

Content-based Image Retrieval Using Data Fusion Strategy

데이터 융합을 이용한 내용기반 이미지 검색에 관한 연구

Woojin Paik*
Sun-Eun Jung**
Gi-Young Kim**
Euigun Ahn***
Moon-Sun Shin****

ABSTRACT

In many information retrieval experiments, the data fusion techniques have been used to achieve higher effectiveness in comparison to the single evidence-based retrieval. However, there had not been many image retrieval studies using the data fusion techniques especially in combining retrieval results based on multiple retrieval methods. In this paper, we describe how the image retrieval effectiveness can be improved by combining two sets of the retrieval results using the Sobel operator-based edge detection and the Self Organizing Map(SOM) algorithms. We used the clip art images from a commercial collection to develop a test data set. The main advantage of using this type of the data set was the clear cut relevance judgment, which did not require any human intervention.

초 록

지금까지의 정보검색 연구에서 데이터 융합 기법을 이용한 문서 검색은 하나의 알고리즘에 의한 검색에 비하여 많은 경우에 효율성이 높은 결과를 얻을 수 있었다. 하지만 이미지 검색에서 상이한 알고리즘을 이용한 다수의 검색 결과를 합쳐 하나의 검색결과를 얻는 데이터 융합 기법의 사용은 많지 않았다. 이 연구에서는 소벨 연산자를 이용한 윤곽선 검출과 자기조직화 지도 알고리즘에 의한 두 검색 결과를 융합하여 각각의 알고리즘에 의한 검색결과보다 높은 효율성을 보여주는 방법을 제시하였다. 이 연구에서는 상용 클립아트 이미지를 이용하여 사람의 주관적인 적합성 판단을 배제한 검색 실험 데이터를 만들어 사용하였다.

Keywords: content-based image retrieval, data fusion, Sobel edge detection algorithm, Self-Organizing Map(SOM) algorithm, clip art images
내용기반 이미지 검색, 데이터 융합, 소벨 윤곽선 검출, 자기조직화 지도, 클립아트 이미지

* Associate Professor, Dept of Computer Science, Konkuk U.(wjpaik@kku.ac.kr)
** Undergraduate Students: Dept of Computer Science, Konkuk U.
(jseking85@kku.ac.kr, progmoni@gmail.com)
*** Graduate Student, Dept of Library & Information Science, Yonsei U.(ahn98@yonsei.ac.kr)
**** Corresponding Author, Lecturer, Dept of Computer Science, Konkuk U.(msshin@kku.ac.kr)
■ Received : 13 May 2008 ■ Revised : 23 May 2008 ■ Accepted : 5 June 2008
■ Journal of the Korean Society for Information Management, 25(2): 49-68, 2008.
[DOI:10.3743/KOSIM. 2008.25.2.049]

1. Introduction

Most of the information retrieval experiments claimed that the data fusion strategy, namely multiple evidence combination using multiple query formulations or multiple retrieval schemes, were more effective than the single evidence based retrieval(Hsu and Taksa 2005). Although, there were a number of studies describing the evidence combination strategies for the image retrieval tasks, the data fusion occurred mainly at the feature combinations(Yavlinsky, et al. 2004). There were also studies, which combined the textual description and the image features, when searching for the web images(Aslandogan and Yu 2000).

Our goal was to develop a highly effective content-based image retrieval system using the data fusion strategy of combining the results from two different retrieval schemes. One scheme generated a similarity ranked list of images with respect to the query image. This scheme was based on the Sobel edge detection algorithm, which identifies all the lines, that outlines the objects in an image, as the feature representation of the images. The other scheme was the Self Organizing Map(SOM) algorithm, which generates a set of clusters of the similar images without any relation to the query image. The SOM cluster based image retrieval occurs by identifying the cluster, of which the query image should have been a member. Thus, the SOM cluster based image retrieval systems returns boolean type results namely a set of images similar to the query image without any rank or order.

We used a data fusion method commonly employed by the two-stage retrieval systems where the fast boolean retrieval results were rearranged according to the results from the ranking-based retrieval system(Mase et al. 2005).

2. Previous Works

There were a number of studies showing the retrieval performance improvements by taking into accounts of the different results of the different query formation(Saracevinc and Kantor 1988; Belkin et al. 1994; Fox and Shaw 1994). Lee(1997) showed why the evidence combination of different retrieval runs produced similar relevant documents while retrieving different non-relevant documents. Vogt and Cottrell(1999) reported the linear combination model of information retrieval results fusion where the combined results was generated by a weighted sum of the scores from each of the component systems. They found that the linear combination model works best when the component systems could achieve high effectiveness where there was a large overlap of relevant documents and a small overlap of non-relevant documents.

Aslam and Montague(2001) defined an external metasearch as the after-the-fact combining of complete search engines in comparison to another type of metasearch where various types of internally available evidences within one search engine were combined to produce single result. Our approach can be described as a type of an external meta-

search scheme.

In many image retrieval studies, the main focus was on the use of relevance feedback to determine the relative weighting of the features extracted from the images (Rui et al. 1998; Rui and Huang 1999). It is necessary to have at least a few relevant images for the relevance feedback scheme to work. Thus, the quality of the initial retrieval was important in this case. Some of the combination strategies used in the image retrievals were taking the minimum, maximum, and the sum of the similarity scores of the retrieved image. In addition, BordaFuse, which simulated the process of voting by treating the images as candidates and features as voters. The final similarity scores for each image was calculated by adding rank values based on each feature (Yavlinsky et al. 2004).

Our data fusion-based image retrieval system used the component systems. One was based on the Sobel edge detection algorithm generated image features. The edge detection is considered to be an essential information in image retrieval. It is because that identifying the edges inside one image is the first step toward recognizing geometric shapes within one image (Amata and Lecce 2003). The Sobel operator is used to compute the gradient of the image intensity at each point. The result gives the direction of the largest possible increase from light to dark and the rate of change in that direction. This reveals how the image changes at that point. Thus, it is possible to determine whether that part of the image represents an edge (Gonzalez and Woods 2002). PicSom image retrieval sys-

tem (Sjoberg 2006) used the Sobel edge detection results as the image features.

Another component of our system was based on the SOM algorithm. Paik et al. (2007) described how the SOM algorithm was used to implement a content-based trademark image retrieval system. They used the visual features, which were derived from the gray histogram representation of the images. aiSomPic (Chesnut 2004) and PicSOM (Sjoberg 2006) were another studies using the SOM algorithm to develop image retrieval systems. aiSomPic used color histogram, color area, texture histogram, and texture area as the bases for generating the image features. PicSom used image shape fourier transformation result, color, and texture to represent the images. Both systems clustered the images using the SOM algorithm. The images in the same cluster were considered to be similar.

3. Research Method

To test our image retrieval system, we develop a test data set using a commercially available clip art images. In the original clip art collection, there were 800,000 images, which were organized by 86 major categories. The major categories were further divided into hundreds of sub-categories (Nova Development Corp. 2004). We randomly selected 19 sub-categories excluding the ones representing the English alphabet transformations based shapes. Then, six images, which belonged to these sub-categories were chosen randomly. In

summary, there were 114 images in the test data set. Only the images in the same sub-category were considered to be similar. This means that there were six similar including the query image itself and 108 not similar images for each image in the test data set. This artificial creation of the test data set allowed us to avoid the subjective human relevance judgments. The size of each image was 100 pixel by 100 pixel. The categories and their

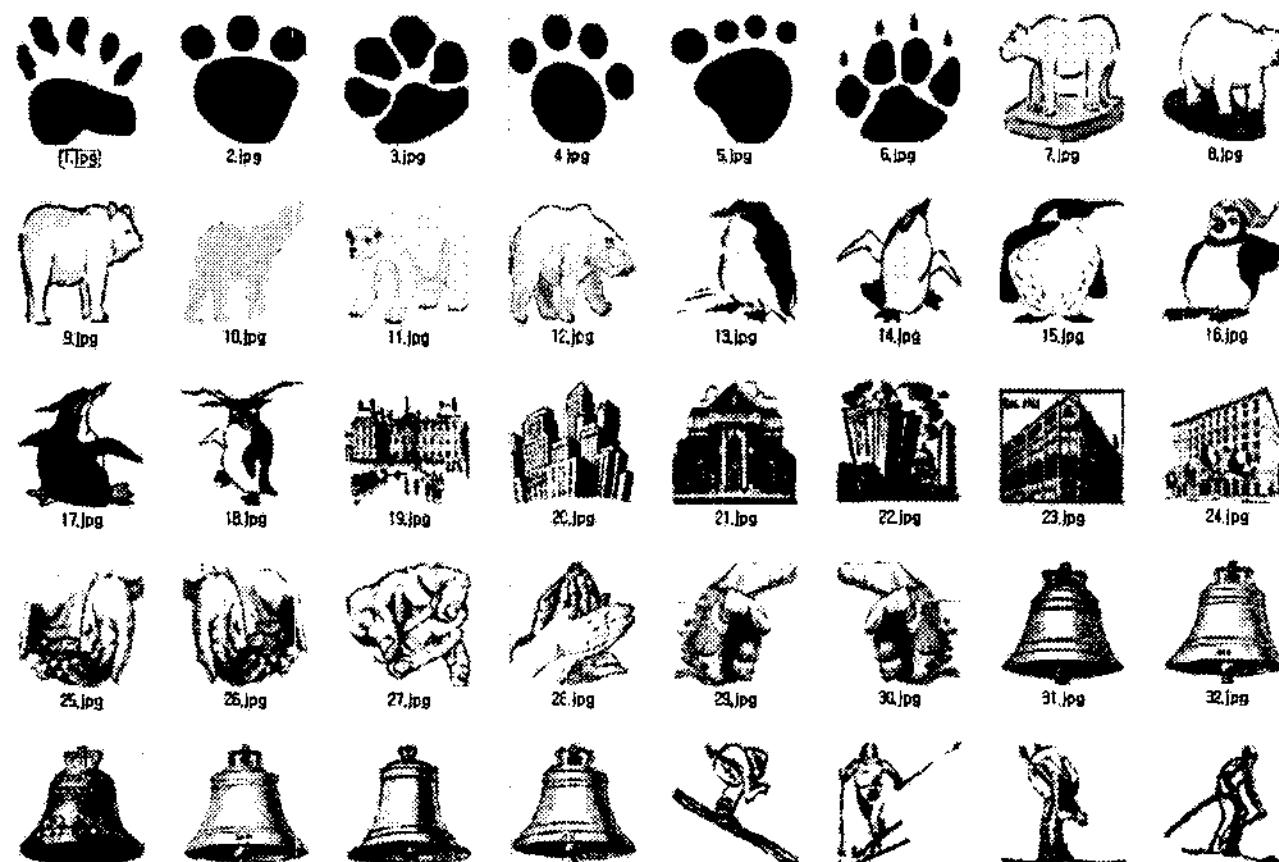
descriptions are shown in the <Table 1>.

The <Figure 1> shows some of the images in the test data set. The first six images belong to the ‘sole of foot’ sub-category. The second set of six images belong to the ‘bear’ sub-category.

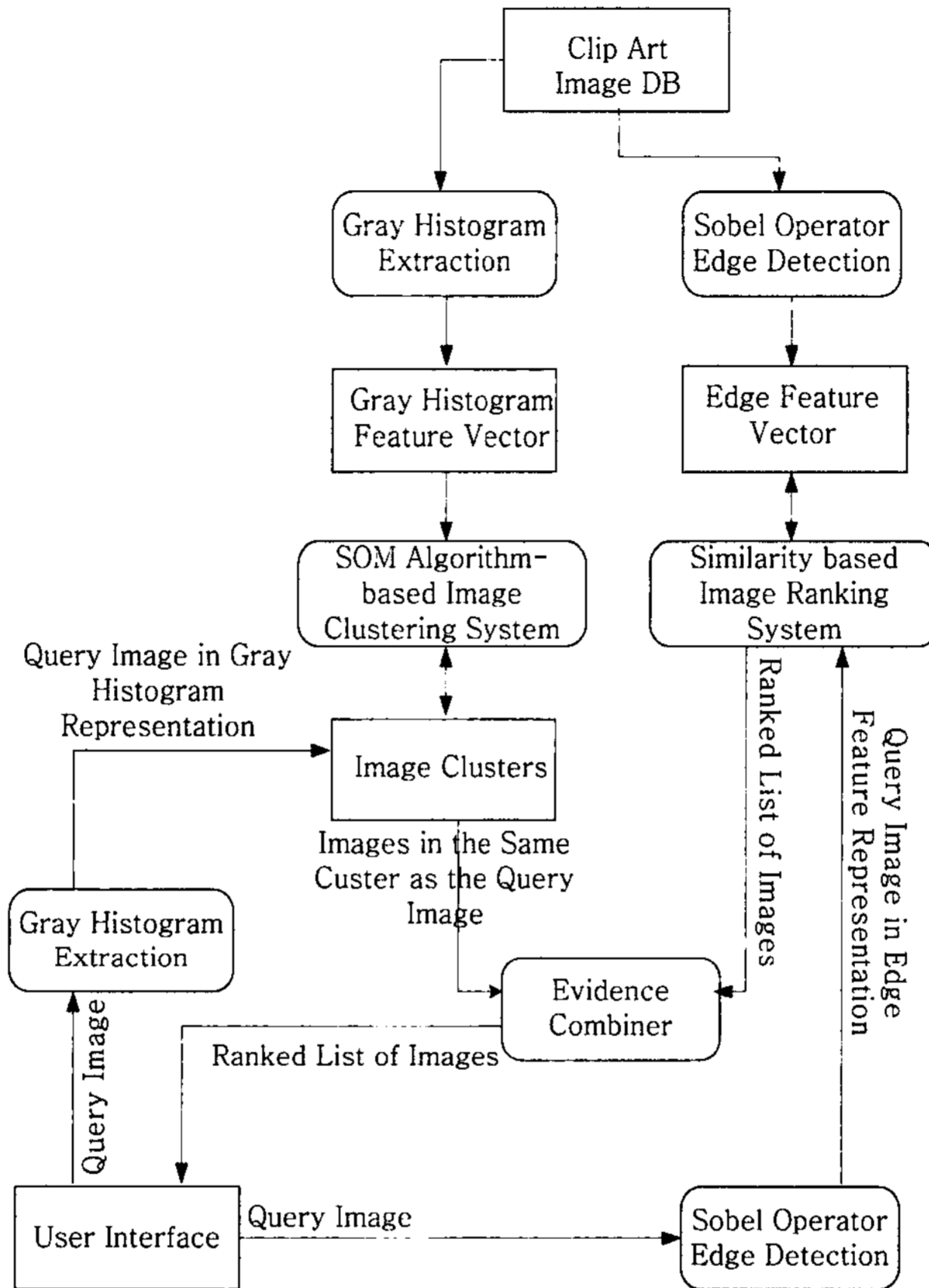
The <Figure 2> shows the data flow diagram of the system developed for the experiments conducted in this study. The test data set was stored as a Clip Art Image Database.

<Table 1> 19 Clip Art Sub-categories

Category ID	Description	Category ID	Description
1	sole of the foot	11	fish
2	bear	12	person playing baseball
3	penguin	13	bat
4	building	14	UFO
5	hand movements	15	traffic signal
6	bell	16	gun
7	person skiing	17	wild flower
8	book	18	racing car
9	person playing basketball	19	computer
10	organ		



<Figure 1> Sample Images from the Test Data Set



<Figure 2> Data Flow Diagram of the System Used in this Study

For each image in the database, two types of features were extracted. One was the edge information extracted using the Sobel algorithm. The edge detection results were stored as the edge feature vectors. The other was the gray histogram representation of the image, which was stored as the gray histogram feature vector.

The query image was also processed by the Sobel operator based edge detection module to generate the query image in edge feature representation. This query representation was compared with the edge feature vector form of each

image in the clip art image database to generate the ranked list of images. This ranked list was passed to the evidence combiner.

The gray histogram feature vectors were used to cluster images using the SOM algorithm. The results from this process were stored as the image clusters. Once the query image was entered through the user interface, the query image was processed by the gray histogram extraction module to produce the query representation in gray histogram representation. This query representation was fed to the same formula, which was used to assign

each database image to the cluster, to find the most appropriate cluster with the similar images from the clip art image database. Then, the images in the selected cluster were passed to the evidence combiner.

The evidence combiner re-ranked edge feature based ranked list based on the images in the same cluster as the query image. This re-ranked list was passed to the user interface for the users to evaluate the image retrieval results.

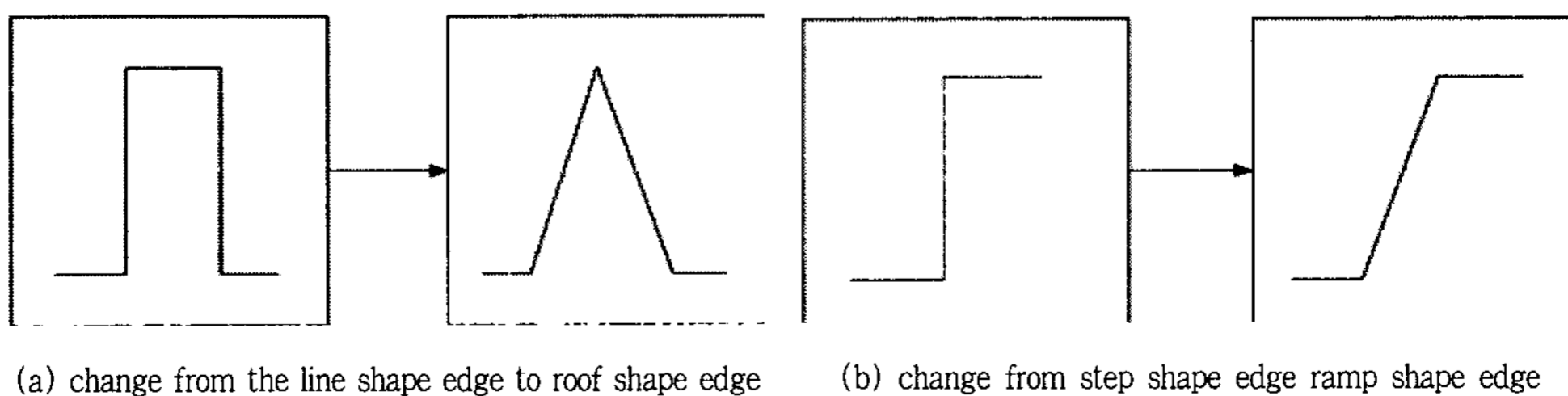
In this study, we evaluate the data fusion based image retrieval results against the known set of relevance judgments. Since the test data set was developed in such a way that there are six images from the same sub-category, each image has five similar images and 108 not similar images. This relevance judgments were used to measure the retrieval effectiveness of the Sobel algorithm based ranked list of the images, the SOM algorithm based clustering results, and the data fusion derived re-ranked lists. We compared three retrieval effectiveness figures to determine the relative advantage of the data fusion method against the Sobel and SOM algorithms.

3.1 Edge Detection Algorithm based Image Retrieval Using Sobel Operators

Edges refer to the boundary between an object and its surrounding background. Edges signal the change of the brightness value from low to high and vice versa. Edges are regarded as useful feature in image processing as they include various information such as the object shape or direction, which can be obtained regardless of image's color or texture. It is common to measure the slope of the edges to get the directional information about the image(Zio and Tabbone 1998).

It is also common to preprocess the images to reduce the noise before extracting the edge information. This preprocessing step often alters the edge shape as shown in the <Figure 3>.

There are many edge detection algorithms. One common type is referred as the gradient method, which detects the edges by looking for the maximum and minimum in the first derivative of the image. The Sobel algorithm is based on the gradient method. It calculates a two dimensional spatial gradient measurement on an image. The Sobel



<Figure 3> Edge Shape Change based on Image Preprocessing

algorithm computes the approximate absolute gradient magnitude at each point in an input gray-scale image. A pair of 3x3 convolution masks, one estimating the gradient in the x-direction, which is represented as columns, and the other estimating the gradient in the y-direction, which is represented as rows, are used in detecting edges using the Sobel algorithm.

A convolution mask is smaller than the actual image. As a result, the mask slides over the image, manipulating a square of pixels at a time. The Sobel masks are shown in the <Figure 4>. The magnitude of the gradient is calculated using [1] (Green 2002).

$$|G| = \sqrt{G_x^2 + G_y^2} \quad [1]$$

The <Figure 5> shows the Sobel algorithm based Edge Detection Result for the images shown in the <Figure 5>.

The edge detected images cannot be used directly for the image retrieval tasks. Thus, it is essential to convert the edge detected image into some form of a numeric representation to measure the similarities between the query image against the images in the database.

First, the center of the image was determined. Since the image size is 100 pixel by 100 pixel, the center of the image was the cross section of the middle point of the x-axis and the y-axis. We

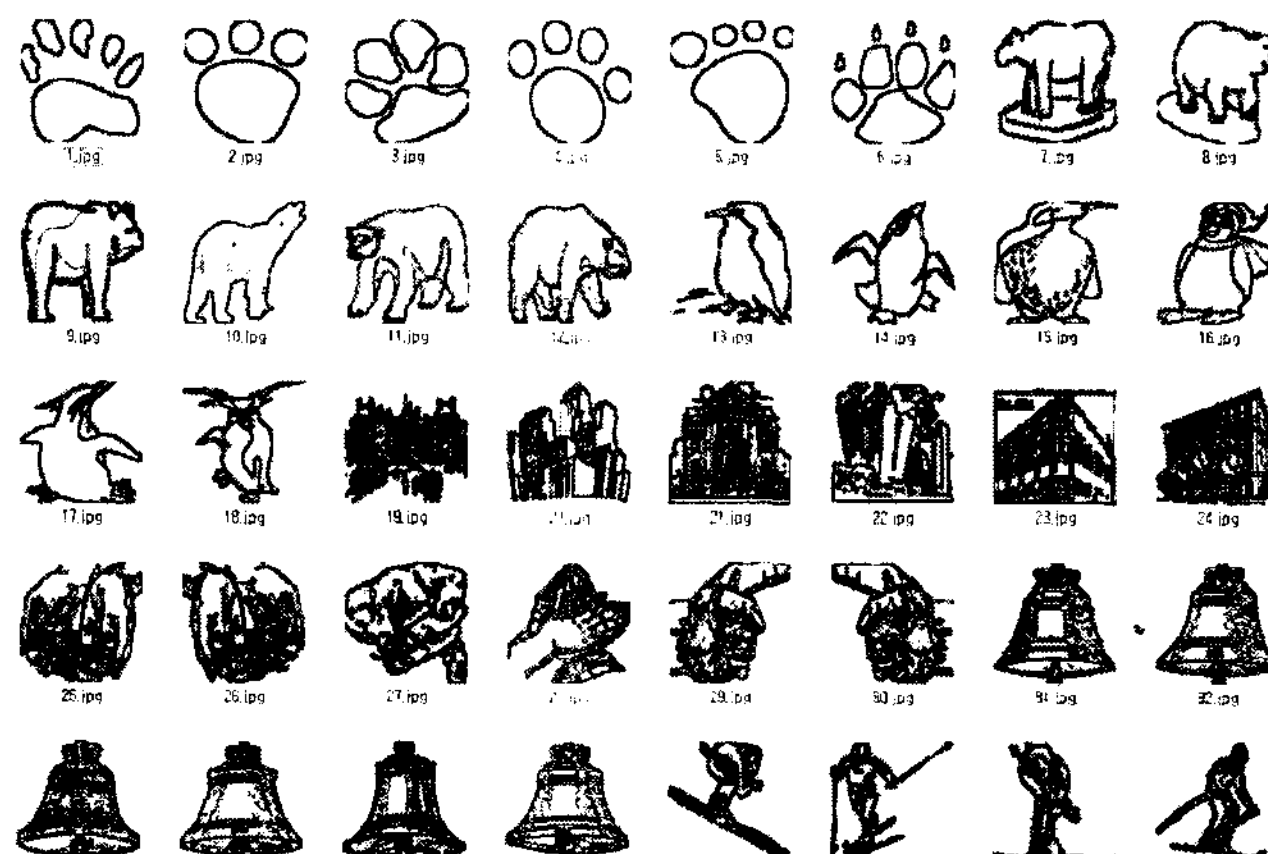
-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

<Figure 4> Sobel masks



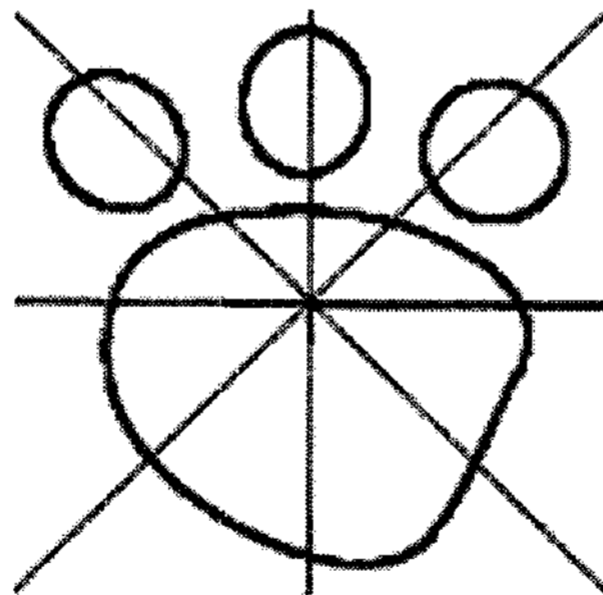
<Figure 5> Sobel algorithm based Edge Detection Result

defined the middle point of the x-axis as the 50th pixel from the left end of the image. The middle point of the y-axis was also defined as the 50th pixel from the top of the image.

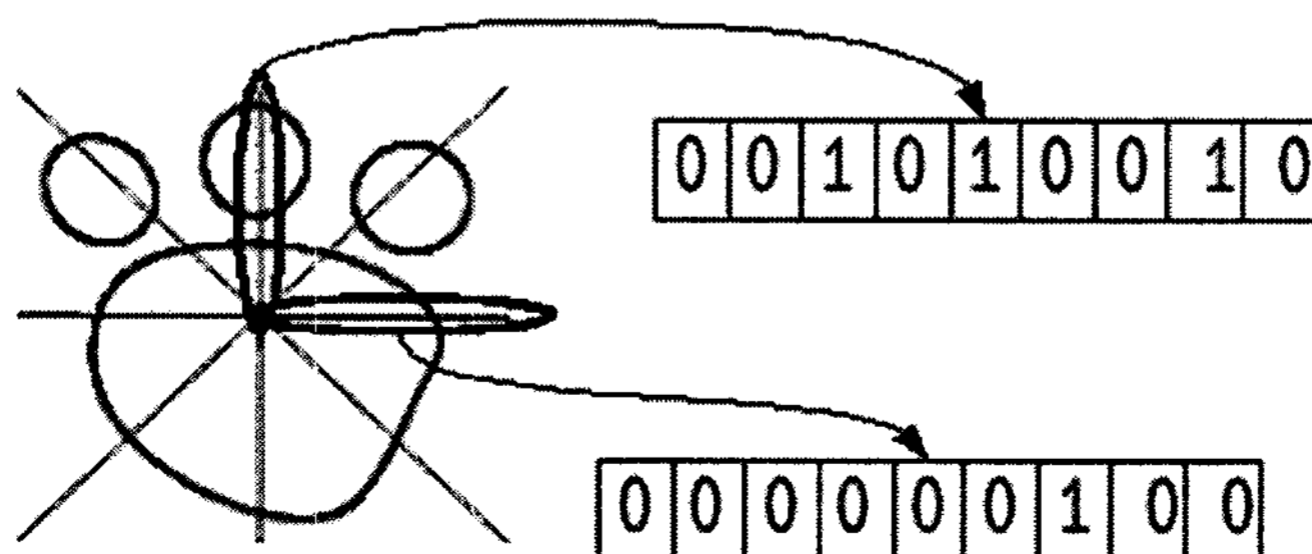
Eight straight lines were drawn from the center of the image to the end of the image by rotating from the 0th, 45th, 90th, 135th, 180th, 225th, 270th, and 315th degrees. The <Figure 6> shows an edge detected image with the lines drawn according to this method. This is one of the sole of a foot type clip art images.

For each line, the number of black pixels, which crosses the line, were counted. The resulting counts for an image became an eight dimension array where the first array element represented the num-

ber of black pixels crossed the line drawn to the 0th degree from the center of the image. The <Figure 7> shows an example of this counting method. For the line drawn to the 0th direction, the first two pixels from the center were white but the third pixel was black. The fourth pixel was white and the fifth pixel was black again. The eighth pixel was black. The nine cell table, which is linked to the 0th degree line shows this identification results. Similarly, only the seventh pixel from the center in the 90th degree line was the black. In summary, there were three black pixels crossing the line drawn to the 0th direction. However, there were only one black pixel crossing the line drawn to the 90th degree direction. Thus,



<Figure 6> Eight Lines Drawn on the Edge Detected Image



<Figure 7> Converting Edge Detected Image into Numerical Representation

this image should be represented as an eight dimension array with the values corresponding to 3,3,1,1,1,1,1,3. This information was stored in the database along with the image identification to be used during the similarity calculation between the query image to the database images.

Rotated images can cause significant degradation of the retrieval effectiveness in image retrieval especially when using the shape oriented features such as the edge information. In our study, we generated seven alternate edge detection-based representation of the database images by rotating the images 45 degree angle at one time. The numerical representation of each image based on the black pixel counting was generated again at each rotation. This means that each database image has eight numerical representations.

The ranked list generation by measuring the similarities between the query image and the database images were carried out by calculating the Euclidean distance between the images in eight dimension space. Since the query image also went through the edge detection and the eight dimension numeric array generation stages. Both the query image and the database images had the comparable representation. However, it should be noted that the query image was not rotated and thus had only one numeric representation. The Euclidean distance was calculated using [2].

$$\sqrt{(X_1 - Y_1)^2 + (X_2 - Y_2)^2 + \dots + (X_7 - Y_7)^2 + (X_8 - Y_8)^2}$$

[2]

where X_i is the i th value of query image's numeric representation and Y_i is the i th value of database image's numeric representation

After the Euclidean distance based similarity was measured between the query image and one of the database image, eight similarity scores are generated as there were eight numeric representation for each database image due to the image rotation. The minimum similarity score was selected as the final similarity score between the query image and the database image. The minimum is chosen as the best similarity score represented as the Euclidean distance is zero. Zero means that two images were identical.

The <Table 2> shows the partial output of the edge detection based similarity ranking by using the image with the identification one as the query image. The query image was compared with all 114 database images but this table only shows the similarity ranking for the first ten database images.

The value in the first column of the table is the identification for the database images. The second column value is the Euclidean distance calculation based similarity score without any rotation. The third column value is the similarity score when the database image was rotated 45 degree. The fourth to the ninth columns show the respective similarity scores when the database image was rotated to the degree shown in the first row.

The tenth column value is the minimum value among the similarity scores in the second to the ninth columns. The final column is the similarity

<Table 2> Edge Detection Based Similarity Ranking

ID	0 degree	45 degree	90 degree	135 degree	180 degree	225 degree	270 degree	315 degree	minimum (bet. 0, 315)	rank(J)
1	0.00	10.10	12.57	12.25	11.75	12.25	12.57	10.10	0.00	1
2	10.44	10.05	9.43	10.34	10.63	11.00	10.34	10.44	9.43	58
3	6.40	15.65	15.46	14.39	15.78	16.34	14.93	12.37	6.40	20
4	10.58	10.68	9.80	7.87	8.72	10.68	10.95	10.20	7.87	36
5	10.44	8.31	5.92	9.00	10.44	9.85	9.64	10.91	5.92	14
6	9.54	12.77	13.00	10.54	8.19	6.86	11.45	8.77	6.86	26
7	9.17	10.00	10.49	9.70	7.21	7.75	8.60	6.00	6.00	16
8	9.80	11.83	11.75	11.49	10.86	8.25	11.14	6.78	6.78	25
9	10.77	10.49	9.06	9.80	9.90	8.72	3.46	8.72	3.46	2
10	12.29	12.29	12.29	12.29	12.29	12.29	12.29	12.29	12.29	74

based ranking from the lowest to the highest scores. The first database image was the same as the query image and thus the rank is one. We discarded this value as it is a redundant value namely comparing an image with itself. However, we used this information to check whether our system works properly.

3.2 Self Organizing Map(SOM)

Algorithm based Image Retrieval

The SOM algorithm based image clustering and retrieval system, that was used in this study, consisted of the learning and the retrieval components. The learning component generated the feature vectors for the image clustering by applying the gray histogram analysis. The stored image feature vectors were used by the SOM algorithm to cluster the images. The clustering process was repeated for a pre-determined number of times to get the best possible result.

The gray histogram represents the images as a numeric array of 256 dimensions by counting the number of pixels, which belong to each gray scale, starting from 0th dimension representing the

white color and 255th dimension representing the black color.

SOM is a type of neural network algorithm, which transforms the high dimensional data into low dimension, which can be visualized especially in two-dimension space, using the competitive learning algorithm. The clustering occurs as a by product of the dimension reduction process (Kohonen 2001). SOM does not require feedback process during the learning stage. Thus, the learning occurs comparatively faster than other machine learning algorithms. Since the SOM is capable of continuous learning, it is possible to cluster the new images as they are added to the database. Furthermore, the SOM is an unsupervised learning algorithm, which does not require pre-determined positive or negative learning examples. Thus, the algorithm is appropriate for the image retrieval tasks where it is difficult to generate example clusters in advance.

SOM is expected to cause different parts of the network to respond similarly to certain input patterns. The weights of the neurons in SOM are assigned with the small random values. As a learning example is fed to the network, its Euclidean

distance to all weight vectors is computed. The best matching unit is referred to as the neuron with weight vector most similar to the input. The best matching unit and its neighboring neurons' weights are adjusted towards the input vector. The degree of adjustment decreases with time and with distance from the best matching unit (Kohonen 2001). The weight vector $W_v(t)$ for a neuron was updated using [3].

$$W_v(t+1) = W_v(t) + \Theta(v, t) \alpha(t)(D(t) - W_v(t)) \quad [3]$$

where $\alpha(t)$ is a monotonically decreasing learning coefficient and $D(t)$ is the input vector.

The neighborhood function $\Theta(v, t)$ depends on the lattice distance between the best matching unit and neuron v . In this study, we used the Mexican Hat function [4] as the neighborhood function.

$$h(\rho, t) = \exp\left(-\frac{\rho^2}{\sigma^2(t)}\right) \left(1 - \frac{2}{\sigma^2(t)}\rho^2\right) \quad [4]$$

- ρ : Distance to the Winning Neuron
- t : Current Learning Cycle Count

The neighborhood function shrinks with time as well. At the beginning when the neighborhood is broad, the self-organizing takes place on the global scale. When the neighborhood has shrunk to just a couple of neurons the weights are converging to local estimates. This process is repeated for each input vector for a predetermined number of cycles. Eventually, the network associates output

nodes with groups or patterns in the input data set. There will be one single winning neuron during mapping. The winning neuron is the one with weight vector positioned closest to the input vector. The winning neuron can be determined by calculating the Euclidean distance between input vector and weight vector. We used [5] to compute the Euclidean distance (Vesanto and Alhoniemi 2000).

$$D_{ij} = |X^l - W_{ij}| = \sqrt{(x_1 - w_{ij1})^2 + \dots + (x_n - w_{ijn})^2} \quad [5]$$

(i and j : Index Value at the Output Layer)

After all learning cycles were completed, the system stores the weight vector values assigned to the output layer and the corresponding coordinates in the database. In addition, the similar image sets, which were clustered on each neuron in the output layer and the coordinates of the neurons were also stored in the database. This information was used to calculate the winning neuron representing the most appropriate cluster for the query image to be assigned so that the images in the selected cluster could be returned as the image retrieval output.

More specifically, the SOM algorithm based clusters can be conceptualized as shown in the <Figure 8>. If we assume that the figure is a graphical depiction of the output layer of the SOM network and each cell is a neuron, Each cell represents a cluster, which includes the similar images. If the query image belongs to the cluster labeled as 0 then all the database images in that cluster are

considered to be similar by the SOM based image retrieval system. This is similar to how the boolean-based information retrieval systems operate where there is no ranking but a binary value of either relevant or non relevant assigned for each document in the database.

However, the geometrically closer clusters are similar to each other than the farther apart clusters in SOM. Thus, the clusters, which are assigned to the cells labeled as 1, are closer to the cluster in the cell 0 in comparison to the clusters assigned to cell labeled as 2. Thus, we might be able to increase the number of similar images found by combining the images in the cell 0 cluster with the images in the cell 1 clusters. We might also be able to further increase the number of similar images found by combining the previous images with the images in the cell 2 clusters. We consider the cell 9 clusters were too far from the original cell 0 cluster. Thus, we did not expand the retrieved clusters to cell 9 clusters. We experimented this hypothesis and reported the result in the experiment section.

3.3 Data Fusion

The SOM algorithm based image retrieval divided the test data set into two groups of similar and not similar images with respect to the query image. Thus, SOM retrieval returned a set of non-ranked images, which were considered to be relevant. However, the edge detection based image retrieval produces a similarity ranked list as the retrieval output. Thus, we used a method similar to the one employed the two-stage information retrieval systems to combine two different types of retrieval results. The main idea is to assign the edge detection-based rank order separately on the retrieved and not retrieved cluster output. This means that any database image, which was not a member of the retrieved cluster cannot be ranked higher than the database images, which belonged to the retrieved cluster. The <Table 3> shows the retrieval result when the query image was the same as the database image with the identification 3. The format of the table is the same as the <Table 2>. However, there is an additional column referred to as Cluster ID. The value in this last column

9	9	9	9	9	9	9	9	9
9	9	9	9	9	9	9	9	9
9	9	2	2	2	2	2	9	9
9	9	2	1	1	1	2	9	9
9	9	2	1	0	1	2	9	9
9	9	2	1	1	1	2	9	9
9	9	2	2	2	2	2	9	9
9	9	9	9	9	9	9	9	9
9	9	9	9	9	9	9	9	9

<Figure 8> SOM algorithm based Clustering Conceptualization

shows whether the database images belonged to the cell 0, 1, 2, or 9 clusters. The database images in the cell 0 cluster means that they were in the same cluster as the query image in the most strict sense. The <Table 3> currently shows the top ten database images by the edge detection based similarity rank. It should be noted that the database images with the identification 1 to 6 were the images from the same clip art sub-categories and thus the similar images. By excluding the database image with the identification 3, which is identical to the query image, the edge detection-based image retrieval system found one more similar image with respect to the query image as the 4th ranked image. However, the cluster id of this image with the identification 1 is 0 and thus the SOM-based image retrieval considered this image to be similar to the query image.

By applying the two-stage retrieval type data fusion method to combine the edge detection based ranking and the SOM-based clustering output, we get the following top ten database images as shown in the <Table 4>. Now, all the database images,

which belonged to the cell 0 cluster ranked at the top part of the list according to the edge detection-based image retrieval ranking. Then, the database images, which did not belong to the cell 0 clusters were placed below according to the edge detection-based image retrieval ranking. The <Table 4> shows much more improved retrieval result. After the data fusion process, two additional similar images appeared in the top ten retrieved image set. The image with the identification 5, which was not in the top ten list based on the Sobel edge detection retrieval shown in the <Table 3>, is now third ranked image excluding the query image itself from the ranked list. The image with the identification 4 is fourth ranked image. The rank of the image with the identification 1 jumped up three ranks after the data fusion process. We compared the data fusion result by considering 1) cell 0 cluster as the SOM retrieved cluster; 2) cell 0 and 1 clusters as the SOM retrieved clusters; and cell 0, 1, and 2 clusters as the SOM retrieved clusters.

<Table 3> Edge Detection Based Similarity Ranking & SOM-based Clustering Result

ID	0 degree	45 degree	90 degree	135 degree	180 degree	225 degree	270 degree	315 degree	minimum (bet. 0, 315)	rank	Cluster ID
3	0.00	17.78	17.83	17.78	18.44	17.78	17.83	17.78	0.00	1	0
16	15.87	16.43	13.64	5.48	16.43	14.14	17.09	16.12	5.48	2	9
89	20.95	20.81	6.24	19.52	18.14	18.68	21.19	21.28	6.24	3	9
96	6.24	19.42	20.12	19.57	22.38	21.93	22.78	19.77	6.24	3	9
1	6.40	12.37	14.93	16.34	15.78	14.39	15.46	15.65	6.40	5	0
73	13.23	16.52	6.56	15.39	13.75	15.65	14.18	16.16	6.56	6	9
83	23.17	21.47	21.19	21.42	6.56	22.65	20.27	21.52	6.56	6	9
17	13.00	16.16	6.86	16.82	15.97	12.69	17.00	14.93	6.86	8	9
28	19.18	16.49	6.93	15.30	18.55	18.33	18.97	20.54	6.93	9	9
32	15.56	15.94	13.64	15.62	14.14	15.36	7.21	11.75	7.21	10	9

〈Table 4〉 Data Fusion Result

ID	0 degree	45 degree	90 degree	135 degree	180 degree	225 degree	270 degree	315 degree	minimum (bet. 0, 315)	rank	Cluster ID
3	0.00	17.78	17.83	17.78	18.44	17.78	17.83	17.78	0.00	1	0
1	6.40	12.37	14.93	16.34	15.78	14.39	15.46	15.65	6.40	5	0
77	12.53	12.69	14.46	12.77	13.75	10.54	11.53	11.53	10.54	58	0
5	14.56	13.86	10.58	12.41	14.35	14.63	13.78	14.42	10.58	59	0
4	14.59	14.53	14.66	12.85	12.61	14.39	15.39	14.73	12.61	74	0
22	21.66	17.69	24.27	19.10	22.29	16.46	23.77	23.90	16.46	94	0
86	20.07	25.55	22.87	23.09	24.10	23.77	24.27	25.63	20.07	102	0
76	14.66	14.66	12.85	14.53	14.53	15.52	14.73	14.59	12.85	76	1
108	18.11	15.94	13.49	18.55	17.26	18.55	15.23	13.93	13.49	80	1
103	24.23	24.43	20.71	22.91	23.04	26.89	21.84	24.23	20.71	103	1

4. Experiment Results

We first measured the retrieval effectiveness of the Sobel edge detection and SOM based image retrieval separately. For the SOM based image retrieval, we measured the recall and precision figures for three cases. The first case was considering only the images in the cell 0 cluster as the retrieved images. The second case was including the images from both the cell 0 cluster and immediately surrounding cell 1 clusters as the retrieved images. The final case was treating the images in the cell 0 cluster, cell 1 clusters, and cell 2 clusters as the retrieved images by the system.

Then, we combined the ranked list from the Sobel edge detection based retrieval and the three types of retrieved set of images from the SOM-based system to generate a new ranked lists. The data fusion resulted three ranked lists were evaluated to determine whether the data fusion had achieved better than the single component system. We also wanted to find out the best combination strategy out of three possible data fusion methods.

The test data set was 114 clip art images, which were organized into 19 sub-categories of six images. Every image in the test data set was used as the query images. For each query image, the other five images belonging to the same sub-category were judged to be similar and the images belonging to the other 18 sub-categories were judged to be not similar. The query image always have one exact duplicate in the test data set as the query images were also from the test data set. The retrieval of the query image itself was ignored in calculating the effectiveness figures to ensure a fair evaluation.

We used 11-Point Interpolated Average Precision to generate a single evaluation figure for the Sobel edge detection based image retrieval system as well as the data fusion results. To compute the 11-Point Interpolated Average Precision, the interpolated precision is measured at 11 recall levels, which begins at 0.0 and ends at 1.0 with 0.1 increment. The interpolated precision at a particular recall level R means the highest precision found for any recall level higher than R. The measured

interpolated precisions at 11 recall levels are averaged over the set of queries to generate the 11-Point Interpolated Average Precision(Manning et al. 2008).

The <Table 5> shows the statistical analysis result of the 11-Point Interpolated Average Precision over 114 query images. The average of 11-Point Interpolated Average Precisions was 0.18.

Since the SOM-based image retrieval system is a boolean retrieval system. Thus, it is possible to generate a single precision and recall value for the retrieved set of images. The precision was computed by dividing the number of similar images

retrieved by the number of all retrieved images. The recall was computed by dividing the number of similar images retrieved by the number of all similar images. The <Table 6> shows the statistical analysis of the retrieval effectiveness over 114 query images by considering only the images in the cell 0 cluster as the retrieved images. The <Table 7> shows the effectiveness figure for the case where the images in the cell 0 and cell 1 clusters were treated as the retrieved images. The <Table 8> the effectiveness figure for the case where the images in the cell 0, cell 1, and cell 2 clusters were regarded as the retrieved images.

<Table 5> Sobel Edge Detection based Image Retrieval: 11-Point Interpolated Average Precision over 114 Query Images

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
11-Point Interpolated Average Precision	114	.04	.60	.1825	.13758	.019
Valid N (listwise)	114					

<Table 6> SOM based Image Retrieval: Average Precision and Recall over 114 Query Images using cell 0 cluster as the retrieved cluster

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Precision	114	.00	1.00	.2687	.35681	.127
Recall	114	.00	.60	.1474	.19742	.039
Valid N (listwise)	114					

<Table 7> SOM based Image Retrieval: Average Precision and Recall over 114 Query Images using cell 0 and 1 clusters as the retrieved clusters

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Precision	114	.00	.45	.1441	.11233	.013
Recall	114	.00	1.00	.4860	.34994	.122
Valid N (listwise)	114					

<Table 8> SOM based Image Retrieval: Average Precision and Recall over 114 Query Images using cell 0, 1, and 2 clusters as the retrieved clusters

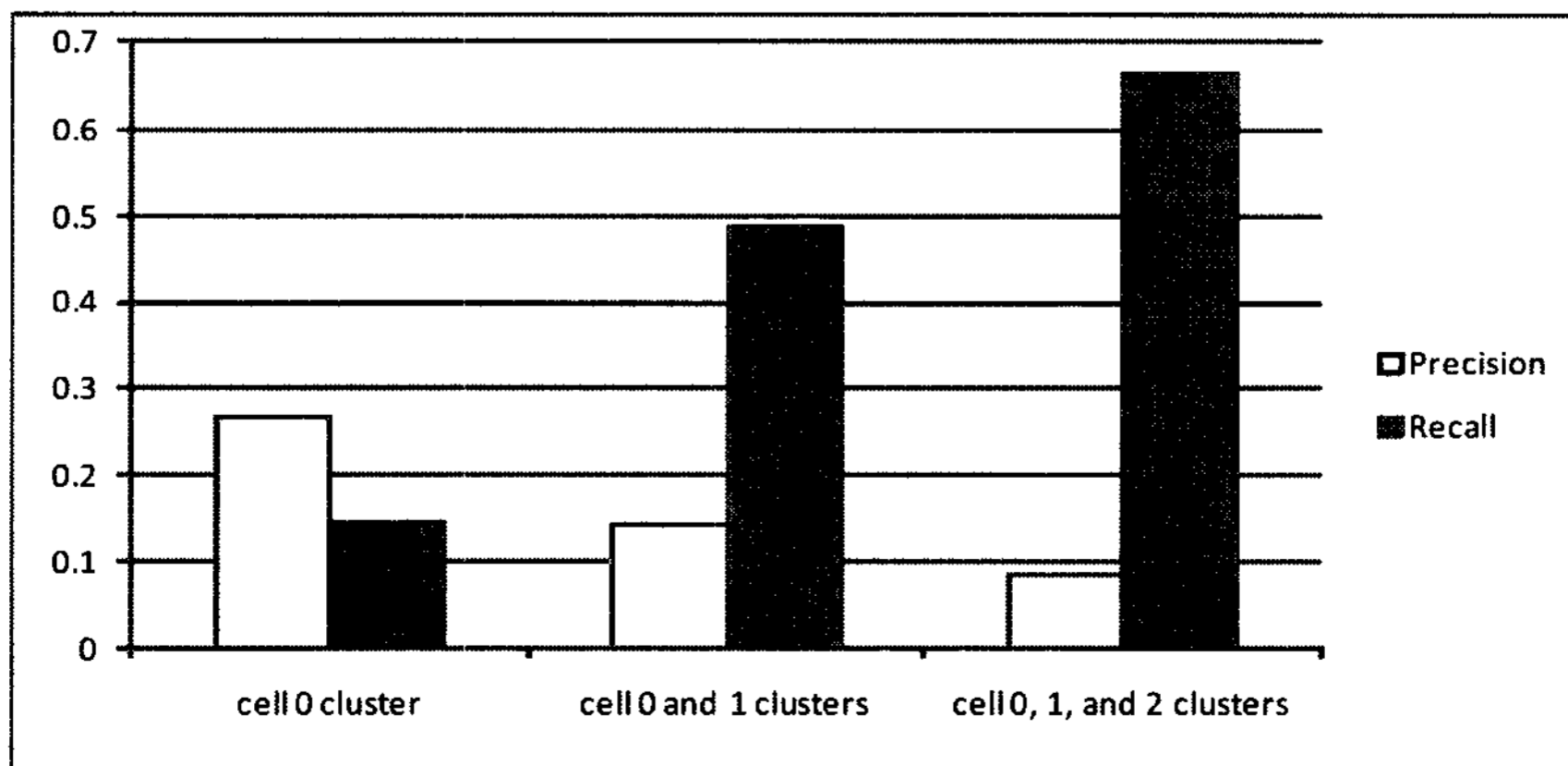
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Precision	114	.00	.19	.0843	.04395	.002
Recall	114	.00	1.00	.6649	.31200	.097
Valid N (listwise)	114					

By comparing three cases shown in the Tables 6, 7, and 8, we were able to determine that the image retrieval also behaves similar to the typical information retrieval systems in terms of how the precision value changed as the recall value changed. The precision value decreased as the recall value increased. The <Figure 9> shows how the precision and recall changes as the criteria for the retrieved images changes.

Finally, we experimented the data fusion of the Sobel edge detection based ranked list with three cases of SOM-based clustering. Two stage retrieval method of data fusion was used. The <Table 9> shows the statistical analysis of the retrieval effec-

tiveness of various data fusion schemes over 114 query images. The figures on the 'No Data Fusion' line represent the baseline performance of the Sobel edge detection image retrieval system without any data fusion with the SOM-based results. The 11-Point Interpolated Average Precision based numbers reported for this line is the same as the one reported in the <Table 5>.

The values on the row labeled as 'Data Fusion Cell 0' represent the data fusion of the ranked list from the Sobel edge detection based image retrieval system and the images belonging to the cell 0 cluster retrieved from the SOM-based image retrieval system.



<Figure 9> SOM-based Image Retrieval Effectiveness Change

<Table 9> Data Fusion based Image Retrieval Effectiveness: 11-Point Interpolated Average Precision over 114 Query Images

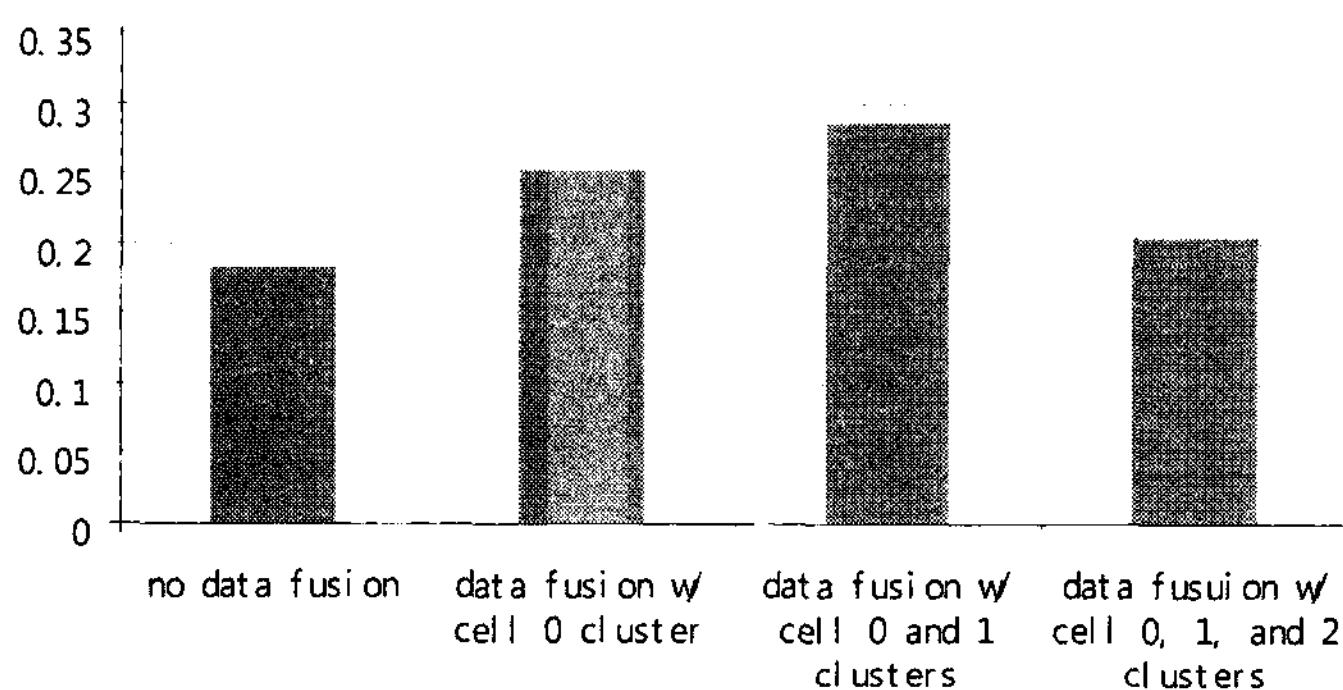
	N	Minimum	Maximum	Mean	Std. Deviation	Variance
No Data Fusion	114	.04	.60	.1825	.13758	.019
Data Fusion Cell 0	114	.04	.73	.2534	.19945	.040
Data Fusion Cell 0+1	114	.04	.87	.2857	.21411	.046
Data Fusion Cell 0+1+2	114	.04	.73	.2042	.16051	.026
Valid N (listwise)	114					

The numbers on the ‘Data Fusion Cell 0+1’ line represent the data fusion of the ranked list from the Sobel edge detection based image retrieval system and the images belonging to the cell 0 and 1 clusters retrieved from the SOM-based image retrieval system. Finally, The numbers on the ‘Data Fusion Cell 0+1+2’ line represent the data fusion of the ranked list from the Sobel edge detection based image retrieval system and the images belonging to the cell 0, 1, and 2 clusters retrieved from the SOM-based image retrieval system.

The <Figure 10> visually compares the average(i.e. mean) values for the 11-Point Interpolated Average Precision figures. It clearly shows the

advantage of the data fusion over the component system. In addition, the <Figure 10> reveals that the best data fusion strategy is to combine the Sobel ranked list with the images in the cell 0 and 1 clusters based on the SOM-based image retrieval system.

By conducting the paired samples T-test result between the worst performing ‘no data fusion’ and the best performing ‘data fusion cell 0+1’, we found that there is significant difference between two cases in terms of the 11-Point Interpolated Average Precision value. The statistical analysis result is shown in the <Table 10>.



<Figure 10> Changes of the Average 11-Point Interpolated Average Precision over 114 Query Images by different Data Fusion Schemes

<Table 10> Paired Sample T-test between the baseline no data fusion and the data fusion between the Sobel ranked list and the images in the cell 0 and 1 clusters

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
	-.103	.20473	.01917	-.14121	-.06524	-5.383	113	.000

5. Conclusion & Future Works

In comparison to the retrieval effectiveness achieved by each component systems such as the Sobel edge detection and SOM based image retrieval systems, the two-stage data fusion method achieved the comparatively better 11-point interpolated average precision. We believe that this result is encouraging as it showed that the image retrieval can also be benefited from the evidence combination based on multiple retrieval schemes.

In addition, our findings suggests one way to improve the image retrieval efficiency through the use of the two-stage information retrieval model.

As we have done during the data fusion process, the SOM-based image retrieval system, which does fast boolean retrieval, can act as the first cut filter. Then, the Sobel edge detection based image retrieval system can assign ranks only to the retrieved images passed from the first cut filter. This data fusion should reduce the overall processing time and thus improve the system efficiency.

We also found that the data fusion using the images from the cell 0 and 1 clusters and the Sobel ranking produced the best retrieval effectiveness. This is also an interesting finding. However, this and other findings will be further verified with a larger test data set.

References

- Amato, A. and Lecce, V.D. 2003. "Edge Detection Techniques in Image Retrieval: The Semantic Meaning of Edge." *Proceedings of 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communication*, 143-148.
- Aslam, J.A. and Montague, M. 2001. "Models for metasearch." *Proceedings of the 24th ACM SIGIR Conference on Research and Development in Information Retrieval*, 276-284.
- Aslandogan, Y.A. and Yu, C.T. 2000. "Multiple evidence combination in image retrieval:

- Diogenes searches for people on the Web.” *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and development in Information Retrieval*, 88-95.
- Belkin, N.J., Kantor, P.B., Cool, C., & Quatrain, R. 1994. “Combining evidence for information retrieval.” *Proceedings of the Second Text Retrieval Conference*, 35-44.
- Chesnut, Casey 2004. “Self Organizing Map AI for Pictures.” Generation 5. [cited 2007.8.1]. <http://generation5.org/content/2004/aiso_mpic.asp>.
- Fox, E.A. and Shaw, J.A. 1994. “Combination of multiple searches.” *Proceedings of the Second Text Retrieval Conference*, 243-252.
- Green, B. 2002. “Edge Detection Tutorial.” [cited 2008.6.7]. <<http://www.pages.drexel.edu/~weg22/edge.html>>.
- Gonzalez, R. and Woods, R. 2002. *Digital Image Processing*. (2nd ed.) Upper Saddle River, NJ. Prentice Hall.
- Hsu, D.F. and Taksa, I 2005. “Comparing Rank and Score Combination Methods for Data Fusion in Information Retrieval.” *Information Retrieval*, 8: 449-480.
- Kohonen, K. 2001. *Self-Organizing Maps*. (3rd ed.) Berlin. Springer.
- Lee, J.H. 1997. “Analyses of multiple evidence combination.” *Proceedings of the 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 267-276.
- Manning, C.D., Raghavan, P., and Schütze, H. 2008. *Introduction to Information Retrieval*. New York, NY. Cambridge University Press.
- Mase, H., Matsubayashi, T., Ogawa, Y., Iwayama, M., & Oshio, T. 2005. “Proposal of two-stage patent retrieval method considering the claim structure.” *ACM Transactions on Asian Language Information Processing*, 4(2): 190-206.
- Nova Development Corp. 2004. *Nova Art Explosion 800,000 Clip Art*. Calabasas, CA.
- Paik, W., Lee, J., Shin, M., Ahn, E. Ham, E. & Shin, M. 2007. “An Evaluative Study on the Content-based Trademark Image Retrieval System Based on Self Organizing Map (SOM) Algorithm.” *Journal of Korea Society for Information Management*, 24(3): 321-341.
- Rui, Y., Huang, T.S., and Mehrotra, S. 1998. “Relevance feedback techniques in interactive content-based image retrieval.” *Storage and Retrieval for Image and Video Databases*, 25-36.
- Rui, Y. and Huang, T.S. 1999. “A novel relevance feedback technique in image retrieval.” *Proceedings of the Seventh ACM international conference on Multimedia*, 67-70.
- Saracevic, T. and Kantor, P.B. 1988. “A study of information seeking and retrieving. III Searchers, searches, overlap.” *Journal of the American Society for Information Science*, 39: 197-216.
- Sjoberg, M., Muurinen, H., Laaksonen, J., &

- Koskela, M. 2006. "PicSOM experiments in TRECVID 2006." In *Online Proceedings of the TREC Video Retrieval Evaluation 2006*. [cited 2007.8.1]. <<http://www-nlpir.nist.gov/projects/tvpubs/tv6.papers/picsom.pdf>>.
- Vesanto, J. and Alhoniemi, E. 2000. "Clustering of The Self-Organizing Map." *IEEE Transaction on Neural Networks*, 11(3): 586-600.
- Vogt, C.C. and Cottrell, G.W. 1999. "Fusion via a linear combination of scores." *Information Retrieval*, 1(3): 151-173.
- Yavlinsky, A., Pickering, M.J., Heesch, D., & Ruger, S. 2004. "A comparative study of evidence combination strategies." *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 3: 1040-1043.