Cluster Oriented Image Retrieval System with Context Based Color Feature Subspace Selection

Ali Ridho Barakbah[1,2] and Yasushi Kiyoki[3]
[1] *Electronic Engineering Polytechnic Institute of Technology, Indonesia*
[2] *Graduate School of Media and Governance, Keio University, Japan*
[3] *Faculty of Environmental Information, Keio University, Japan*
Email: ridho@mdbl.sfc.keio.ac.jp, kiyoki@sfc.keio.ac.jp

Abstract

This paper presents a cluster oriented image retrieval system with context recognition mechanism for selection subspaces of color features. Our idea to implement a context in the image retrieval system is how to recognize the most important features in the image search by connecting the user impression to the query. We apply a context recognition with Mathematical Model of Meaning (MMM) and then make a projection to the color features with a color impression metric. After a user gives a context, the MMM retrieves the highest correlated words to the context. These representative words are projected to the color impression metric to obtain the most significant colors for subspace feature selection. After applying subspace selection, the system then clusters the image database using Pillar-Kmeans algorithm. The centroids of clustering results are used for calculating the similarity measurements to the image query. We perform our proposed system for experimental purpose with the Ukiyo-e image datasets from Tokyo Metropolitan Library for representing the Japanese cultural image collections.

Keywords: Image retrieval, Mathematical Model of Meaning, clustering, Pillar algorithm.

1. Introduction

The World Wide Web has become a significant source of information, including image data. Everyday abundant information is transformed and collected into huge databases which makes difficult in processing and analyzing data without the use of automatic approaches and techniques. Related to image data, many researchers and developers developed an efficient image searching, browsing, and retrieval systems in order to provide better ways and approaches for such kinds of activities.

The image retrieval systems based on the contents are attracting and challenging in research areas of image searching. Many content-based image retrieval (CBIR) systems have been proposed and widely applied to both commercial purposes and research systems. The system analyzes the content of an image by extracting primitive features such as color, shape, texture, etc. Most approaches have been introduced to explore the content of an image and identify the primary and dominant features inside the image.

QBIC [3] introduced an image retrieval system based on color information inside an image. VisualSeek [7] represented a system by diagramming spatial arrangements based on representation of color regions. NETRA [8] developed a CBIR system by extracting color and texture features. Virage [6] utilized color, texture, and shape features for the image retrieval engine. CoIRS [10] also introduced a cluster oriented image retrieval system based on color, shape, and texture features. Veltkamp and Tanase [9] and Liu et al [11] presented a survey to many image retrieval systems using diverse features. Barakbah and Kiyoki introduced an image retrieval system by combining color, shape and structure features [12].

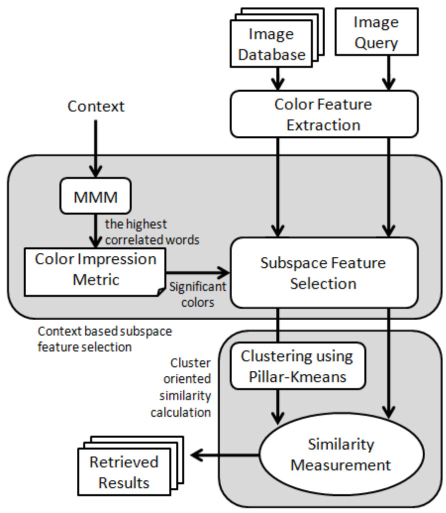

Figure 1. System architecture of our proposed image retrieval system

In this paper, a cluster based image retrieval system with context recognition mechanism for selection subspaces of color features. Our idea to implement a context in the image retrieval system is how to recognize the most important features in the

Ali Ridho Barakbah, Yasushi Kiyoki, "Cluster Oriented Image Retrieval System with Context Based Color Feature Subspace Selection", Industrial Electronics Seminar (IES) 2009, October 21, 2009, Surabaya, Indonesia.

image query by connecting the user impression to the system. We apply a context recognition with Mathematical Model of Meaning (MMM) and then make a projection to the color features with a color impression metric for subspace feature selection. The system then applies clustering for the image database and the image query using K-means clustering after optimized by Pillar algorithm. Figure 1 shows the system architecture of the proposed system.

We organize this paper as follows. In Section 2, the context recognition mechanism using MMM is described. Section 3 discusses the feature extraction and subspace selection. The cluster oriented similarity measurement in our image retrieval system using Pillar-Kmeans algorithm is described in Section 4. Section 5 describes the experimental results using the Ukiyo-e image datasets and discusses the performance analysis, and then followed by concluding remarks in Section 6.

2. Context Recognition Mechanism

Our idea to implement a context in the image retrieval system is how to recognize the most important features in the image query by connecting the user impression to the system. To recognize a context given by a user, we apply Mathematical Model of Meaning (MMM). When a user gives a context, the MMM retrieves the highest correlated words to the context. These representative words are projected to the color impression metric to obtain the most significant colors for subspace feature selection. In this section, the outline of the Mathematical Model of Meaning (MMM) is briefly reviewed. This model has been presented in [2], [4] and [5] in detail.

2.1. An overview of the Mathematical Model of Meaning

In the Mathematical Model of Meaning, an orthogonal semantic space is created for semantic associative search. Retrieval candidates and queries are mapped onto the semantic space. The semantic associative search is performed by calculating the correlation of the retrieval candidates and the queries on the semantic space in the following steps:

(1) A context represented as a set of impression words is given by a user, as shown in Figure 2(a).
(2) A subspace is selected according to the given context as shown in Figure 2(b).
(3) Each information resource is mapped onto the subspace and the norm of A1 is calculated as the correlation value between the context and the information resource, as shown in Figure 2(c).

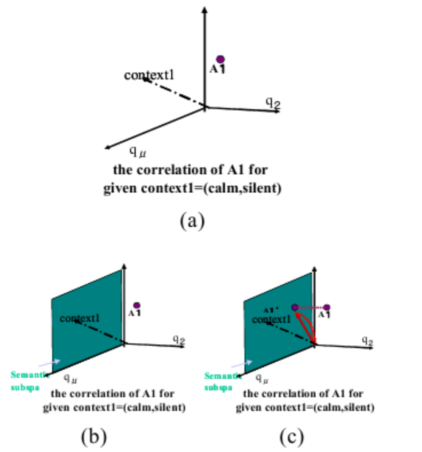

Figure 2. Semantic associative search in MMM

In MMM, the semantic interpretation is performed as projections of the semantic space dynamically, according to contexts, as shown in Figure 3.

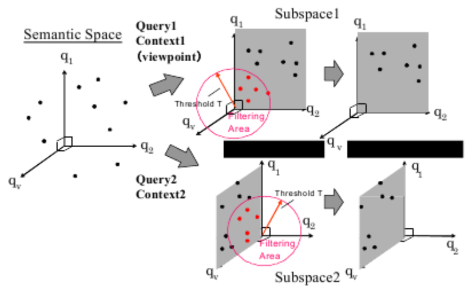

Figure 3. Semantic interpretation according to contexts in MMM

2.2. The outline of semantic associative search in MMM

The outline of the MMM is expressed as follows:

(1) A set of m words is given, and each word is characterized by n features. That is, an m by n matrix $\boldsymbol{M}$ is given as the data matrix.
(2) The correlation matrix $\boldsymbol{M}^T\boldsymbol{M}$ with respect to the n features is constructed from the matrix $\boldsymbol{M}$. Then, the eigenvalue decomposition of the correlation matrix is computed and the eigenvectors are normalized. The orthogonal semantic space $\boldsymbol{MDS}$ is created as the span of the eigenvectors which correspond to nonzero eigenvalues.
(3) Context words are characterized by using the n features and representing them as n-dimensional vectors.

(4) The context words are mapped into the orthogonal semantic space by computing the Fourier expansion for the n-dimensional vectors.
(5) A set of all the projections from the orthogonal semantic space to the invariant subspaces (eigen spaces) is defined. Each subspace represents a phase of meaning, and it corresponds to a context or situation.
(6) A subspace of the orthogonal semantic space is selected according to the user's impression expressed in n-dimensional vectors as context words, which are given as a context represented by a sequence of words.

The most correlated information resources to the given context are extracted in the selected subspace by applying the metric defined in the semantic space. The highest correlated words to the context are the representative words for lists of color impressions.

3. Feature Extraction and Subspace Selection

In this section, we describe the color feature extraction in the image database and the image query and then how to select the subspace of these color features.

3.1. Color feature extraction

The system extract color features using 130 basic color features of Color Image Scale [1]. These features consist of discretization of RGB color space based on human impression. The features contain 120 chromatic colors and 10 achromatic colors.

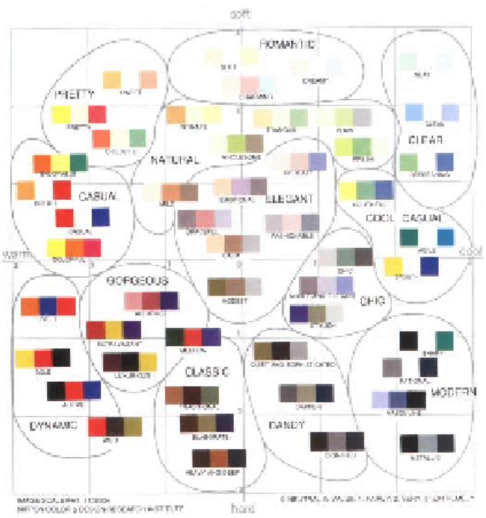

Figure 4. Key image impression words in Color Image Scale

3.2. Subspace feature selection

After obtaining the representative words for lists of color impressions described in Section 2, these representative words are then projected to the color impression metric based on Color Image Scale [1] which consists of 130 basic color features and 180 key image impression words, as shown in Figure 4. The projection calculates the relationships between the representative words from MMM and key image impression words in the Color Image Scale. The most significant colors which have the highest values of the projection is obtained and then used for selecting the color features among 130 color features of the image database and the image query.

4. Cluster oriented Similarity Measurement

After applying subspace selection, the system then clusters the image database using our previous work, Pillar-Kmeans algorithm. Pillar algorithm [13] is an algorithm to optimize the initial centroids for K-means clustering. This algorithm is very robust and superior for initial centroids optimization for K-means by positioning all centroids far separately among them in the data distribution.

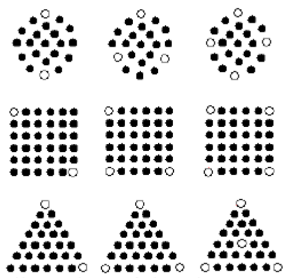

Figure 5. Illustration of locating a set of pillars (white points) withstanding against different pressure distribution of roofs.

Pillar algorithm is inspired by the thought process of determining a set of pillars' locations in order to make a stable house or building. Figure 5 illustrates the locating of two, three, and four pillars, in order to withstand the pressure distributions of several different roof structures composed of discrete points. It is inspiring that by distributing the pillars as far as possible from each other within the pressure distribution of a roof, the pillars can withstand the roof's pressure and stabilize a house or building. It considers the pillars which should be located as far as possible from each other to withstand against the pressure distribution of a roof, as number of centroids among the gravity weight of data distribution in the vector space. Therefore, this algorithm designates positions of initial centroids in the farthest accumulated distance between them in the data distribution.

The Pillar algorithm is described as follows. Let $X=\{x_i \mid i=1,\ldots,n\}$ be data, k be number of clusters, $C=\{c_i \mid i=1,\ldots,k\}$ be initial centroids, $SX \subseteq X$ be identification for X which are already selected in the sequence of process, $DM=\{x_i \mid i=1,\ldots,n\}$ be accumulated distance metric, $D=\{x_i \mid i=1,\ldots,n\}$ be distance metric for each iteration, and m be the grand mean of X. The following execution steps of the proposed algorithm are described as follows:

1. Set $C=\varnothing$, $SX=\varnothing$, and $DM=[]$
2. Calculate $D \leftarrow \mathrm{dis}(X,m)$
3. Set number of neighbors $nmin = \alpha \,.\, n / k$
4. Assign $dmax \leftarrow argmax(D)$
5. Set neighborhood boundary $nbdis = \beta \,.\, dmax$
6. Set i=1 as counter to determine the i-th initial centroid
7. $DM = DM + D$
8. Select ж $\leftarrow x_{argmax(DM)}$ as the candidate for i-th initial centroids
9. SX=SX $\cup$ ж
10. Set D as the distance metric between X to ж.
11. Set $no \leftarrow$ number of data points fulfilling $D \leq nbdis$
12. Assign DM(ж)=0
13. If $no < nmin$, go to step 8
14. Assign $D(SX)$=0
15. $C = C \cup$ ж
16. $i = i + 1$
17. If $i \leq k$, go back to step 7
18. Finish in which C is the solution as optimized initial centroids.

The centroids of clustering results from Pillar-Kmeans algorithm are used for calculating the similarity measurements to the image query. In this case, we use Cosine distance metric for similarity calculation.

5. Experimental System

For experimental study, we use the Ukiyo-e image datasets from Tokyo Metropolitan Library for representing the Japanese cultural image collections. These datasets contain 8743 typical images and artworks of famous paintings in Edo and Meiji era, including Hiroshige, Toyokuni, Kunisada, Yoshitoshi, Kunichika, Sadahige, Kuniteru, etc.

Figure 6. A trial image query

For experimental work, let image id 1, shown in Figure 6, be a trial image query with a context "calm quiet". First, the context is processed by MMM to calculate the highest correlated words to the context. The results of highest correlated words by MMM are "calm", "silent", "clean", "fair", "comfort", "health", "peaceful", etc. These highest correlated words then are projected to the color impression metric to obtain the most significant colors for subspace feature selection. The results of this projection is that the system select 35 most significant colors related to the impression words among 130 color features. This color feature subspace selection is applied for both the image database and the image query. After color feature subspace selection, we apply clustering using Pillar-Kmeans algorithm. In our case, we set 20 numbers of clusters. For similarity measurement, the system calculates the distances between the centroids of clustering results and the image query. The members of the closest similarity distances are selected as retrieved image results. Figure 7 shows the top 15 retrieved image results from the experiment.

Figure 7. The top 15 retrieved image results

For performance analysis, we extract the highest computed impression words from each retrieved image results using color impression metric. Table 1 shows the lists of 10-impression words from each retrieved image results. Table 1 performs 80% (12 of 15) retrieved image results containing "calm quiet" context. If we suppose the context words independently, the experimental results performed 60% (9 of 15) retrieved image results containing "calm" context and 73.33% (11 of 15) retrieved image results containing "quiet" context. Moreover, if we refer to human perception which the given context "calm quiet" may relatively consist of several meanings which are restful, sedate, solemn, sober, placid, tranquil, etc, the experimental results achieved all correct image retrieved results. This experimental result performed that our proposed system is able to reach high precision for image retrieval in accordance with the given context by the users.

6. Conclusion

A cluster oriented image retrieval system with context recognition mechanism for selection subspaces of color features is introduced in this paper. It implements a context in the image retrieval system to recognize the most important features in the image search by connecting the user impression to the query. The system applies a context recognition with Mathematical Model of Meaning (MMM) and then make a projection to the color features with a color impression metric in order to select subspaces of color features. The selected color features of the image database and the image query are then clustered using Pillar-Kmeans algorithm for similarity measurement purpose. The proposed system is examined in the experimental study with the 8743 Ukiyo-e image datasets from Tokyo Metropolitan Library for representing the Japanese cultural image collections. The experimental results described in Section 5 showed that the proposed system reached 80% precision rate to the given context and all correct results to the impression closeness of the context according the human perception.

Table 1. The impression words of retrieved images

Image id	Impression_words
1	subtle_and_mysterious, nostalgic, dry, provincial, simple_quiet_elegant, calm, assiduous, conservative, japanese, solemn
8	dry, familiar, gentle, large_hearted, agreeable_to, nostalgic, pleasant, mild, calm, amiable
74	dry, simple_quiet_elegant, calm, provincial, gentle_and_elegant, nostalgic, familiar, subtle_and_mysterious, simple_and_appealing, pleasant
14	dry, familiar, calm, gentle_and_elegant, pleasant, simple_quiet_elegant, large_hearted, restful, nostalgic, provincial
89	heavy_and_deep, subtle_and_mysterious, assiduous, bitter, conservative, authoritative, formal, rustic, quiet_and_sophisticated, old-fashioned
2	subtle_and_mysterious, dry, chic, provincial, nostalgic, simple_quiet_elegant, assiduous, exact, calm, formal
40	familiar, dry, restful, calm, amiable, simple_and_appealing, smooth, gentle, agreeable_to, simple_quiet_elegant
9	genteel, delicate, subtle, chic, sedate, elegant, sleek, gentle_and_elegant, graceful, mild
39	familiar, dry, amiable, smooth, gentle, tranquil, agreeable_to, simple_and_appealing, restful, simple_quiet_elegant
134	dry, provincial, chic, simple_quiet_elegant, subtle_and_mysterious, simple_and_appealing, sober, calm, nostalgic, solemn
97	subtle_and_mysterious, dry, provincial, chic, nostalgic, simple_quiet_elegant, exact, gentle, solemn, calm
102	authoritative, heavy_and_deep, intellectual, stout, solemn, formal, earnest, bitter, robust, sublime
93	subtle_and_mysterious, provincial, simple_quiet_elegant, dry, assiduous, modest, sober, chic, calm, nostalgic
85	provincial, rustic, old-fashioned, simple_quiet_elegant, aromatic, assiduous, pastoral, nostalgic, quiet_and_sophisticated, placid
4	dry, familiar, restful, tranquil, simple_and_appealing, pleasant, amiable, gentle_and_elegant, large_hearted, gentle

References

[1] Shigenobu Kobayashi, *Color Image Scale*, 1-st edition, Kodansha International publisher, 1992.

[2] T. Kitagawa, Y. Kiyoki, *A mathematical model of meaning and its application to multidatabase systems*, Proc. 3rd IEEE International Workshop on Research Issues on Data Engineering: Interoperability in Multidatabase Systems, pp.130-135, 1993.

[3] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, W. Equitz, *Efficient and effective querying by image content*, Journal of Intelligent Information Systems 3 (3–4), pp. 231–262, 1994.

[4] Y. Kiyoki, T. Kitagawa, T. Hayama, *A metadatabase system for semantic image search by a mathematical model of meaning*, ACM SIGMOD Record, Vol.23, No. 4, pp.34-41, 1994.

[5] Y. Kiyoki, , T. Kitagawa, Y. Hitomi, *A fundamental framework for realizing semantic interoperability in a multidatabase environment*, International Journal of Integrated Computer-Aided Engineering, Vol.2, No.1 (Special Issue on Multidatabase and Interoperable Systems), pp.3-20, John Wiley & Sons, 1995.

[6] J. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Gorowitz, R. Humphrey, R. Jain, C. Shu, *Virage image search engine: an open framework for image management*, Proc. The SPIE, Storage and Retrieval for Image and Video Databases IV, San Jose, CA, pp. 76–87, 1996.

[7] J.R. Smith, S.F.Chang, *VisualSEEk: a fully automated content-based image query system*, Proc. The Fourth ACM International Conference on Multimedia, Boston, MA, pp. 87-98, 1996.

[8] W.Y. Ma, B.S. Manjunath, *Netra: A toolbox for navigating large image databases*, Multimedia Systems 7 (3), pp. 184–198, 1999.

[9] R.C. Veltkamp, M. Tanase, *Content-Based Image Retrieval Systems: A survey*, Technical Report UU-CS-2000-34, 2000.

[10] H.M. Lotfy, A.S. Elmaghraby, *CoIRS: Cluster-oriented Image Retrieval System*, Proc. 16th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2004) 00, pp. 224-231, 2004.

[11] Y. Liu, D. Zhang, G. Lu, W.Y. Ma, *A survey of content-based image retrieval with high-level semantics*, Pattern Recognition 40, pp. 262–282, 2007.

[12] A.R. Barakbah, Y. Kiyoki, *An Image Database Retrieval System with 3D Color Vector Quantization and Cluster-based Shape and Structure Features*, The 19th European-Japanese Conference on Information Modelling and Knowledge Bases, Maribor, 2009.

[13] A.R. Barakbah, Y. Kiyoki, *A Pillar Algorithm for K-Means Optimization by Distance Maximization for Initial Centroid Designation*, IEEE Symposium on Computational Intelligence and Data Mining (CIDM), Nashville-Tennessee, 2009.