# Colorado Technical University

**Technical Report**
**Computer Science**

## Improving the Performance of Content Based Image Retrieval Systems

Mary Jane Willshire
University of Portland
mailto:willshir@up.edu

Thomas Allen
Air Force Research Labs, Kirtland AFB
mailto:Thomas.Allen@kirtland.af.mil

**Technical Report Number**
**CTU-CS-2003-002**

# Improving the Performance of Content Based Image Retrieval Systems

Mary Jane Willshire, University of Portland*
Thomas Allen, Air Force Research Labs, Kirtland AFB†

**Abstract**

Our world is dominated by visual information and a tremendous amount of such information is being added day-by-day. It would be impossible to cope with this explosion of visual data, unless the images are organized such that we can retrieve them efficiently and effectively. At the core of content-based image retrieval (CBIR) is the requirement that database elements must be indexed to facilitate retrieval in an efficient manner. Most existing image retrieval systems are text-based, but images frequently have little or no accompanying textual information. Problems with text-based access to images have prompted increasing interest in the development of image-based solutions. On the other hand, CBIR relies on the characterization of primitive features such as color, shape, and texture that can be automatically extracted from images themselves. Hence, the field of CBIR focuses on intuitive and efficient methods for retrieving images from a database based solely on the content contained in the images. This study introduces a novel clustering methodology based on the gradient of images using the concept of "spin-up" and "spin-down" states derived from statistical mechanics to improve the speed of retrieval and improve the accuracy of retrieval in comparison to the traditional color histogram $L_1$-norm retrieval methodology. By expanding the interpretation of color in images to include a gradient-based description, a new indexing method for content-based retrieval of images from an image database is developed for the reduction of false positives in the retrieval process.

**Keywords**: CBIR, statistical mechanics, spin-up, spin-down, $L_1$-norm, Minkowski Metrics, gradient, ANOVA, information theory.

## 1. Introduction

At the core of content-based image retrieval (CBIR) systems is the requirement that database elements must be indexed to facilitate retrieval in an efficient manner. Problems with text-based access to images have prompted increasing interest in the development of image-based solutions, primarily in the area of primitive features, such as color, shape, and texture. An area in CBIR research lacking in research is the development of a gradient-based methodology for categorizing images for efficient retrieval of relevant images. This paper presents a gradient-based scheme for clustering images according to a complexity rating and assigning an index number to each cluster in order to improve the accuracy of retrieval in comparison to the traditional way of accomplishing retrieval, namely the color histogram retrieval methodology (commonly referred to as the $L_1$-norm in this paper). In order to show improvement over the $L_1$-norm, this paper introduces a methodology that places images into relevant groups or relevant clusters based on an image complexity index value determined by the gradient. The placement into clusters allows a given query image to choose a relevant cluster as opposed to searching every single image in the database in order to pull out all relevant images. This type of schema in turn reduces the search space which in turn allows the facilitation of fast and efficient retrieval. This CBIR methodology lastly proposes a final pass on the reduced image database using the $L_1$-norm to extract all appropriate images and place them in order of relevance.

## 2. Statement of Problem

In general, "information retrieval" is the process of converting a request for information into a meaningful set of references. With information retrieval, a list of image features and attributes is presented in increasing levels of subjective form and abstraction. A classification of query types based on a similar analysis of image features and attributes was developed by Eatkins and Graham [1]. According to Eatkins and Graham, queries from an image can be aggregated into three levels of increasing complexity:

1) **Level 1** comprises retrieval by syntactic features such as color, texture, shape or the spatial location of image elements. Examples of such queries might include "retrieve all images with red blobs in the middle of the image", "retrieve images that contain blue squares, rectangles, and diamonds", - or most commonly "find more pictures that look like this".

* willshir@up.edu
† Thomas.Allen@kirtland.af.mil

2) **Level 2** comprises retrieval by derived (sometimes known as logical) features, involving some degree of logical inference about the identity of the objects depicted in the image. It can usefully be divided further into retrieval of objects of a given type (e.g. "find pictures of a double-decker bus") and retrieval of individual objects or persons (e.g. "find a picture of the Eiffel tower" or "find a picture of a woman").
3) **Level 3** comprises retrieval by abstract attributes, involving a significant amount of high-level reasoning about the meaning and purpose of the objects or scenes depicted. Again, this level of retrieval can usefully be subdivided into: 1) retrieval of named events or types of activity (e.g. "find pictures of Scottish folk dancing" or "find a picture of a posing fashion model") and 2) retrieval of pictures with emotional or religious significance ("find a picture depicting suffering").

The clustering technique introduced in this paper is tested using Level 1 retrieval which is generally the focus of CBIR research and systems development [2]. At the present time, Level 2 and Level 3 queries are considered harder to implement, as exemplified by several decades of computer vision techniques.

## 3. $L_1$-norm Metric

The $L_1$-norm for the color histogram captures the distribution of colors in an image or region of an image. A detailed account and analysis of the metrical properties of the color histogram space is given in [3]. The $L_1$-norm is defined as

$$d_{l1}(I,H) = \|I - H\|_{L1} = \sum_{l=1}^{n} (\|i_l - h_l\|) \quad (1)$$

where $I$ and $H$ are two color histogram vectors. The $L_1$-norm for the color histogram in general is a subset of an M-dimensional vector space and forms the face of an M-dimensional simplex. Therefore, in order for two distinct histograms $I$ and $H$ to be distinguishable from one another, they must be separated by a non-zero distance $t$. This property is called *t-difference* and describes the $L_1$-norm for the color histogram space as a Hausdorf space. As a result, the value of $t$ depends on the composition of the image data set.

### 3.1. Advantages and Disadvantages of the *$L_1$-norm* for the Color Histogram

Swain and Ballard's seminal paper [4] matched images based solely on their color. The distribution of color was represented by color histograms, and formed the images' feature vectors. The similarity between a pair of images was then calculated using the $L_1$-norm similarity measure between their histograms called the "normalized histogram intersection". This approach became very popular due the following advantages [5]:

1) Robustness: The color histogram is invariant to rotation of the image on the view axis and changes in small steps when rotated or scaled. It is also insensitive to changes in image and histogram resolution and occlusion.
2) Effectiveness: There is a high percentage of relevance between the query image and the extracted matching of images,
3) Implementation simplicity: The construction of the color histogram is a simple scanning of the image, to get the color values, discretization of the color values to the resolution of the histogram, and building the histogram using color components as indices.
4) Computational Simplicity: The histogram computation (not the retrieval computation) has $O(M^2)$ complexity for images of size MxM.
5) Low storage requirements: The color histogram size is significantly smaller than the image itself, assuming color quantization.

There are some problems with the color histogram though, namely:

1) Different images may have similar or identical color histograms. For instance, as an extreme case, randomly scrambling the positions of pixels in an image leaves its histogram unaffected despite massive changes in the image content.
2) Images taken under different ambient lighting may produce different histograms. This has been partly addressed by applying color constancy normalization [5].
3) Although the complexity for a single image match is linear O(n) without sorting the results (where n represents the number of different colors or resolution of the histogram), a sort from least to greatest match performed on the results worsens the complexity of the retrieval to O[nlog(n)]. This, coupled with the color histogram computation of $O(M^2)$ can make the retrieval computational time prohibitive for large databases.

### 3.2. Distance Metrics and Similarity

The mathematical concept of similarity is the mapping of items to positive real numbers, typically in the normalized range of [0,1]. Smith [6, 7] empirically tested the $L_1$-norm against the Minkowski Metrics, such as the $L_2$-norm (Euclidean distance), Hamming distance, Chebyshev distance, the quadratic form distances, and others to determine retrieval performance. His conclusion was that there was little difference in retrieval accuracy and performance among the Minkowski metrics. Additionally, the $L_1$-norm was found to be statistically more robust to measurement outliers than the $L_2$-norm. Therefore, this research uses the $L_1$-norm (on a second pass) to compute the similarity between two points in a gradient feature space and a color histogram space. This decision is based primarily on the fact that the $L_1$-norm is a smaller computational effort compared to other distance functions (as implied by Smith) [6, 7].

## 4. Shape Based Research

One of the primitives comprising image analysis is “shape”. Queries for shapes are generally achieved by selecting an example image provided by the system or by having the user sketch a shape. The primary mechanism used for shape retrieval include identification of features such as lines, boundaries, aspect ratio, and circularity, and by identifying areas of change or stability via region growing and edge detection. In current research, two major steps are involved in shape feature extraction. They are: object segmentation and shape representation [8, 9].

Image retrieval based on object shape is considered to be one of the most difficult aspects of content-based image retrieval because of difficulties in low-level image segmentation and the variety of ways a given 3D object can be projected into 2D shapes. Several segmentation techniques have been proposed so far, which include the global threshold-based technique [10], the region growing technique [11], the split and merge technique [12], the texture-based technique [13], the color-based technique [14], and the model-based technique [15]. Generally speaking, it is difficult to do a precise segmentation due to the complexity of the individual object shape, the existence of shadows, noise, etc. Gradient-based research in CBIR would make these issues a little easier to handle since indexing would only depend on performing a convolution filter on the image in order to produce an image that depends only on edge-information (or gradient-information) only. No satisfactory gradient-based CBIR technique has been devised to date.

Once objects are segmented, their shape features can be represented and indexed. In general, shape representations can be classified into three categories [16]:

1) Boundary based: a representations based on the outer boundary of the shape
2) Region-based: based on the entire region
3) Combined representations: the combined effort of the above two techniques.

## 5. Gradient Based Research

The present shape-based research makes use of extremely useful and powerful features, but the above methodologies discussed in Section 4.0 generally run into two problems according to Gagaudakis [5]:

1) They mostly require the image to be partitioned into regions from which shape descriptors can then be extracted. Unless the segmentation is directed by the user (e.g., [17]) segmentation algorithms are prone to fail, especially when new situations due to different imaging modality or object type are present [18].
2) Determining effective shape descriptors for complex natural objects is a complex process and remains an active area of research [19, 20].
3) None of the shape-based methods exploit gradient-based feature extraction which has the potential for reducing the number of false hits as compared to the color histogram as implied by Tao and Grosky: "As color plays an important role in image composition, many color indexing techniques have been studied. Although global color histograms and moments have been proven to be very useful for image indexing, they do not take *color based spatial information* into account. Thus, when the image collection becomes very large, many false hits frequently occur. " [21].

The gradient operator has been the basis for various approaches of image differentiation [22] but has not been used as an indexing scheme for CBIR or used as a mechanism for clustering. Gradient-based features that are extracted from an image have the potential for being extremely useful and powerful for improving retrieval above and beyond the color histogram for the following reasons:

1) Humans are sensitive to the shape trend of an entire image. Given two images, one that describes the landscape of a mountain and the other half the body of a man, the human would consider these two images as not similar even though they have, for the sake of argument, color distributions that are similar. However, if there are two images depicting half the body of a man, but have different color distributions the human recognizes that the two images are similar in view of the shape trend of their color distributions. The implication here is that discontinuities in edges of an image can be used as a means for distinguishing between different images with similar histograms.
2) If two objects are taken under different ambient lighting, the same image may produce different color histograms. This has been partly addressed by applying color constancy normalization but for the most part is still a problem in image retrieval. Using the gradient-based features of an image may be an invariant to ambient lighting since the edges (peaks and valleys) of an image is under consideration in gradient-based feature extraction.

## 6. Enumeration of Gradient States Methodology

This section discusses in detail a clustering methodology derived from the physics of spin as discussed by Kittel and the concept of statistical mechanics of learning as discussed in [23], both concepts synthesized and applied to CBIR process.

### 6.1. Sub-blocking

Many researchers suggest that using sub-blocking (sometimes called sub-imaging, both color feature and spatial relations) is a better solution to the image retrieval problem if applied appropriately to the spatial content of images [24-29]. Others contend that in order to extend the $L_1$-norm feature or apply it appropriately, a natural approach is to divide the whole image into sub-blocks and extract features from each of the sub-blocks [24, 26]. For the most part, the proponents of sub-blocking agree that sub-blocking is important for two primary reasons:

1) Images have the potential to be invariant to translation, rotation, and scaling based on extracting features from a block;
2) Sub-blocking has the potential for decreasing the storage and computational complexity of an image database.

As examples, Gorkani and Picard discriminate between photos of city scenes and photos of landscape scenes by sub-blocking images in a database of about 1000 images [25]. Yiu [27] used sub-blocking features to classify indoor and outdoor scenes using a database of 500 images. Thor [28] made use of sub-blocking techniques for developing an image color vector technique for image retrieval.

In the experiment conducted for this study, the images from the University of Washington database are preprocessed by dividing the images into 9 (3x3) sub-blocks based on similar research conducted by the following investigators in CBIR research: Ko, Lee, and Byun [29]. Ko, Lee, and Byun's research appeared to fit more closely to the research presented in this study in terms of their extensive analysis of sub-blocking techniques using geometric information (Ko, Lee, and Byun, tagged each sub-block with a single numerical moment value). As it turned out, Ko, Lee, and Byun's research was based on some earlier research conducted by Gong [30] who presented an alternative approach to reducing the computational complexity in color indexing by proposing the use of local histograms. In this paper, Gong partitions an image into 9 (3x3) sub-images. Next, for each sub-image, a color histogram is generated. The content of the image is represented by the histogram of the entire image and the histograms of the sub-images. However, this technique only considered color information. Geometric information is required to retrieve more precise images. Ko, Lee, and Byun [29] extended Gong's research by proposing a way for extracting geometrical information from an image by calculating central moments for each of the 9 (3x3) non-overlapping block-based segmentation of the image in order to represent a more meaningful spatial color information and reduce the index size. The images used for similarity retrieval in the sub-blocking technique of Ko, Lee, and Byun appeared to be invariant to translation and rotation. Using the principles derived from Ko, Lee, and Byun, the 9 (3x3) sub-blocks methodology is utilized in this research to obtain an efficient clustering and indexing methodology for an image database.

The goal of the above sub-blocking research is to provide a CBIR system that is capable of automatically generating associations between the low-level (sub-blocks) and semantic-level feature representations of an image database. With non-overlapping sub-blocks, the existing research implies that a certain amount of "fuzzy-ness" (or smearing of the image) is allowed to be incorporated in the spatial distribution of color information [31], allowing for variations in similar images to still remain relevant (within some acceptable limit), and therefore allowing them to remain within the same class or cluster of similarity. This concept is extended to spatial information in the form of the gradient for an image database.

### 6.2. Gradient Spin Excess as a Measure of Image Complexity

As stated earlier, the $L_1$-norm can provide reasonable discriminating power in image retrieval process, but it tends to give too many false positives when the image collection is too large. The *enumeration of gradient states (EGS) methodology* makes use of the gradient between each successive sub-block for determining a clustering and indexing scheme for efficient retrieval, to include the reduction of false positives.

When one calculates the gradient between each sub-block (see [22]), the *enumeration of gradient states methodology* tags a positive (increasing) gradient with a +1 and a negative (decreasing) gradient with a -1 (very similar to placing the image into a "active" and "passive" state as with the case of the "active" and "passive" state of the binary state variable *S* as discussed in the statistical mechanics of learning discussed in [32]). Symbolically, we represent the +1 as an up arrow (the gradient rising for values equal to or greater than 0) and the –1 as a down arrow (the gradient falling for negative values). Also, a rise is defined or referred to in this discussion as a spin up and a fall is referred to a spin down, terms that are adapted from statistical mechanics [33]. A single image can have N possible pixel-moments of orientation of the gradient (where N is the product of a PxQ pixel image, where P is the pixel-width and *Q* is the pixel-length). One orientation for a 4x6 image is depicted in Figure 1.

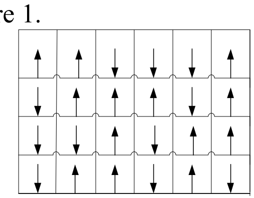


**Figure 1** Pixel Moments in a typical image

The "moment" term is a general term referring to the orientation of the spin (either a spin up or down); hence, each pixel is referred to as a pixel-moment in the *enumeration of gradient states methodology* based on the direction of the gradient. As an example, each cell for the 4 by 6 image in Figure 1 represents a pixel tagged with either a +1 (spin up) or a –1 (spin down). We can also *enumerate* such an arrangement of pixel-moments in the following manner:

$\uparrow\uparrow\downarrow\downarrow\downarrow\uparrow\downarrow\uparrow\uparrow\uparrow$ ...

(hence the term *enumeration of gradient states*). The total possible arrangement of N pixel-moments is $2^N$. The gradient orientations are assumed to be arranged in a definite order, and we might number them in sequence from left to right on the image, for example. On this convention, we can write:

$\uparrow_1\uparrow_2\downarrow_3\downarrow_4\downarrow_5\uparrow_6\downarrow_7\uparrow_8\uparrow_9\uparrow_{10} \cdots$

One observes from Figure 1 that the image appears to represent a statistical mixture of pixel-moments of gradient spin-up and gradient spin-down. How does one place a metric to an image based on this statistical mixture of pixel-moments? This is accomplished by calculating the difference between the spin-up states and the spin-down states [33] of the pixel-moments:

$$\text{(number-up)} - \text{(number-down)} = S \qquad (2)$$

The research claims that the value of *S* is sufficient for measuring the complexity of an image and thereby representing the amount of *edge/complexity information(or gradient information)* contained in an image. Since *S* is a single real value and not a vector value, the research has reduced the feature gradient space from an *n*-dimensional vector space to a one dimensional real value, namely *S*.

From Equation 2, the gradient spin excess value *S* is a measure of apparent random changes or random variations in the gradient of an image. This measure of random variations is interpreted as a measure of complexity for an image, and this concept is extended for grouping or clustering images based on complexity of the term *S*. The clustering of images is based on equal levels of gradient complexity *S*. Therefore, The gradient spin excess is a way of designating an image's level of complexity by a single number *S*.

The images in Figure 2 are examples of level of complexity of images based on the gradient spin excess value *S*. The absolute value of the spin excess |*S*| is used rather than S in order to take into account images that are similar within the domain of numbers that are equal but opposite in sign (derived from experimental observations). The item to note in Figure 2a and Figure 2b is that large values of |*S*| designate images of lesser complexity and smaller values of |*S*| designate images of greater complexity.

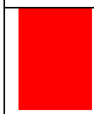
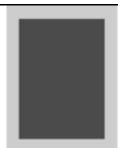
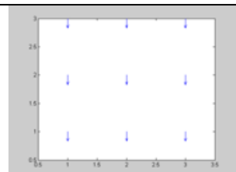
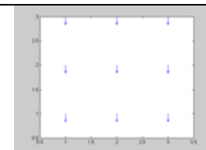
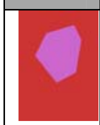
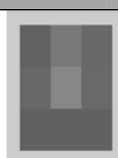
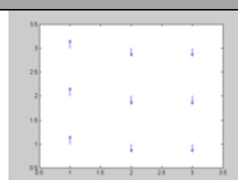
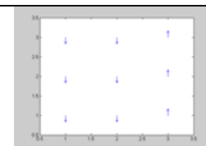
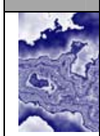
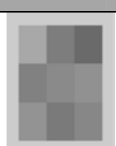
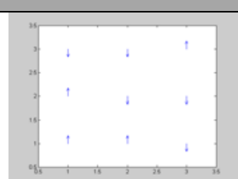
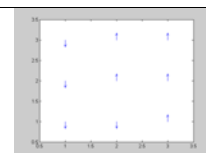
| Image | 3x3 Sub-blocks | Pixel Moment Orientations | \|S\| | | Image | 3x3 Sub-block | Pixel-Moments | \|S\| |
|---|---|---|---|---|---|---|---|---|
|  |  |  | 9 | |  |  |  | 9 |
|  |  |  | 3 | |  |  |  | 3 |
|  |  |  | 1 | |  |  |  | 1 |

**Figure 2 Examples of levels of complexity**

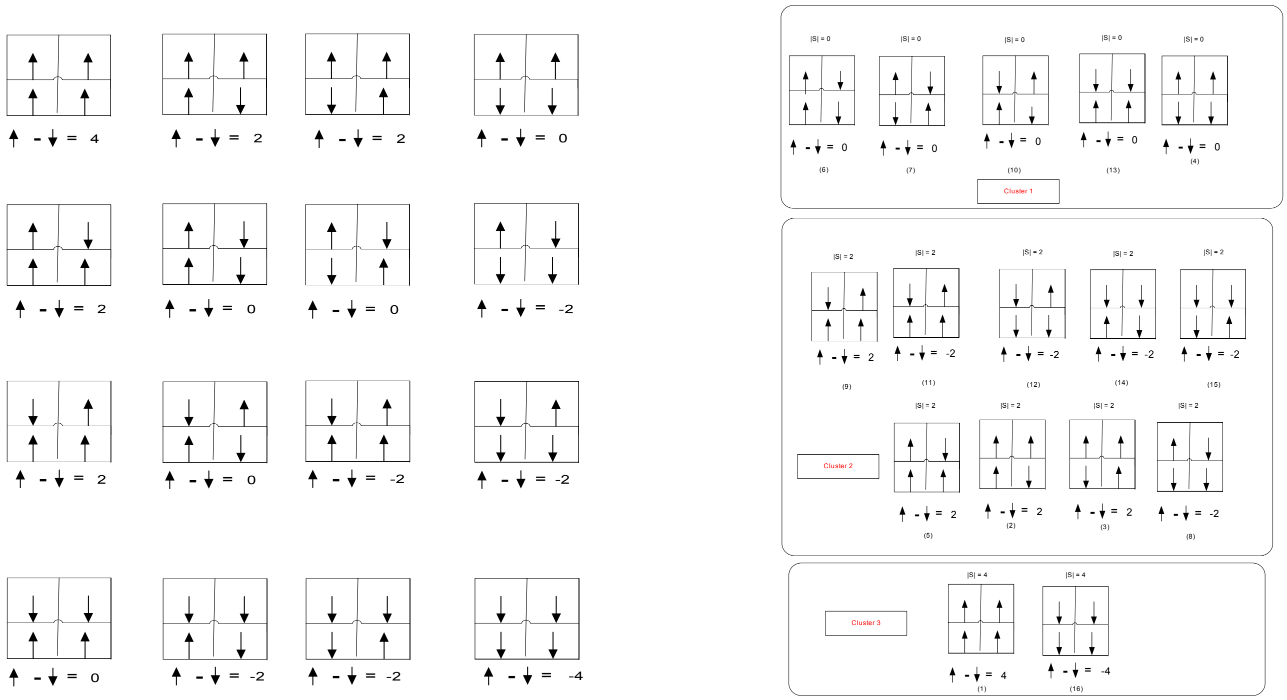
From this observation, one can assert from examining the level of disorder of pixel moments a possible one-one-correspondence of image complexity and that of information theoretical concepts. Suffice it to say, this concept of complexity derived from the gradient is applied to clustering an image database by grouping images based on levels of complexity.

### 6.3. Gradient Theoretic Clustering using Pixel Moments

This section describes the steps involved in clustering an image database using the enumeration of gradient states methodology. For demonstration purposes, images divided into 4(2x2) non-overlapping sub-blocks are used to describe the clustering methodology and process, and are depicted in the series of figures described below. The actual experiment for the research used images divided into 9 (3x3) non-overlapping sub-blocks based on the suggested research of Ko, Lee, and Byun [29].

If we consider the ideal case that consists of 4(2x2) non-overlapping sub-blocks for an image database of 16 images, all the possible object orientations of pixel moments in an image based on 4 (2x2) non-overlapping image sub-blocks are shown in Figure 3 with the corresponding value of gradient spin excess *S* (represented as $\uparrow - \downarrow$). Figure 3(a) represents $2^4$ or 16 possible combinations of states for an image using four pixel-moments. Taking the absolute value of the gradient spin excess *S* and collecting like values of |*S*|, one obtains the clusters depicted in Figure 3(b). In a similar way, the case of the 9(3x3) sub-blocking used for this study yields a total of 5 clusters, with gradient spin excess |*S*| values of 1, 3, 5, 7 and 9.

Figure 3 represents the pre-processing stage that prepares the database for query image retrieval. In the post-processing stage the query image searches for a matching cluster value of |S|. In addition, the query image is allowed to search adjacent clusters in order to take into account images that have extreme rotation, translation, and scaling. Examples of clustering with rotation, translation, and scaling from the University of Washington ground truth database are shown in Figure 4. Figures 4(a) to 4(d) represent clusters of similar images in various poses, while Figure 4(e) is an example of two similar images, one occluded in comparison to the other (both images located in adjacent clusters).




(a) (b)

**Figure 3 All possible spin orientations of 4 (2x2) Sub-Blocks (a) and Clusters obtained from edge information or gradient spin excess (b)**

**a) Two images in the same cluster**

**Translated images.**

**b) Two images in the same cluster**

**Scaled images .**

**c) Two images in same cluster**

**Similar images.**

**d) Two images in the same cluster**

**Scaled and reversed images.**

**e) Two images in two different clusters**

**Left image occluded**

**Figure 4 Clustering of relevant images (examples from the University of Washington Database)**

The fundamental idea for searching clusters in general using the EGS algorithm (for a 9(3x3) sub-blocked image database) is depicted in Figure 5. If the query image maps to cluster 3 based on the |*S*|-value match, the two adjacent clusters 2 and 3 are also searched. If an endpoint is chosen, such as 1 or 5, the algorithm chooses the next cluster, shown as arrows emanating from 1 to 2, and 5 to 4 in Figure 6. The time complexity for this part of the retrieval is O(1). After the query image marks its starting cluster and consequently chooses adjacent clusters, the $L_1$-norm is executed on this subset of clusters**.**

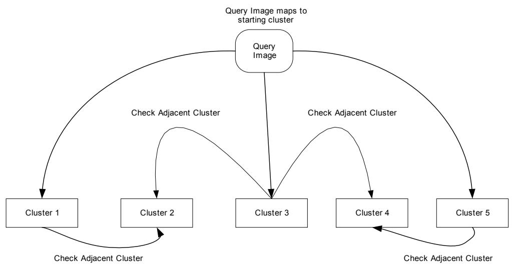



**Figure 5** The heuristic used for choosing the appropriate clusters**.**

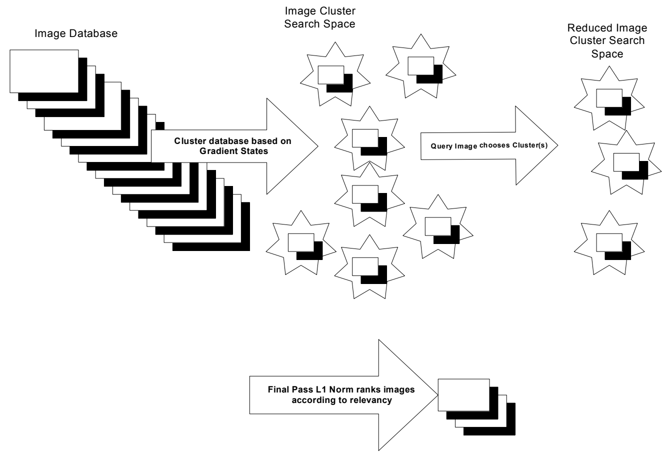



**Figure 6** Enumeration of Gradient States Process

### 6.4. EGS Algorithm

Although a clustering technique has been proposed for clustering images into similar or relevant clusters based on the complexity metric (the gradient spin excess), there is still a need to further filter out the images that most closely relate to the query. The use of this clustering methodology as the sole measure of similarity is inappropriate without the use of an additional technique. Color histogram comparisons using the $L_1$ norm are used as a second pass over a chosen cluster to pull out the most relevant images since two dissimilar images may be in the same cluster due to identical values of complexity (gradient spin excess). After all relevant images are pulled out and sorted according to the $L_1$ norm criterion, the user then examines this ranked list starting from the top image.

In this discussion, we do not assert that the enumeration of gradient states methodology is capable of providing a meaningful similarity measure based on gradient spin excess values alone. This should not be very surprising to the reader since such an assertion would suggest that a single real number contains more information than a vector for distinguishing between two images (as with the $L_1$-norm). The vector always contains more information than the single real number, particularly since the single number is an aggregation of the vector. Hence, we integrate the gradient states methodology with the $L_1$-norm in order to improve the performance of the retrieval process. The fundamental idea for the enumeration of gradient states (EGS) algorithm is graphically summarized in Figure 6.

## 7. Testing

During the testing phase, a master database consisting of 670 images from the University of Washington (located at the following URL: http://www.cs.washington.edu/research/imagedatabase/groundtruth) was used. This database was selected for the following reasons:

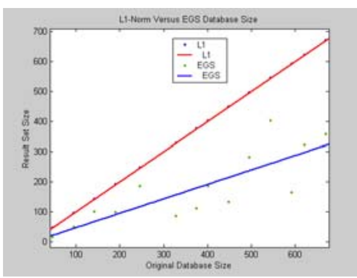
1. The master image database provided at the above URL contained a wide variety of realistic scenes (no synthetic scenes were used), such as images of animals, humans in various activities, landscapes, and architecture. The image database was initially assumed not to be dominated by any class of images (e.g., galaxies) and is approved by many researchers as a good start for CBIR research [34].
2. The 670 images provided by the University of Washington Computer Science Department was adequate for CBIR testing [34]. The master database was randomly ordered for testing.
3. The image database was conveniently divided into various themes by the University of Washington Computer Science Department which included a complete collection of relevant or similar groupings of images as determined by a set of experts from the University of Washington [35].
4. The benchmark was based on fourteen query images randomly chosen (in order to minimize human subjectivity) from similarity groups of the University of Washington database.

## 8. Discussion of Results

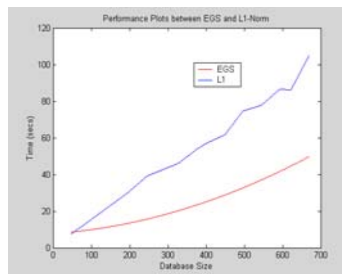
Various precision-recall measurements depicted in Table 1 were performed in order to verify the consistent behavior of the EGS methodology in comparison with the $L_1$-norm. The four metrics used were: 1) recall and precision, 2) average precision at seen relevant images, 3) r-precision, and 4) harmonic mean. All four techniques are standard techniques used in CBIR research for performance evaluation and are explained in detail in [36]. As depicted in Table 1, the EGS algorithm for all four techniques exceeds the retrieval performance of the $L_1$-norm with an overall average performance of 12% improvement over the $L_1$-norm. In addition to precision improvement, for each of the 14 query images used in the experiment, the intermediate search space produced by the EGS algorithm reduced the master database by an average factor of 2.4:1, or approximately 2/5$^{th}$ the master database (see Figure 8a). The runtime performance of the EGS algorithm compared to the $L_1$-norm across 14 databases of increasing size is depicted in Figure 8b. The average time of improvement ratio over the $L_1$-norm is approximately 2.4:1. In other words, the EGS algorithm is faster than the $L_1$-norm since the EGS result database set is commensurately smaller than the result sets for the $L_1$-norm.

| **Technique** | **% increase over $L_1$-norm** |
|---|---|
| | |
| Precision vs. Recall | 17 |
| Average Seen Precision | 12 |
| Average R-Precision | 9 |
| Harmonic Mean | 10 |
| Overall Average | 12 |

**Table 1 Various recall-precision measurements**




**(a)Reduced database set (b) Top curve: $L_1$-norm**
**Top curve: $L_1$-norm/Bottom Curve: EGS**




**(b) Runtime performance:**
**Top curve: $L_1$-norm/ Bottom Curve: EGS**

**Figure 8 Performance Curves**

The test results show a consistent behavior of performance for the EGS algorithm over that of the $L_1$-norm. Such consistent behavior gives confidence in the data gathered during the testing phase. The EGS algorithm is an improvement over one reported source which is also compared to the $L_1$-norm, namely Gagaudakis [5] who reported a 10% precision improvement of their algorithm over the $L_1$-norm. The tests show that the EGS algorithm warrants further research attention.

## 9. ANOVA Study

Analysis of Variance (ANOVA) was used in this study to show that there is a significant level of performance difference between EGS and $L_1$-norm precision measurements and performance curves. The purpose of one-way ANOVA is to find out whether data from several groups have a common mean or to determine whether groups are actually different in measured characteristics. Currently, no standard metric exists for measuring the performance of algorithms in the CBIR field. Hopefully, this research will open the door for such a standard using ANOVA techniques.

ANOVA was applied to the EGS retrieval performance values depicted in Table 1 and Figure 8 using a pre-set level of significance of 10%. ANOVA indicated that the p-value (the probability of observing the given sample result under the assumption that the null hypothesis is true) is less than the preset value of 10%, favoring the EGS algorithm's performance over the $L_1$-norm.

## 10. The Issue of Standardizing CBIR Research

To date, the issue of image database standards for CBIR research is not addressed adequately by CBIR researchers. With an acceptable standard image database, one can compare different algorithms more efficiently. One must realize that the art of evaluating image retrieval is quite chaotic in that researchers design different algorithms and then test the performance on their own test-beds [37]. Rao, Srihari, Zhu and Zhang comment on this issue: "… there is no common test-bed, and there is no theory about how to compare different test-beds. The lack of a uniform evaluation methodology is clearly a limiting factor in the development of the multimedia retrieval field." It is the attempt of this research to use an image database that is gaining acceptance in the CBIR community, namely the University of Washington image database.

In further emphasizing the lack of existing standards in CBIR research, Muller, Squire, Marchand-Maillet, and Pun have the following to say about this issue [34]:

"Several problems must be addressed in order to create a common image collection. The collection must be available free of charge and without copyright restrictions, so that images can be placed on the web and used in publications. *The greatest problem is to create a collection with enough diversity to cater for the diverse, partly specialized domains in CBIR such as medical images, car images, face recognition and consumer photographs"* (emphasis added). They go on to say that "an alternative approach is for CBIR researchers to develop their own collection. Such a project is underway at the University of Washington in Seattle (Annotated groundtruth database, Department of Computer Science and Engineering, University of Washington, http://www.cs.washington.edu/research/imagedatabase/groundtruth/). This collection is freely available without any copyright and offers annotated photographs of different regions and topics. … The collection size should be sufficiently high that the trade-off between speed and accuracy can be evaluated." This is the standard database used for this research and is supported by the above researchers. Although the database has some shortfalls as indicated at the beginning of this section, it is considered by Muller, Squire, Marchand-Maillet, and Pun as a good start for any research in CBIR for using different themes or topics.

## 11. Conclusion and Future Effort

The objective of this research was to show that a new clustering technique integrated with the $L_1$-norm held some promise as an indexing scheme for retrieval of relevant images. The primary concept of relevancy employed in this study is that similarity for images is based on the similarity of the absolute value of the gradient spin excess designated as $|S|$, derived from the difference in the direction of pixel moments of gradient spin-up and gradient spin-down. This value of $|S|$ was also interpreted as a complexity value in which images with similar complexities fall into similar clusters. As a result, there are many aspects of this research open for future investigations. Some future areas of consideration are as follows:

1. Scalability: The research briefly examined sub-blocks and their relationship to the size of the database. Higher sub-blocks on a fixed database size produce higher precision during the retrieval process. A point is reached in this process in which the precision does not improve. This is to be expected (with a constant image database) since as the sub-blocks are increased, there becomes a point where the EGS algorithm retrieves only those clusters predominately represented by the relevant images. It was also noticed in this preliminary investigation of scalability, as the sub-blocks increase, there are fewer and fewer number of false positives that appear within the retrieved clusters. The issue of scalability is to be addressed in a future article.
2. Information Theory: The enumeration of gradient states poses an interesting investigation into the field of information theory and the relationship of the gradient states methodology to entropy and to statistics in general. The notion of gradient spin excess is a concept derived from entropic properties of thermal physics [33].

The application of "spin" to image interpretation and its application to the field of CBIR poses many questions for future exploration. The concept of complexity and clustering as it relates to shape (the gradient) was discussed in this paper. One area that warrants further investigation is the extension of this concept of complexity to other visual features, such as texture and regions, and possibly the formulation of a consistent general theory of image complexity.

## REFERENCES


[1] J. Eakins and M. Graham, "Content-Based Image Retrieval: A Report to the JISC Technology Applications Programme," Institute for Image Data Research, University of Northumbria http://www.unn.ac.uk/iidr/report.html, 1999.

[2] J. Zachary, "An Information Theoretic Approach To Content Based Image Retrieval,"Doctoral Thesis at *Computer Science*. Baton Rouge: Louisiana State University, 2000.

[3] M. Stricker and M. Swain, "The Capacity and the Sensitivity of Color Histogram Indexing," Communications Technology Lab, ETH-Zentrum, Technical Report Technical Reprot 94-05, 1994 1994.

[4] M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Vision*, **vol. 7**, pp. 11-32, 1991.

[5] G. Gagaudakis and P. L. Rosin, "Incorporating shape into histograms for CBIR," *Pattern Recognition*, **vol. 35**, pp. 81-91, 2002.

[6] J. R. Smith, "Integrated Spatial and Feature Image Systems: Retrieval, Analysis, and Compression,"Doctoral Thesis at *Columbia University Department of Computer Science*: Columbia University, 1997.

[7] J. R. Smith and S. Chang, "Tools and Techniques for Color Image Retrieval for Color Image Retrieval," presented at SPIE Proceedings of Storage and Retrieval for Image and Video Databases IV, San Jose, CA, vol. 2670, pp. 426-437 1995.

[8] A. Goodrum, "Image Information Retrieval: An Overview of Current Research Techniques, Promising Directions, and Open Issues.." *Journal of Visual Communication and Image Representation*, **vol. 3**, pp. 63-67, 2000.

[9] Y. Rui, T. Hung, and S.-F. Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," *Journal of Visual Communication and Image Representation*, pp. 1-23, 1999.

[10] D. C. Tseng, Y. F. Li, and C. T. Tung, "Circular histogram thresholding for color image segmentation," presented at Proceedings of the Third International Conference of Document Analysis and Recognition, Montreal, Canada, vol. 2, pp. 673-676 1995.

[11] N. Ikonomatakis, K. N. Plataniotis, M. Zervakis, and A. N. Venetsanopoulos, "Region growing and region merging image segmentation," presented at Digital Signal Processing Proceedings, DSP 97., 1997 13th International Conference, vol. 1, pp. 299-302 1997.

[12] M. D. G. Montoya, C. Gil, and I. Garcia, "Load balancing for a class of irregular and dynamic problems: region growing image segmentation algorithms," presented at Proceedings of the Seventh Euromicro Workshop, Funchal, Portugal, pp. 163-169 1999.

[13] O. Schwartz and A. Quinn, "Fast and accurate texture-based image segmentation," presented at Image Processing, Proceedings on International Conference, Laussane, Switzerland, vol. 3, pp. 121-124 1996.

[14] R. Battiti, "Low level image segmentation with high level 'emergent properties': color based segmentation," presented at Industrial Applications of Machine Intelligence and Vision, International Workshop, Tokoyo, Japan, vol. 1, pp. 128-132 1989.

[15] A. Hoogs and R. Bajcsy, "Model-based learning of segmentations," presented at Pattern Recognition, Proceedings of the 13th International Conference, vol. 4, pp. 494-499 1993.

[16] Y. Rui, A. C. She, and T. S. Huang, "Modified Fourier descriptors for shape representation—a practical approach," presented at Proc. of First International Workshop on Image Databases and Multimedia Search, Amsterdam, The Netherlands 1996.

[17] E. N. Mortensen and W. A. Barret, "Interactive segmentation with intelligent scissors," *Graphical Models and Image Processing*, **vol. 60**, pp. 349-384, 1998.

[18] E. M. Gurari and H. Wechsler, "On the difficulties involved in the segmentation of pictures," *IEEE Trans.*, **vol. 4**, pp. 304-306, 1982.

[19] I. L. Dryden and K. V. Mardia, *Statistical Shape Analysis*. New York: Wiley & Sons, 1998.

[20] M. Peura and J. Iivarinen, "Efficiency of simple shape descriptors," presented at The 3rd International Workshop on Image Analysis and Processing, Singapore, pp. 443-451 1997.

[21] W. I. Grosky and Y. Tao, "Spatial Color Indexing: A Novel Approach for Content-Based Image Retrieval," presented at IEEE 6th International Conference on Multimedia Computing and Systems (ICMCS'99), Florence, Italy, pp. 530-535 1999.

[22] R. Gonzales and R. Woods, *Digital Image Processing*. Reading, MA: Addison Wesley, 1992.

[23] A. Engel and V. d. Broeck, *Statistical Mechanics of Learning*. Cambridge, UK: Cambridge University Press, 2001.

[24] C. Faloutsos, M. Flickner, W. Niblack, Petkovic, W. Equitz, and R. Barber, "Efficient and effective querying by image content," *Technical Report, IBM Research Report*, 1993.

[25] M. Gorkani and R. Picard, "Texture orientation for sorting photos at a glance," presented at Proceedings of the 1994 International Conference on Computer Vision and Pattern Recognition, Seattle, Washington, vol. 1, pp. 459-464 1994.

[26] T. S. Chua, K.-L. Tan, and B. C. Ooi, "Fast signiture-based color-spatial image retrieval," presented at Proc. IEEE Conf. on Multimedia Computing and System, vol. 1, pp. 362-369 1997.

[27] E. Yiu, "Image Classification using color cues and texture orientation,"Masters Thesis at *Department of EECS*. Boston: MIT, 1996.

[28] P. Thor, "Indexing a Digital Image Database with Color Vectors,"Doctoral Thesis at *Computer Science*. Colorado Springs: Colorado Technical University, 1999.

[29] B. Ko, H.-S. Lee, and H. Byun, "Image Retrieval Using Flexible Image Subblocks," presented at Proceedings of the 15th Symposium on Applied Computing, Como, Italy, vol. 2, pp. 574-578 2000.

[30] Y. Gong, H. Zhang, H. Chuant, and M. Skauuchi, "An Image database system with content capturing and fast image indexing abilities," presented at Proceedings of the International Conference on Multimedia Computing and Systems, Boston, MA, pp. 121-130 1994.

[31] D. Stan and I. Sethi, "Mapping low-level image features to semantic concepts," presented at Storage and Retrieval for Media Databases , *Proceedings of SPIE*, San Jose, California, vol. 4315, pp. 1-15 2001.

[32] A. Engel and V. D. Broeck, *Statistical Mechanics of Learning*. Cambridge, UK: Cambridge University Press, 2001.

[33] Kittel, *Thermal Physics*. New York: John Wiley and Sons, 1969.

[34] Muller, Squire, and Marchand-Maillet, "Performance evaluation in content-based image retrieval: overview and proposals," *Pattern Recognition Letters*, **vol. 22**, pp. 593-601, *2001*.

[35] L. Shapiro and Y. Li, "Personal Contact and Web site investigation at University of Washington," 2002.

[36] R. Baeza-Yates and B. Ribeiro-Neto, *Modern Information Retrieval*. Menlo Park, CA: Addison-Wesley, 1999.

[37] A. Rao, R. K. Srihar, L. Zhu, and A. Zhang, "A Method for Measuring the Complexity of Image Databases," presented at IEEE Transactions on Multimedia, vol. 2, pp. 160-173 2002.