

Chung-Sheng Li ; John R. Smith ; Vittorio Castelli; Lawrence D. Bergman Show Abstract In this paper, the performance of similarity retrieval from a database of earth core images by using different sets of spatial and transformed-based texture features is evaluated and compared. A benchmark consisting of 69 core images from rock samples is devised for the experiments. We show that the Gabor feature set is far superior to other feature sets in terms of precision-recall for the benchmark images. This is in contrast to an earlier report by the authors in which we have observed that the spatial-based feature set outperforms the other feature sets by a wide margin for a benchmark image set consisting of satellite images when the evaluation window has to be small 32 X 32 in order to extract homogenous regions. Consequently, we conclude that optimal texture feature set for texture feature-based similarity retrieval is highly application dependent, and has to be carefully evaluated for each individual application scenario. Rodney Long ; George R. We are now working 1 to determine utility of data directly derived from the images in our databases, and 2 to investigate the feasibility of computer-assisted or automated indexing of the images to support image retrieval of images of interest to biomedical researchers in the field of osteoarthritis. To build an initial database based on image data, we are manually segmenting a subset of the vertebrae, using techniques from vertebral morphometry. From this, we will derive and add to the database vertebral features. The customized-queries approach first classifies a query using the features that best differentiate the major classes and then customizes the query to that class by using the features that best distinguish the subclasses within the chosen major class. This research is motivated by the observation that the features which are most effective in discriminating among images from different classes may not be the most effective for retrieval of visually similar images within a class. This occurs for domains in which not all pairs of images within one class have equivalent visual similarity. We apply this approach to content-based retrieval of high-resolution tomographic images of patients with lung disease and show that this approach yields The traditional approach that performs retrieval using a single feature vector yields only Chi-Ren Shyu ; T. Tony Cai; Lynn S. Broderick Show Abstract In the picture archiving and communication systems PACS used in modern hospitals, the current practice is to retrieve images based on keyword search, which returns a complete set of images from the same scan. Both diagnostically useful and negligible images in the image databases are retrieved and browsed by the physicians. In addition to the text-based search query method, queries based on image contents and image examples have been developed and integrated into existing PACS systems. Most of the content-based image retrieval CBIR systems for medical image databases are designed to retrieve images individually. However, in a database of tomographic images, it is often diagnostically more useful to simultaneously retrieve multiple images that are closely related for various reasons, such as physiological contiguity, etc. For example, high resolution computed tomography HRCT images are taken in a series of cross-sectional slices of human body. Typically, several slices are relevant for making a diagnosis, requiring a PACS system that can retrieve a contiguous sequence of slices. In this paper, we present an extension to our physician-in-the-loop CBIR system, which allows our algorithms to automatically determine the number of adjoining images to retain after certain key images are identified by the physician. Only the key images, so identified by the physician, and the other adjoining images that cohere with the key images are kept on-line for fast retrieval; the rest of the images can be discarded if so desired. This results in large reduction in the amount of storage needed for fast retrieval. Roberts Show Abstract DARWIN is a computer vision system, which helps researchers identify individual bottlenose dolphins, *Tursiops truncatus*, by comparing digital images of the dorsal fins of newly photographed dolphins with a database of previously identified dolphin fins. In addition to dorsal fin images, textual information containing sighting data is stored for each of the previously identified dolphins. The software uses a semiautomated process to create an approximation of the fin outline. The outline is used to formulate a

sketch-based query of the dolphin database. The system utilizes a variety of image processing and computer vision algorithms to perform the matching process, which is necessary to identify those previously identified fins, which most closely resemble the unknown fin. The program presents the database fin images to the researcher in rank order for comparison with the new fin image. A hierarchy of retrievable units is automatically constructed by integrating information from different media. The hierarchy provides a compact, yet meaningful, abstraction of the broadcast news data, similar to a conventional table of content that can serve as an effective index table, facilitating the capability of browsing through large amounts of data in a nonlinear fashion. The recovery of the semantic structure of the data further enables the automated solutions in constructing visual representations that are relevant to the semantics as well as in establishing useful relationships among data units such as topic categorization and content based multimedia hyperlinking. Preliminary experiments of integrating different media for hierarchical segmentation of semantics have yielded encouraging results. Some of the results are presented and discussed in this paper.

Safranek Show Abstract Multimedia information systems are experiencing a tremendous growth as a direct consequence of the popularity and pervasive use of the world wide web. As a consequence, it is becoming increasingly important to provide efficient and flexible solutions for accessing and retrieving multimedia data. Images and video are emerging as significant data types in multimedia systems. And yet, most commercial systems are still text and keyword based and do not fully exploit the image content of these systems. We believe that there is an opportunity to build a novel interactive multimedia system for some specific applications in electronic commerce. In this paper, we present an overview of our approach, the rationale behind it and the problems that are inherent in building such a system. We address some of the technical issues in representing and analyzing image primitive features. These are the building blocks of any such systems. They can be generalized into a much broader range of applications as well.

Lewis Show Abstract We developed a content-based retrieval scheme for texture by using text-based description. The texture technique is based on our previous work, which uses very simple texture primitives, such as edges and plain regions to generate features. Other methods that apply complicated statistics can be difficult to transcribe into understandable forms for normal users. Unlike these other methods, with the simplicity of our features, we can express them in terms of simple language. Hence, we can bridge the gap between semantics and computed features. A number of benefits can be achieved, which open a new horizon for content-based retrieval with texture. For example, the user can request a texture image without necessarily knowing what types of textures are stored.

Alan Hanjalic ; Reginald L. Lagendijk ; Jan Biemond Show Abstract In this paper, we present the concept of an efficient semiautomatic system for analysis, classification and indexing of TV news program material, and show the feasibility of its practical realization. The only input into the system, other than the news program itself, are the spoken words, serving as keys for topic prespecification. After the basic analysis steps on a news program stream, including the processes of shot change detection and key frame extraction, the system automatically represents the news program as a series of longer higher-level segments. Each of them contains one or more video shots and belongs to one of the coarse categories, such as anchorperson news reader shots, news shot series, the starting and ending program sequence. The segmentation procedure is performed on the video component of the news program stream and the results are used to define the corresponding segments in the news audio stream. In the next step, the system uses the prespecified audio keys to index the segments and group them into reports, being the actual retrieval units. This step is performed on the segmented news audio stream by applying the wordspotting procedure to each segment. As a result, all the reports on prespecified topics are easily reachable for efficient retrieval. This is not longer practical, due to the growing size of these collections. That is why systems for image indexing based on their content are needed. Due to the use of large vector features, we adopted the pyramid trees are used for creating the index structure. The block diagram of the system is presented and the functionality of each block is explained. The features used are presented as well.

Hari Sundaram ; Shih-Fu Chang Show Abstract This paper presents algorithms to deal with problems associated with indexing high-dimensional feature vectors, which characterize video data. Indexing

high-dimensional vectors is well known to be computationally expensive. Our solution is to optimally split the high dimensional vector into a few low dimensional feature vectors and querying the system for each feature vector. This involves solving an important subproblem: Once we formulate the retrieval problem in terms of a retrieval model, we present an optimality criterion to maximize the number of results using this model. The criterion is based on a novel idea of using the underlying probability distribution of the feature vectors. A branch-and-prune strategy optimized per each query, is developed. This uses the set of features derived from the optimality criterion. Our results show that the algorithm performs well, giving a speedup of a factor of 25 with respect to a linear search, while retaining the same level of recall. Minoru Takahata; Hidetaka Kuwano; Shoji Kurakake; Chikashi Matsuda; Kazutoshi Nishimura Show Abstract We have developed a TV-on-demand system, which provides playback of a television program after a period ranging from a few seconds to one week after broadcast, and have conducted usage trials in cooperation with a television station in Nagano Prefecture of Japan. This system has been achieved through the development of various technologies, such as automatic updating of stored television programs and contents retrieval by telop characters. Users in the trials can begin playback of a television program immediately after its broadcast has begun. This paper presents applied technologies and some experimental results, and also addresses a new direction of information retrieval system based on the evaluation of the usage trials. Dan Schonfeld ; Dan Lelescu Show Abstract In this paper, we present topics related to tracking of video objects in compressed video databases, within the context of video retrieval applications. The structure of the video compression standards is exploited in order to avoid the costly decompression operation. This is achieved by utilizing motion compensation - a critical prediction filter embedded in video compression standards - to eliminate and interpolate the desired method for template matching. Occlusion analysis, filtering and motion analysis are used to implement fast tracking of objects of interest on the compressed video data. Being presented with a query in the form of template images of objects, the system operates on the compressed video in order to find the images or video sequences, where those objects are present and their positions are in the image. This enables the retrieval and display of the query-relevant sequences. Jau-Yuen Chen; Charles A. Bouman ; John C. Dalton Show Abstract In this paper, we describe a new approach to managing large image databases, which we call active browsing. Our method is based on a similarity pyramid data structure, which hierarchically organizes the database, so that it can be efficiently browsed. At coarse levels, the similarity pyramid allows users to view the database as large clusters of similar images. We discuss relevance feedback for the browsing process, and argue that it is fundamentally different from relevance feedback for more traditional search-by-query tasks. We propose two fundamental operations for active browsing: Both of these operations depend on a user-defined relevance set, which represents the image or set of images desired by the user. Yeung Show Abstract Image search has been actively studied in recent years. On the other hand, image browsing has received limited attention. Image browsing refers to the process of presenting forms of overview or summary of the image relationships, thus facilitating a user to navigate across the data set and find images of interest.

DOWNLOAD PDF STORAGE AND RETRIEVAL FOR IMAGE AND VIDEO DATABASES VII

Chapter 2 : Tables of Contents for Storage and Retrieval for Image and Video Databases VII

Note: Citations are based on reference standards. However, formatting rules can vary widely between applications and fields of interest or study. The specific requirements or preferences of your reviewing publisher, classroom teacher, institution or organization should be applied.

In this paper the incoming video is systematically and efficiently reduced via a frame selection procedure which takes advantage of the fact that the incoming video is encoded using one of several existing DCT- based standards. The procedure is performed in the frequency domain prior to video decoding. Further refinement in the frame selection step is achieved using a robust metric based upon the color histogram of the selected subset of decoded frames. The procedure is presented in detail and several examples are exhibited. Jain Show Abstract Visual information systems require a new insertion process. Prior to storage within the database, the system must first identify the desired objects shots and episodes , and then calculate a descriptive representation of these objects. This paper discusses the steps in the insertion process, and some of the tools we have developed to semi-automatically segment the data into domain objects which are meaningful to the user. Image processing routines are necessary to derive features of the video frames. Models are required to represent the desired domain, and similarity measures must compare the models to the derived features. Suh-Yin Lee; Huan-Ming Kao Show Abstract With the progression of multimedia technology and the trend of visualization of man-machine interface, video data will become a kind of fundamental resource of data as general as text and graphics. When the data size is large, retrieving data will become a time-consuming process. Much research has been done to facilitate the retrieval of data, e. Similarly, the video data also encounter the same problem. It is very difficult to generate video indexes automatically, because the index is application-dependent and only the user knows what indexes he actually needs. In this paper a mechanism for video data indexing is proposed which is based on the concept of objects and object motion with interactive annotation. This mechanisms provides an efficient method to access video substance without browsing the whole video data. A motion representation for the track of a moving object is presented. The access methods of video queries are introduced. A prototype of the video indexing system is implemented. Autosophy databases are educated very much like a human child to organize their own internal data storage. Input patterns, such as written questions or images, are converted to points in a mathematical omni dimensional hyperspace. The input patterns are then associated with output patterns, such as written answers or images. Omni dimensional information storage will result in enormous data compression because each pattern fragment is only stored once. Pattern recognition in the text or image files is greatly simplified by the peculiar omni dimensional storage method. Video databases will absorb input images from a TV camera and associate them with textual information. Self-aligning autosophy databases may lead to a new generation of brain-like devices. Ramin Samadani ; Cecilia S. This data set provides a realistic testbed for developing algorithms for scientific image databases. Scientists studying the aurora currently need to browse through large numbers of images to find events suitable for further scientific studies. They select or reject an image based on a variety of visual queues, including shape, size, and intensity. This paper describes a system currently under development for selecting interesting events based on image content. We use boundaries from the images to outline the aurora, and then to extract features that relate to shape, size, and intensity. These features are then input into a supervised decision tree classifier. The system retrieves images of potential interest to the user. The user makes the final decision regarding the use of the images retrieved. The node and the branch of an ARG denotes a classified region and a spatial relationship between adjacent regions, respectively. Similarity retrieval thereby turns out to be equivalent to graph matching. The similarity retrieval process of the system is as follows: The method employs a new similarity measure between graph representations of images. The measure is effective for drawing images that describe logical meaning by their structure. Most of the currently available image database systems offer retrieval functions called key word retrieval, where users specify key

words such as titles, attributes, and categories of themes. But it is not easy for the users to select suitable key words according to the purpose of retrieval. So recently some retrieval functions called similarity retrieval have been proposed, where users specify key images by means of examples, sketches, and icons. We are developing a drawing image database system that stores plant diagrams. The system scans, recognizes, and stores plant diagrams. Then users can refer to any parts of the diagrams according to their needs. The system is used as a help to plant observation and control. To realize similarity retrieval for logically structured drawings like plant diagrams, we introduced a graph representation of drawings, which is suitable to deal with their logical structure. Then we defined a similarity measure between them. In this paper, effectiveness of the similarity measure and applicability to plant diagrams are discussed and some experimental results are shown.

John Stonham Show Abstract We address some of the problems of accessing database images which do not contain any indexing information and investigate methods of automating search strategies which currently rely on human operators to match the target against a number of images in the database. Such problems might include the extraction of facial photographs from a library given a suspect or the registration of new trademarks whose uniqueness must be assured. The object of the retrieval mechanism is to narrow down the search space for final perusal by the human operator. We present a neural network based coding scheme to retrieve images from a database according to the degree of similarity with a target image. The code represents each image with respect to a set of feature archetypes learned by the neural network during a training phase. We introduce a novel neural network learning law which performs an extremely efficient implementation of principal component analysis and maximizes the amount of information conveyed by the code. We present results using a database of machine printed fonts and discuss how the image size, the database diversity, and code length affect the efficacy of the retrieval mechanism.

Swain Show Abstract The general problem of object recognition is difficult and often requires a large amount of computing resources, even for locating an object within a single image. One important optimization is to take advantage of interaction with the user to find out what types of variation are expected in the database, and to rely on the user to discriminate between similar-looking objects. Another is to create appropriate data structures off-line to speed on-line searches. We are building a tool, called FINDIT, for locating the image of an object from within a large number of images of scenes which may contain the object. The user outlines an object in an image that he wants to find in the database, and specifies the constraints on the transformations of the object that are expected to occur. The program acts as a filter to quickly reduce the possible number of candidates to a number small enough to be perused by the user. FINDIT chooses an appropriate search algorithm depending on the selection of constraints by the user.

Shann Show Abstract We report a database language for visual retrieval which allows queries on image feature information which has been computed and stored along with images. The language is novel in that it provides facilities for dealing with feature data which has actually been obtained from image analysis. The MVQL constructs are mainly based on proven operators from the field of digital image analysis. An example is the Hough-group operator which takes as input a specification for the objects to be grouped, a specification for the relevant Hough space, and a definition of the voting rule. The output is a ranked list of high scoring bins. The query could be directed towards one particular image or an entire image database, in the latter case the bins in the output list would in general be associated with different images. We have implemented MVQL in two layers. The command interpreter is a Lisp program which maps each MVQL line to a sequence of commands which are used to control a specialized database engine. In the paper we outline the language and provide examples of useful queries. We also describe our solution to the engineering problems associated with the implementation of MVQL.

Roger David Hersch Show Abstract Professionals in various fields such as medical imaging, biology, and civil engineering require rapid access to huge amounts of uncompressed pixmap image data. In order to fulfill these requirements, a parallel image server architecture is proposed, based on arrays of intelligent disk nodes, each disk node being composed of one processor and one disk. Pixmap image data is partitioned into rectangular extents, whose size and distribution among disk nodes minimize overall image access times. Disk node processors are responsible for maintaining both the data

structure associated with their image file extents and an extent cache offering fast access to recently used data. Disk node processors may also be used for applying image processing operations to locally retrieved image parts. This contribution introduces the concept of an image oriented file system, where the file system is aware of image size, extent size, and extent distribution. Such an image oriented file system provides a natural way of combining parallel disk accesses and processing operations. The performance of the proposed multiprocessor-multidisk architecture is bounded either by communication throughput or by disk access speed. However, when disk accesses are combined with low-level local processing operations such as image size reduction zooming , close to linear speedup factors can be obtained by increasing the number of intelligent disk nodes. Hiroaki Sakamoto; Kazutoshi Nishimura; Yutaka Ishibashi ; Hirotaka Nakano Show Abstract Advanced visual information retrieval systems supporting both video and images need to have flexible system design so that their system configurations can easily be enhanced. It is therefore desirable to separate the features of a central system into three parts: In this architecture, unscheduled arrivals of data blocks at a back-end network cause two problems: To solve these problems, we have designed a new multimedia integrated switching system MISS that uses a fully connected crossbar switch to combine servers. According to simulation results and estimates based on queuing theory, MISS greatly reduces video frame fluctuation and halves the average image transfer delay. These effects have been confirmed in an experimental visual communication system built around MISS. Yie-Tarng Chen; Rangasami L. Kashyap; Arif Ghafoor Show Abstract In this paper, a top-down data placement methodology for a large interactive multimedia information system MMIS on a single spindle multi-disk environment such as a Jukebox is presented. The objective of this work is to minimize average disk seek time as well as the number of platter switches for Jukebox. A large data placement problem can be divided into a number of small data placement problems by weighted graph decomposition. The Kernighan-Lin partitioning algorithm is recursively applied for this purpose. Once the graph is fully partitioned, the objects in the same subgraph are assigned to the same disk. The data placement within a disk is divided into two stages, global data placement and detailed data placement. The expected access patterns of global data placement are modeled as a time-homogeneous ergodic Markov Chain, from which the stationary probability for each node of the browsing graph can be found. Based on these probabilities, we define an expected access cost. Then, the problem of global data placement is posed as an optimization problem, and various clustering and storage layout algorithms are proposed. Xiaobo Li; Zhiyong Liu Show Abstract In past decades, many storage schemes for large images on parallel computers have been proposed to provide simultaneous access to various subsets of the pixels.

Chapter 3 : dblp: Storage and Retrieval Methods and Applications for Multimedia

Tables of Contents for Storage and Retrieval for Image and Video Databases VII. and retrieval in large image databases [] video retrieval and tracking.

Their performances are continuously improving and their base principles span a wide range of diversity. However, there are a number of factors that are ignored when dealing with images which should be dealt with when using videos. These factors are primarily related to the temporal information available from a video document. While these factors may complicate the querying system, they also may help in characterising useful information for the querying. The temporal information firstly induces the concept of motion for the objects present in the document. When in CBIRS, it is the list or organisation of such object which is search for, video retrieval may imply the retrieval of a behaviour of an object throughout the document. Two video documents may therefore contain the same objects but little relevance may be found between the two in this search context. It is therefore essential to encode within the indexing of a video document the behaviour of all objects throughout the document. A video document can typically be split into a hierarchical structure see section 4. Another issue in video retrieval is the complexity of the querying systems. A very elaborated retrieval systems would allow flexibility for the user to specify its query parameters. Query-by-example systems typically used in CBIRS require the user to show the system one or more document similar to what he is looking for. Although this seems natural for images the whole content of an image may be scanned in one glance , this querying process is far more complicated to adapt into the context of video documents. This technology is fairly recent and it is currently necessary to examine where it would just replace existing systems, where it can really bring some improvement and where it will open new possibilities. By definition, a CBVRS aims at assisting a human operator user to retrieve a video sequence target within a potentially large database. Three major cases may be distinguished. The user has a specific sequence in mind and knows it is included within the database. In this case, the target is unique and corresponds to specific criteria. The user will be able to describe fairly precisely the targeted sequence and will be able to see at the first glance whether a suggested sequence corresponds to his wish. Since he knows that the document is in the database, the user will keep querying until he finds the document. In this case, an indexing using text such as keywords or title should suffice. The user has a specific video in mind and does not know if the document he is looking for exists in the database. The problem here is to provide a precise search tool so that the user can quickly make a decision as to whether the target is in the database or otherwise. The user simply searches for a document by referring to its topic or some event occurring within it. In this case, the search should be hierarchical, in order to guide the user through the search. Relevance feedback is important in this case since it would be the means for user interaction by allowing the user to filter the response of the system. An important issue in this case is the representation under which the response is given. Since the user is not supposed to know the documents, he should be able to scroll all responses quickly for selecting the relevant or irrelevant ones. It is therefore essential the each document is represented by a short but complete and comprehensive description. Potential users may arise from different horizons. The need for keeping complete archives induces a large volume of typically short documents. The specificity of news report video documents makes an automated CBVR very attractive to retrieve documents with respect to a specific topic, place or appearance of a given character. One should also note that the audio stream attached to all such documents as well as captions generally present in news reports form important cues for an automated retrieval process. Here again, documents are typically short. An example of retrieval may concern the fact of retrieving all document w. Such a characteristic is generally difficult to express using text so that a visual-based approach is highly desirable. Here again, the characteristics of such documents may be difficult to express based on given textual annotations stored on a database. For this type of specific applications, one may think of retrieving documents on the basis of some dance step described by e. Two types of educational documents can be distinguished. Firstly, lectures where

the main content is given by the audio stream and practical courses when both the visual contents and the audio stream are of importance. Video archiving Medical applications. The major issues for ensuring the usability of such systems emerge as follows. Obtaining a compact and complete video sequence representation Providing the user different search strategies adapted to the type of search he is doing. The next section reviews the internal representation for automatically deriving a representation of a multimedia document. Spatial scene analysis This section first reviews some visual document processing operations that will be essential for automatically extracting an extensive description of a document. Feature extraction aims at characterising a list of properties called feature vector or document signature for each component pixel, frame region, frame, sequence of a video document. Feature extraction operations rely on the analysis of the human visual system HVS and range from simple statistics to elaborated model-based filtering techniques. Another distinction separate unsupervised i. The analysis of elements such as colour and texture aim at characterising features in the spatial space as opposed to the temporal domain. Experience acquired from CBIR studies may be fully transferred to video in this case. More specific to the video domain is the use of temporal and possibly audio information to characterise further a document. Some key-advances made in these directions are reviewed below. Colour feature space Colour is an important cue for measuring the similarity between visual documents. Historically, colour features have been the first features used in the context of CBIR. Colour statistics are used for measuring global or local dissimilarities. Global colour features are analysed through histograms. These histograms offer the advantage of being invariant under rotation, translation and many other geometric operations. However, such a global analysis does not allow for characterising the spatial organisation of the colour within the 2D spatial domain. There is therefore a need for refining this technique with using colour layout features. Features encoding colour organisation within the document are often based on blocks and are either blind i. A feature vector is attached to each unit part of the spatial domain, and it is the relationship between neighbouring image parts which is encoded as feature. Whatever their exact definition is, the comparison of all these features calls for the definition of distance measures between colours and colour histograms. Colour distances are defined implicitly through the use of a given colour space. Colour spaces such as YIQ and were designed based on empirical measurements of the HVS so that the Euclidean distance defined in the respective colour space fairly represents a perceptual distance see for example [16] for a thorough study. Colour statistics based on training or more generally supervised learning may allow to distinguish between pixel colours having a specific semantics e. Once histograms have been obtained, they may be compared using histogram distances see [33] for examples. Texture feature space The analysis of textures requires the definition for a local neighbourhood corresponding to the basic texture pattern. It makes no sense to study the texture of an isolated pixel. Typically, the analysis is done via the mapping of the texture onto the response of one or a bank of pre-defined filters against the image wavelets, Gabor filters. Once again, this response space aims at allowing the use of a similarity measures between feature vectors. The response image may therefore be analysed using a similar set of tools than for the analysis of colour. Another approach defines textons [12 , 23] as the basic builders for any texture. Each texture is decomposed using these building blocks and the parameters of the local texture are obtained. Typical texture features include orientation and coarseness [12]. In [31], texture models are learned so that geodesic active contours are able to segment texture regions then considered as uniform or consistent patches , thus extending the classic snake-based segmentation algorithms. Supervised learning of textures is done via texture samples. One classic texture database is the Brodatz texture album. An analysis of colour and texture dissimilarity measures for CBIR is proposed in [33]. Supervised feature spaces More complex features may be defined for parsing the contents of a video document. One example of such feature is the development of face detection algorithms. Human faces are widely recognised as a useful cue for video indexing. Here again, the same applies to image indexing but more information can be obtained in the context of video indexing. The recognition of persons present in an image allow for further classification. The same classification readily applies to video. In a video document, a further step can be made by exploiting the frequency of occurrence of one person for the

classification of the document with respect to the ID of that person. Furthermore, face localisation may constitute the starting point of the segmentation of the person within the video document. Motion information obtained through that segmentation will allow further classification. Face localisation algorithms are mostly based on supervised techniques such as Neural Networks [9] or HMM [27 , 26]. The retrieval of text in a video document may also improve its understanding. It is often the case that textual annotations are readily available within the document itself. Using text specificity such as geometrical shape and contrast, captions or credits may be extracted and processed through OCR for completing the document indexing see e. Finally, combining models for understanding may permit high-level interpretation [32].

Temporal analysis The temporal dimension of a video document contains an information that is specific to this type of document. The temporal analysis of that document typically requires its partitioning into basic elements. It is now recognised that this partitioning can operate at four different levels of granularity. Each frame is treated separately. There is no or little temporal analysis at this level. A shot is a set of contiguous frames all acquired through a continuous camera recording. The partitioning of the video into shots generally does not refer to any semantic analysis.

DOWNLOAD PDF STORAGE AND RETRIEVAL FOR IMAGE AND VIDEO DATABASES VII

Chapter 4 : CiteSeerX " Adaptive Storage and Retrieval of Large Compressed Images

Storage and Retrieval for Image and Video Databases VII, San Jose, CA, USA, January , SPIE Proceedings , SPIE , ISBN Image Retrieval Applications.

The Vibe video database system: An update and further studies by Cuneyt M. In this paper we extend the shot transition detection component of the ViBE video database system to include gradual scene changes. We also present results on the performance of our cut detection algorithm using a large test set. The performance of two other techniques are compared against our method. Show Context Citation Context We can define these transitions using the following video edit model: In this paper we describe a framework of analyzing programs belonging to different TV program genres using Hidden Markov Models and pseudo-semantic features derived from video shots. Clustering using Gaussian mixture models is used to determine the order of the models. Results for initial genre class Results for initial genre classification experiments using two simple features derived from video shots are given. Consequently, there has been a great interest in designing and building systems that organize and search video data based on its content. Since the smallest meaningful unit of video is the shot, this may be achieved. Furthermore, assume that the random variables X_1, X_2 , Farag, Hussein Abdel-Wahab , " The increasing use of multimedia streams nowadays necessitates the development of efficient and effective methodologies and systems for manipulating databases storing these streams. These systems have various areas of application such as video-on-demand and digital libraries. The importance of video The importance of video content-based retrieval CBR systems motivates us to explain their basic components in this chapter and shed light on their underlying working principles. In general, a content-based retrieval system of video data consists of the following four stages: Each one of the above stages will be reviewed and expounded based on our experience in building a Video Content-based Retrieval VCR system that has been fully implemented from scratch in JAVA Language Moreover, current research directions and outstanding problems will be discussed for each stage in the context of our VCR system. This paper describes the ongoing development of tools for assisting the management of large volumes of visual data in film and television postproduction. By integrating recently developed content-based image retrieval techniques into an existing asset tracking framework, we can facilitate reuse of archived material by making the contents of the archive searchable on the basis of visual content.

Chapter 5 : Content-based Video Retrieval: An overview

Storage and Retrieval for Image and Video Databases VII, San Jose, CA, USA, January , SPIE Proceedings , SPIE , ISBN [contents] 6.

Chapter 6 : dblp: Storage and Retrieval for Image and Video Databases

Venue: in Storage & Retrieval for Image and Video Databases, VII, M.M Yeung, B.L. Yeo and C. A. Bouman Eds. Proc. SPIE Enabling the efficient storage, access and retrieval of large volumes of multi-dimensional data is one of the important emerging problems in databases. We present a framework for.

Chapter 7 : CiteSeerX " Citation Query Video and image databases: Who cares

Currently, much work has been done in the area of image storage and retrieval. However, the overall performance has been far from practical. A highly integrated wavelet-based image management system is proposed in this paper.

DOWNLOAD PDF STORAGE AND RETRIEVAL FOR IMAGE AND VIDEO DATABASES VII

Chapter 8 : Storage and Retrieval for Image and Video Databases | () | Publications | Spie

Advanced Search >. Home > BrowseVolume Browse Proceedings.

Chapter 9 : Vittorio Castelli - Publications

You, Color-WISE: a system for image similarity retrieval using color, in Storage and Retrieval for Image and Video Databases VI, San Jose, , pp. 6 X. Wan and C.-C. Jay, Kuo, Color distribution analysis and quantization for image retrieval, in Storage and Retrieval for Image and Video Databases IV, SPIE, , pp.