

Using Latent Semantic Index for Content-Based Image Retrieval

Andy¹, Bernardus Ari Kuncoro²

^{1,2}Master of Information Technology, Bina Nusantara University, Jl. Kebon Jeruk Raya No. 27, DKI Jakarta, Indonesia

Abstract: *In this paper, the latent semantic indexing (LSI) based method was used to the various image feature extracted matrix in order to perform content-based image retrieval. The feature extraction techniques include color histogram, color auto-gram, color moment, gray-scale, and wavelet moment. The implementation of LSI here is to achieve an improved image retrieval performance, because it reduces the size of input image matrix by determining parameter of k. The objective of this paper is to know the performance of LSI-based CBIR using precision and recall parameters for each feature extractions.*

Keywords: image feature extractions, latent semantic indexing, content-based image retrieval

1. Introduction

The number of digital multimedia information available is increasing. Large amounts of text, images, speech, and video are converted to digital form. Also, the birth of social media platforms contributes the explosion of digital multimedia information. With huge data available, getting the relevant data that queried by users is challenging, whereas more and more people have high expectation retrieving the desired information precisely and relevantly. Those facts make the multimedia retrieval system become important topic research in this digital era [1].

One of the forms of multimedia data other than text, audio, and video is image. Having huge set of images data contributes challenges in building retrieval system in term of speed and relevancy. In recent years, many researchers have proposed an efficient method of image searching and retrieval. However, the systems present today are mostly text-based or manually annotated by text-based keywords. Users query by set of keywords, instead of looking into the contents of the image, in which this system matches the query to the keywords present in the database[2],[3]. The disadvantages of this system are: first, the manual annotation is a labor process and exhausting, second, the rich features of an image cannot be described by using keywords only.

The disadvantages of text-based image retrieval motivate researchers to build a method called Content-Based Image Retrieval (CBIR). CBIR is a technology that in principle helps organize digital image archives according to their visual content. This system distinguishes the different regions present in an image based on their similarity in color, pattern, texture, shape, etc. and decides the similarity between two images by reckoning the closeness of these different regions. The CBIR approach is much closer to how we humans distinguish images[2].

According to [3], CBIR can be categorized based on the type of features used for retrieval, which could be either low-level or high-level features. Low-level features include color, texture, shape, and spatial relations, while high-level features include the semantic features, which are difficult due to the fact that the visual image feature descriptors extracted from an image cannot (as yet) be automatically translated reliably

into high-level semantics[3].

In this paper, Latent Semantic Index, a method to reduce the complexity of calculation of the matrix was proposed in CBIR. The aim of using LSI here is to make CBIR more efficient and fast, as the features extracted are multidimensional. The basic idea of LSI is that, LSI searches for something that is closer to represent the underlying semantics of an image. The searching is done by using matrix computation, in particular Singular Value Decomposition (SVD)[4]. The objective of this research is to know the performance of LSI-CBIR system towards different low-level feature extractions.

2. Dataset

The dataset used in this work can be downloaded from Wang Database (<http://wang.ist.psu.edu/docs/related/>). The actual database contains 1000 images, but in this research only first 500 images were used. Among 500 images, there are 5 categories, in which each category contains 100 images. The size of images is various, ranged from 384x256 for landscape layout and 256x384 for portrait layout. The dataset has filename that has number from 0 to 499, which consist of that 5 categories in numerical sequence in each categories.

3. Methodology

In this research, the process to search image consist of three big steps: feature extraction, creating the database using LSI, and image retrieval of query images. To produce more accurate retrieval result, this research will use various features and then compare each feature to know which feature produce best result for specific category using Precision and Recall parameter.

3.1 Feature Extractions

The feature extraction techniques used in this work[5] are color histogram [6], [7], color auto correlogram[8], color moment, mean-amplitude and energy of Gabor wavelet [9], wavelet moment [10], and grayscale. They were applied to all images. The matrix of each image was reconstructed to become 1 dimensional matrix by transposing each rows of

original matrix of feature extraction. The size of the 1 dimensional matrix is $m \times I$, where m equals to the size of row \times column from original matrix. As a result, the size of final matrix is $m \times 500$.

3.2 Creating Database using LSI

The basic idea of LSI is to transform the attribute-image vectors to some low-dimensional space, which approximate the original vectors as good as possible. The new dimensions are no longer bond to an individual attribute, e.g. pixels, but should denote concepts encompassing several attributes. Let the symbol A denotes the $m \times n$ attribute-image matrix, where m is the number of element of feature-extracted matrix of each image, and n is the number of images in database. The LSI procedure involves a Singular value decomposition (SVD) of the attribute-image matrix A . The aim of SVD is to compute decomposition

$$A = USV^T, \quad (1)$$

where $S \in R^{m \times n}$ is a diagonal matrix with nonnegative diagonal elements called the singular values, $U \in R^{m \times m}$ and $V \in R^{n \times n}$ are orthogonal matrices. The columns of matrices U and V are called the left singular vectors and the right singular vectors respectively. The decomposition can be computed so that the singular values are sorted by decreasing order.

For large image database, the full SVD decomposition (2) is memory and time consuming operation. Hence, the reduction of dimension is necessary. It was done by sorting elements by decreasing singular values (U and V), and eliminates the smallest elements in each matrix, that leads to a reconstructed with smaller (k, k) -diagonal matrix S_k . In this work, simply the parameter of k was chosen. Afterwards, the new image vector coordinates in this reduced k -dimensional space was revealed. Rows of V holds eigenvector values. As a result, each individual image can be represented by k -dimension vector. The process of LSI can be illustrated in Figure 1.

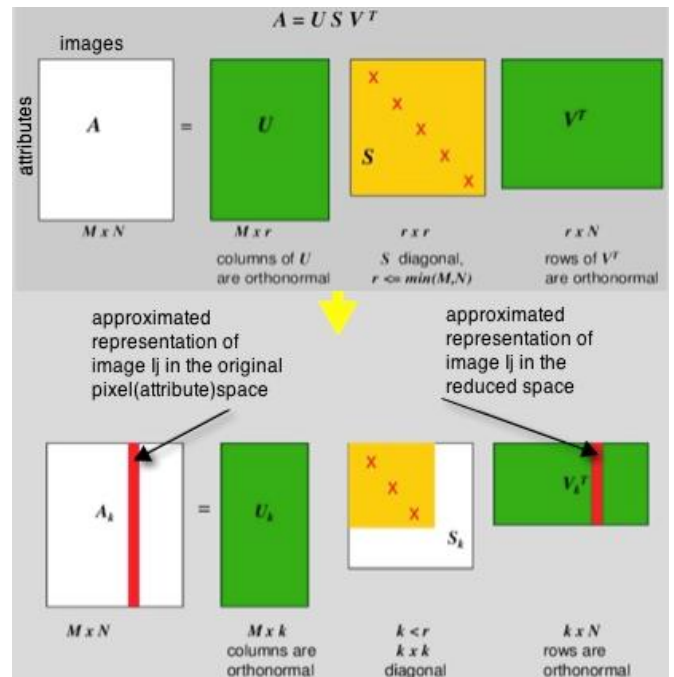


Figure 1: Latent Semantic Indexing Process in Image Data

3.3 Image Retrieval

Each query (q) images produced features vector with LSI approach. The coordinate query vector (q_c) is represented by the following formula.

$$q_c = q^T U_k S_k^T \quad (2)$$

Where k represents the rank reduction of dimension, in this research, $k = 10$. The similarity was calculated with the images database loaded before retrieval using the cosine similarity matrix formula. Each query was retrieved up to 25 images.

$$sim(q, I) = \frac{q^T I}{\|q\| \|I\|} \quad (3)$$

4. Results and Discussion

In this work, three images from each category were used as query images. Then the average of the precision from each images in each feature were calculated. Among 7 features used in the experiment, color moment has 6-dimensional features which is lower than the designated k value (10), thus the k in color moment is set to 6, which is the maximum dimensional features value of color moment.

Table 1: Class 1 CBIR Precision Result

| Query Image | Color Histogram | Color Auto correlogram | Color Moment | Mean Amplitude Wavelet | Energy Gabor | Wavelet Moment | GrayScale |
|-------------|-----------------|------------------------|--------------|------------------------|--------------|----------------|-----------|
| 0.jpg | 76% | 20% | 60% | 16.00% | 52.00% | 68.00% | 36.00% |
| 11.jpg | 80% | 16% | 36% | 56.00% | 40.00% | 36.00% | 12.00% |
| 22.jpg | 72% | 32% | 72% | 84.00% | 88.00% | 52.00% | 28.00% |
| Average | 76% | 23% | 56% | 52% | 60% | 52% | 25% |

Table 2: Class 2 CBIR Precision Result

| Query Image | Color Histogram | Color Auto correlogram | Color Moment | Mean Amplitude Wavelet | Energy Gabor | Wavelet Moment | GrayScale |
|-------------|-----------------|------------------------|--------------|------------------------|--------------|----------------|-----------|
| 102.jpg | 76% | 12% | 100% | 60.00% | 48.00% | 60.00% | 44.00% |
| 122.jpg | 48% | 16% | 72% | 52.00% | 52.00% | 32.00% | 16.00% |
| 147.jpg | 44% | 44% | 48% | 56.00% | 56.00% | 32.00% | 40.00% |
| Average | 56% | 24% | 73% | 56% | 52% | 41% | 33% |

Table 3: Class 3 CBIR Precision Result

| Query Image | Color Histogram | Color Auto correlogram | Color Moment | Mean Amplitude Wavelet | Energy Gabor | Wavelet Moment | GrayScale |
|-------------|-----------------|------------------------|--------------|------------------------|--------------|----------------|-----------|
| 202.jpg | 36% | 32% | 72% | 48.00% | 60.00% | 64% | 36% |
| 232.jpg | 20% | 2% | 60% | 56.00% | 56.00% | 28% | 16% |
| 242.jpg | 68% | 12% | 48% | 48.00% | 52.00% | 60% | 40% |
| Average | 41% | 15% | 60% | 51% | 56% | 51% | 31% |

Table 4: Class 4 CBIR Precision Result

| Query Image | Color Histogram | Color Auto correlogram | Color Moment | Mean Amplitude Wavelet | Energy Gabor | Wavelet Moment | GrayScale |
|-------------|-----------------|------------------------|--------------|------------------------|--------------|----------------|-----------|
| 302.jpg | 76% | 2% | 88% | 56.00% | 52.00% | 76% | 60% |
| 332.jpg | 96% | 36% | 68% | 92.00% | 80.00% | 28% | 72% |
| 342.jpg | 16% | 12% | 36% | 84.00% | 76.00% | 36% | 28% |
| Average | 63% | 17% | 64% | 77% | 69% | 47% | 53% |

Table 5: Class 5 CBIR Precision Result

| Query Image | Color Histogram | Color Auto correlogram | Color Moment | Mean Amplitude Wavelet | Energy Gabor | Wavelet Moment | GrayScale |
|-------------|-----------------|------------------------|--------------|------------------------|--------------|----------------|-----------|
| 402.jpg | 100% | 24% | 100% | 52.00% | 64.00% | 92% | 100% |
| 422.jpg | 92% | 44% | 84% | 72.00% | 68.00% | 60% | 100% |
| 442.jpg | 100% | 20% | 100% | 92.00% | 96.00% | 100% | 100% |
| Average | 97% | 29% | 95% | 72% | 76% | 84% | 100% |

Table 6: Average Features Precision Result

| | Color Histogram | Color Auto correlogram | Color Moment | Mean Amplitude Wavelet | Energy Gabor | Wavelet Moment | GrayScale |
|---------------|-----------------|------------------------|--------------|------------------------|--------------|----------------|-----------|
| Total Average | 67% | 22% | 70% | 62% | 63% | 55% | 49% |

Table 7: Average Classes Precision Result

| | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 |
|---------------|---------|---------|---------|---------|---------|
| Total Average | 49% | 48% | 44% | 56% | 79% |

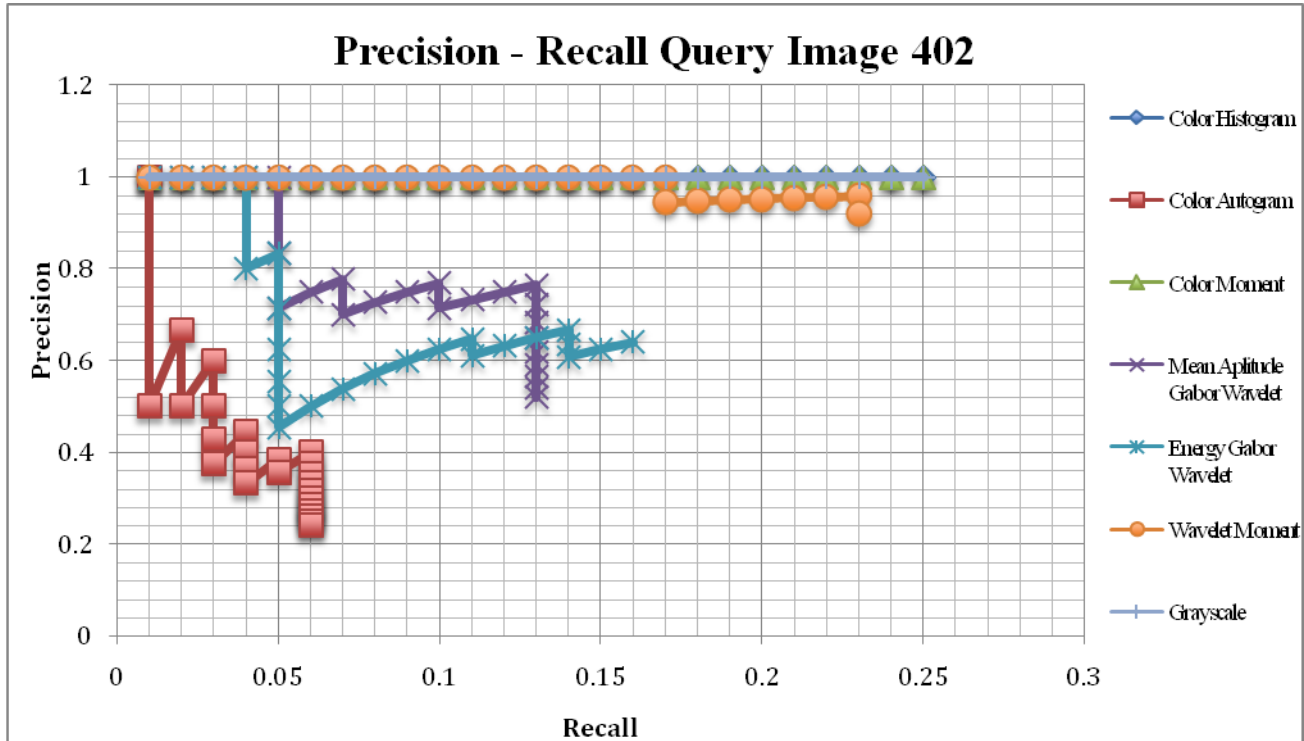


Figure 1: Precision-Recall of Query Image 402

Based on Table 1, 2, 3, 4, and 5, there is no particular feature resulting high consistence precision value among all classes. The main reason of this result is that the color and texture complexity of the images in each classes are different. Based on Table 6, color moment has 70% of average accuracy in

retrieval result or the best among others. Interestingly, during execution of the program, the computation time of color moment was the shortest, most probably due to the lowest number of dimensional features (6 – dimensions).

Table 7 shows that class 5 has the highest average precision at 79%, followed by class 4, class 1, class 2, and class 3 with 56%, 49%, 48%, and 44% average precision respectively. The logical reason of this result is that the images in class 5 are dinosaur animals, which are very similar between images, while class 4 images consist of bus images, which have much similar texture. Meanwhile class 1 images consist of African tribe people, class 2 images consist of beach environment, and class 3 images consist of building images which have various type of color or texture.

Figure 1 illustrates the precision-recall (P/R) graph for image 402 (dinosaur image) as a query. According to that figure, the feature extraction techniques of grayscale, color moment, and wavelet moment give good precision/recall result. It was shown that the points of P/R were located on the upper right of the graph. While mean amplitude, energy wavelet, and color autogram did not give good P/R result for this query.

5. Conclusion

This research shows that LSI can be considerably used to produce good CBIR system in terms of precision and recall, even only one feature extraction technique was used. Although no single feature technique shows high-consistence precision result, the average accuracy of some features are still considered to be good enough, especially the color moment feature extraction technique that performed well with 70% precision with only 6 dimensional features vector. The non-recommended feature techniques to be paired with LSI are color autocorrelogram and grayscale, because the average precision result were below 50%.

For further research, it is recommended to combine the various features to increase the precision result of the CBIR. Another experiment needed to increase LSI performance in CBIR is to change the value of k .

References

- [1] H. M. Blanken, A. P. de Vries, H. E. Blok, and L. Feng, *Multimedia Retrieval*. Berlin Heidelberg: Springer, 2007.
- [2] V. S. V. S. Murthy, E. Vamsidhar, J. N. V. R. S. Kumar, and P. S. Rao, "Content Based Image Retrieval using Hierarchical and K-Means Clustering Techniques," *Int. J. Eng. Sci. Technol.*, vol. 2, no. 3, pp. 209–212, 2010.
- [3] H. H. Wang, D. Mohamad, and N. A. Ismail, "Image Retrieval : Techniques , Challenge , and Trend," *World Acad. Sci. Eng. Technol.*, vol. 36, pp. 718–720, 2009.
- [4] P. Praks, J. Dvorsk'y, and V. Snášel, "Latent semantic indexing for image retrieval systems," in *SIAM Linear Algebra Proceedings, Philadelphia, USA. International Linear Algebra Society (ILAS)*, 2003.
- [5] Y. D. Chun, N. C. Kim, and I. H. Jang, "Content-Based Image Retrieval Using Multiresolution Color and Texture Features," *IEEE Trans. Multimed.*, vol. 10, no. 6, pp. 1073–1084, 2008.
- [6] S. M. R. Devi and C. Bhagvati, "Connected Component in feature space to capture High level semantics in CBIR," *Proc. Fourth Annu. ACM Bangalore Conf. - Comput. '11*, pp. 1–6, 2011.

- [7] Y.-F. Huang and H.-Y. Lu, "Automatic Image Annotation Using Multi-object Identification," *2010 Fourth Pacific-Rim Symp. Image Video Technol.*, pp. 386–392, 2010.
- [8] X. Chen, X. Yang, R. Zhang, A. Liu, and S. Zheng, "Edge Region Color Autocorrelogram: A New Low-level Feature Applied in CBIR," *Broadband Multimed. Syst. Broadcast. (BMSB), 2010 IEEE Int. Symp.*, pp. 1–4, 2010.
- [9] S. M. Singh and K. Hemachandran, "Content-Based Image Retrieval using Color Moment and Gabor Texture Feature," *Intentional J. Comput. Sci. Issues*, vol. 9, no. 5, pp. 299–309, 2012.
- [10] X. Wang, P. Niu, and M. Lu, "A Robust Digital Audio Watermarking Scheme Using Wavelet Moment Invariance," *J. Syst. Softw.*, vol. 84, no. 8, pp. 1408–1421, 2011.

Author Profile



Andy was born in Medan, 1991. His major is Computer Science and interest in game application and Artificial Intelligent major. Currently study in Bina Nusantara University for master degree.



Bernardus Ari Kuncoro received B. Eng. (Sarjana Teknik) of Electrical Engineering from Bandung Institute of Technology in 2008. Currently he is a master student of Bina Nusantara University Graduate Program, Jakarta in Information Technology. His research interest includes Artificial Intelligence, Image Processing, and Data Mining.