedup the processing. Dynamically extracted features consider only manually selected image elements, which are subsequently compared to all image sections in the database and thus enable a precisely object search. Satisfactorily performance results can be achieved with clusters consisting of low-cost PCs.

Feature work includes the development of suitable scheduling strategies for cluster architectures. A dynamic, temporal or permanent migration of the images between the nodes balances the workload. Moreover, suitable image operators for dynamic feature extraction and similarity metrics are necessary.

References

1. W.I. Grosky, R. Jain, R. Mehrota (Eds.): The Handbook of Multimedia information management, Prentice Hall, 1997, 365-404
2. W. Klas, K. Aberer: Multimedia and its Impact on Database System Architectures. In P.M.G. Apers, H.M. Blanken, M.A.W. Houtsma (Eds.): Multimedia Databases in Perspective, pp 31--62, Springer, 1997.
3. O. Kao: Towards Cluster Based Image Retrieval. In Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA), 2000, pp 1307-1315, CSREA Press
4. A. Reuter: Methods for parallel execution of complex database queries. Journal of Parallel Computing, Volume 25, pp 2177-2188, 1999.
5. J.Ashley et. al.: Automatic and semi-automatic methods for image annotation and retrieval in QBIC. In Proceedings of Storage and Retrieval for Image and Video Databases II, pp 24--35, 1995.
6. S.Khoshafian, A. B. Baker: MultiMedia and Imaging Databases, Morgan Kaufmann Publishers, 1996.
7. O. Kao, I. la Tendresse: CLIMS - A system for image retrieval by using colour and wavelet features, Proceedings of the First Biennial International Conference on Advances in Information Systems (ADVIS'2000), to be published
8. A. Del Bimbo: Expressive Semantics for Automatic Annotation and Retrieval of Video Streams, In: Proceedings of the IEEE Conference on Multimedia & Expo, pp 671-674, IEEE Computer Society Press, 2000