Annales Univ. Sci. Budapest., Sect. Comp. 51 (2020) 253–274

A SEMANTIC-BASED IMAGE RETRIEVAL SYSTEM USING A HYBRID METHOD K-MEANS AND K-NEAREST-NEIGHBOR

Le Manh Thanh (Hue, VietNam)
Nguyen Thi Uyen Nhi (DaNang, VietNam)
Nguyen Thi Uyen Thi (Hue, VietNam)
Pham Thi Ai Han (Hue, VietNam)
Van The Thanh (Ho Chi Minh city, VietNam)

Communicated by Attila Kiss
(Received October 4, 2019; accepted February 9, 2020)

Abstract. Semantic-based image retrieval is one of the important problems for the multimedia system and has garnered great interest in recent years. This paper proposes an improved method to build a self-balanced clustering tree, called iC-Tree. We build the Semantic-Based Image Retrieval (SBIR) based on iC-Tree named SBIR_iCT. We propose a hydrid method, which is a combination of K-mean, K-Nearest-Neighbor (KNN) to build iC-Tree, to improve the classification performance of this tree. The system uses images that are segmented into different regions and extracted visual features of each region. Semantic concepts are analyzed from those features based on the proposed C-Tree. The result of the retrieval process is a set of similar images with semantic concepts, which matches the query image and a visual word vector. hen, we design ontology for the image dataset and create the SPARQL query by extracting semantics of image. Finally, the semantic-based image retrieval on iC-Tree (SBIR_iCT) model is created hinging on our proposal. Experiments on 20,000 images of ImageCLEF dataset confirm that our proposed methods improve the retrieval performance, compared with an image retrieval system using K-means clutering, which we proposed earlier. At the same time, this performance is compared with some of recently published methods on the same dataset, show that our proposed methods are effective in handling problems of SBIR.

Key words and phrases: iC-Tree, SBIR_iCT, Hydrid method, K-means, KNN.
The ACM Computing Classification (*1998*): H.2.8, H. 3.3.

1. Introduction

With the development of the Internet, digital image collection is increasing rapidly. Image data plays an important role in many multimedia systems such as geographic information systems (GIS), hospital information systems (HIS), digital library systems (DLS), etc. Therefore, interest in image retrieval has increased. Many techniques have been proposed to perform image retrieval systems, that can be categorized into three types: Text-based Image Retrieval (TBIR) [2], Content-based Image Retrieval (CBIR) [7,8], and Semantic-based Retrieval Image (SBIR) [16]. TBIR can use of the text descriptors, keywords to retrieve images. In CBIR, images are indexed by their visual content such as color, texture, shapes, spatial layout, etc. However, the performance of TBIR, CBIR is still far from user's expectations [5,9]. In SBIR, semantic analysis needs to be incorporated in CBIR to retrieve images, such as using object ontology to define high-level concepts [17,22], using machine learning methods to associate low-level features with query concepts [3, 6], using of both the visual content of images and the textual information obtained from the Web for WWW image retrieval [15,18], etc. SBIR extracts visual features and maps it into semantics to describe content of image. Then system describes semantics and retrieves for related images.

This paper has proposed an improved method to build a self-balanced clustering tree (C-Tree), that we built earlier, called iC-Tree. We used the combination of K-means and K-Nearest-Neighbor (KNN) algorithms to create a data model that supports for the retrieval process. iC-Tree has been built for the purpose for clustering of image in order to improve the efficiency of clustering similarity and data classification, and easily controls the size of the tree as well as reducing the complexity of retrieval process on C-Tree. SBIR based on iC-Tree (SBIR-iCT) is built. Image retrieval process is executed based on the visual feature vectors of the query image. SBIR_iCT identifies the meaning of the query images; then, it retrieves the set of similiar images in visual features and extracts semantic contents of these images, from which it produces a visual word vector. The experiment of SBIR-iCT is performed on 20,000 images of ImageCLEF dataset [4,14]. The performance of SBIR_iCT is compared with an image retrieval system using K-means clutering (SBIR_CT), which we proposed earlier. At the same time, we also compared this performance with some of recently published methods on the same dataset. It shows that proposed methods and model are correct and effective. The paper's contributions include: (1) building an automatic clustering model by improving a self-balanced clustering tree structure (iC-Tree) base on a hybrid method K-means and KNN in order to increase the efficiency of clustering similarities and data classification; (2) proposing model and algorithms of SBIR_iCT to retrieve semantics of similar images; (3) building ontology for image dataset on the basis of triple language RDF and creating a SPARQL command [10,20] to retrieve similar images based on visual word vector; (4) building the experimental application of SBIR_iCT based on proposed model and algorithms.

The rest of this paper is as follows. In section 2 we survey and analyze related works of semantic image retrieval systems. In section 3 we build an improved self-balanced clustering tree based on a hybrid method K-means and KNN; Section 4, the general architecture of SBIR-iCT is described. Section 5, we build the experiment and evaluate the effectiveness of proposed method. Conclusions and future works are presented in section 6.

2. Related works

Semantic-based image retrieval has become an active research topic in recent times. There were many techniques of image retrieval by semantics that have been widely applied in many different digital systems, such as techniques to build a model for mapping between low-level features and high-level semantics [1], query techniques based on ontology to accurately describe semantics for images [18,19], techniques for classification data [11], content based semantics and image retrieval system for hierarchical databases [13], etc.

In 2008, Liu Y., et al. [12] proposed a region-based image retrieval system with high-level semantic learning. Image from dataset was segmented into different regions and was extracted low-level features of each region. During retrieval, a set of images whose semantic concept matches the query is returned. Their semantic image retrieval system allowed users to retrieve images using both query by region of interest and query by keywords, and experimented on 5000 COREL images. However, they only used machine learning techniques to associate low-level image features with query concepts, so high-level concepts could not be define accurately by ontology. Besides, the experiments in this paper were conducted using query by single specified region.

In 2008, Liu Y., et al. [13] proposed a region-based image retrieval system with high-level semantic learning. Image from dataset was segmented into different regions and was extracted low-level features of each region. During retrieval, a set of images whose semantic concept matches the query is returned. Their semantic image retrieval system allowed users to retrieve images using both query by region of interest and query by keywords, and experimented on 5000 COREL images. However, they only used machine learning techniques to associate low-level image features with query concepts, so high-level concepts could not be defined accurately by ontology. Besides, the experiments in this paper were conducted using query by single specified region. Ontology is built from triple language RDF. To query on ontology, we can use the SPARQL query language. G. Gombo and A. Kiss [10] (2017) give the triple patterns, and a query system finds the matching subgraphs. The authors present the PSparkql that extends the Sparkql with parallel query plan. They presented an optimization technique that uses some statistics about the graph (number of predicates, data properties). The query plan will be created with these statistics. They show less edges enough for the evaluation than the Sparkql is using.

In 2015, A.alzu'bi, et al [3] surveyed and evaluated the feasibility of the image query system according to CBIR and SBIR. Machine learning and data mining techniques were also introduced as the basis for building image query problems. From this work, it is shown that the content query and the semantic query are feasible problems and they can be applied to many different multimedia systems.

Yue Cao et al. [24] (2016) used CNN to classify images to generate binary feature vectors. On the base of this, the authors proposed Deep Visual-Semantic Hashing (DVSH) model to identify a set of similar images by semantics. However, this method has not yet mapped visual features to semantics of images. Their semantic image retrieval system experimented on Image. Vijayarajan V. et al. [23] performed image retrieval based on analyzing natural language to create a SPARQL query to find similar images based on RDF image description. This method has not yet implemented classification of image content to perform retrieval; therefore, the search process does not proceed from a given query image.

In 2017, Hakan Cevikalp et al. proposed a method for large-scale image retrieval by using binary hierarchical trees and transductive support vector machines (TSVM) [4]. Experiments of the paper were applied on ImageCLEF series to evaluate the accuracy of the proposed method. However, the binary tree structure in this article was only used to classify images but has not yet created a storage structure for the images. Therefore, the speed of retrieving and searching speed of images were limited. On the other hand, this work has not yet analyzed or extracted the visual semantics of the query images. Jiu M. et al. [11] proposed a multilayered neural network based on different nonlinear activation functions on each layer for image annotation. The SVM technique is applied to layering images at the output layer to extract a semantic level according to visual information for similar pocket-based images from BoW (Bag-of-Words). The method is evaluated on ImageCLEF dataset.

In 2018, Ouiem Bchir et al. [18] performed image retrieval based on the feature vectors extraction of regional objects in order to perform the process of partitioning to speed up image searching. In this method, the authors constructed a semantic mapping between visual features and high-level semantics.

The authors have experimented on ImageCLEF image series and showed the effectiveness of the proposed method. However, this work has not created a search model nor has it built a process of extracting and querying semantics on a given ontology.

In general, the recently approaches focused on providing methods for mapping low-level features to semantic concepts by using supervised or unsupervised machine learning techniques [13,22]; building data models such as graphs, trees or deep learning network to store low-level contents of images [1,11,19]; building ontology to define the high-level concepts [17,19,23], etc. However, the SBIR strongly depend on an external reliable resource such as automatically annotation images, ontology, and learning datasets. Therefore, SBIR is a research field that receives much attention and challenges. On the inherited basis of the previous researches and overcoming the limitations of the related methods as well as creating a system of semantic images retrieval to promote searching results, we propose an improved self-banlanced clustering tree based on the combination of K-means and KNN to storage low-level visual feature vectors. A semantic-based image retrieval system is proposed to search the semantic and URI of the images on this data structure.

3. An improved self-balanced clustering tree

3.1. General description of a self-balanced clustering tree:

A self-balanced clustering tree (C-Tree) is an automatic clustering data mining model, which performs automatically cluster feature vectors of image datasets and improves the performance of similar images retrieval. C-Tree is created based on a combination of hierarchical clustering, partitional clustering and semi-supervisor learning technique. Data stored on C-Tree are low-level visual features and concepts of objects, which are described in the image. Each image is segmented into different regions [13,17]; on each region of the image, the feature vector is extracted based on low-level visual features such as color, position, and shape, etc. Each feature vector is labeled and mapped to a conceptual subclass of images to describe visual semantics for each image area. Each image is extracted with many feature vectors and many semantic descriptions.

Figure 1 describes the node structure of C-Tree. Each node has elements *E*, each element included a feature vector of each region image *f_vector* and an image identifier *id* or a path from the parent node to the child node *Link*.

Internal node has $Link \neq 0$ and $id = 0$, leaf node has $Link = 0$ and $id \neq 0$. Each node has a maximum of M elements.

$\mathbf{E_1}$	f_vector 1	Link 1	id 1
$\mathbf{E_2}$	f_vector 2	Link 2	id 2
		...	
$\mathbf{E_M}$	f_vector M	Link M	id M

Figure 1. A node structure C-Tree

Figure 2 shows the structure of a self-balanced clustering tree. A C-Tree consists of a root, a set of internal nodes and a set of leaf nodes, which are used to partition feature vectors in clusters at leaf nodes. A root is a internal node without a parent node, containing at least 2 elements of internal node, feature vector of element in root is the cluter center of child node, and has the path *Link* to child node. The internal node is similar to the root, however it has one parent node, at least one child. Leaf node is a node without child nodes, and has identifier images *id*.

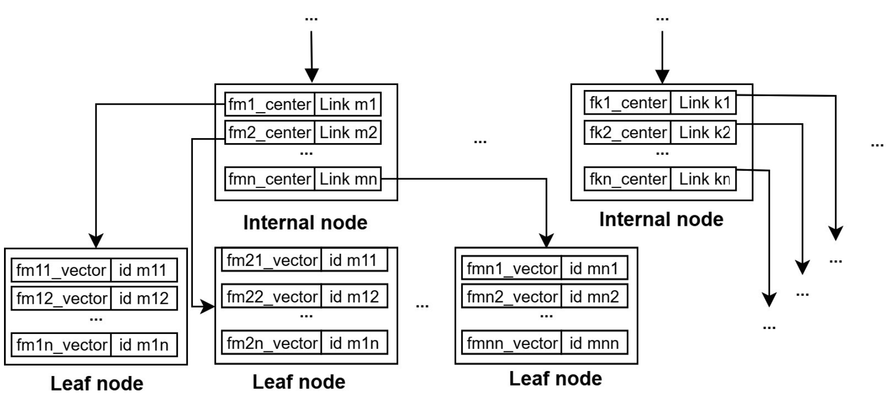

Figure 2. Structure of self-balanced clustering C-Tree

Algorithms for creating C-Tree include: Splitting the node, updating the cluster center, and inserting an element into the tree. We use K-means method to create C-Tree. Rules for creating C-Tree: (1) At the beginning, C-Tree has only one empty root node; (2) Each element is inserted to a leaf node of the C-Tree, which selected the nearest branch in similarity measure; (3) A leaf node is split into k-clusters if the number of elements exceeds M elements; (4) An internal node is split into internal k-nodes if the number of elements exceeds M elements. (5) When inserting element into the appropriate leaf node or splitting nodes, so the tree C-Tree has to be updated from the center from leaves to the root.

Because image data is constantly increasing, so C-Tree must be able to grow. C-Tree grows in the root direction, so the height of the leaf nodes increases equally, so C-Tree is a multi-branched tree that balances in height from the root to the leaf node in all directions.

3.2. An improvement for building a self-balanced clustering tree

In this paper, we propose an improved method for building a self-balanced clustering tree iC-Tree. We have provided a two stage approach for data clustering. At the first we use K-means algorithm to identify clusters in large data sets. At this stage, feature vectors which have the most association with each other will be placed in a cluster based on Euclidean similarity measure. In the second phase, after the data is clustered, we apply the KNN algorithm to find the best neighborhoods and classify the clusters into different classes. Therefore, we used the hybrid method K-means and KNN for clustering image datasets in order to improve the efficiency of clustering similarity. Figure 3 describes proposed model for building a self-balanced clustering tree iC-Tree.

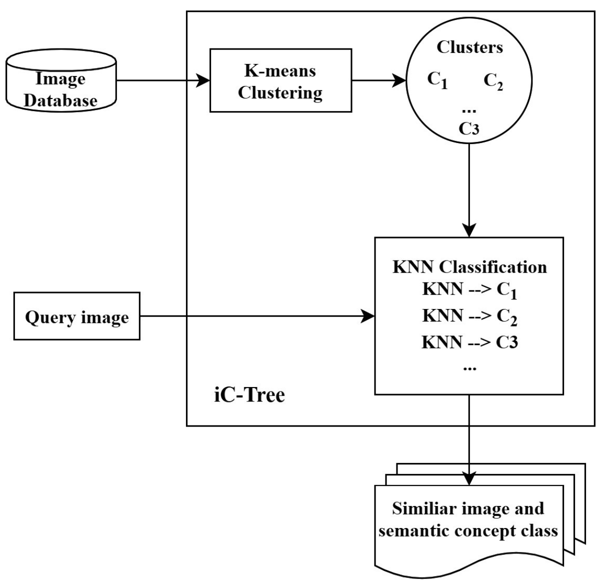

Figure 3. A proposed model for building iC-Tree

The process creating iC-Tree is based on inserting data, updating cluster center and splitting node to cluster feature vectors and the identifier of the images with the metadata of those images. The method of building an improved self-balanced clustering iC-Tree is as follows:

(i) **Step 1:** Initialize root node is empty. Insert the feature vectors in the image data set into the root of the C-Tree.

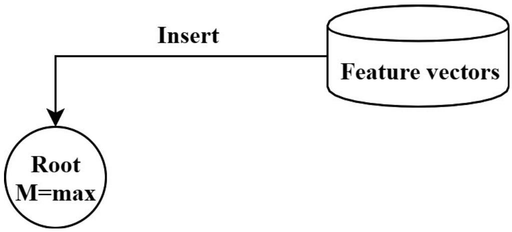

Figure 4. Step 1 of the process creating iC-Tree

(ii) **Step 2:** When the number of E elements in the node/root exceeds M elements, node is split into k-clusters by the KNN algorithm.

(iii) **Step 3:** Update the new cluster center and store it in the parent node. A path from the parent node to the child node is created.

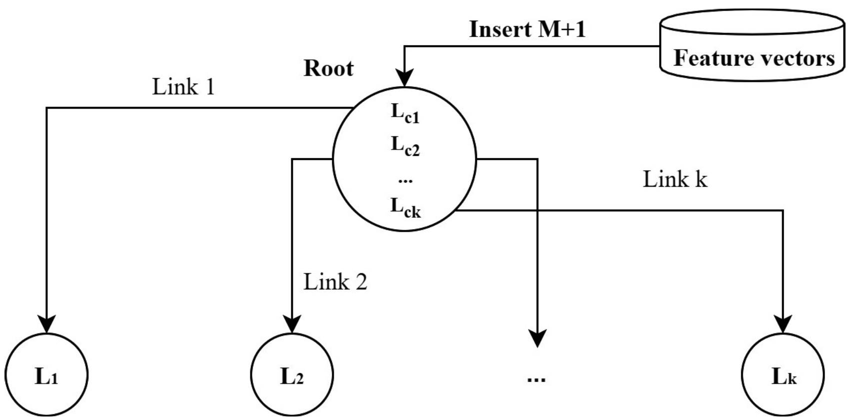

Figure 5. Step 2 and step 3 of the process creating iC-Tree

(iv) **Step 4:** Insert feature vectors in the image dataset into leaf nodes of the C-Tree, which selected the nearest branch in similarity measure based on the K-means algorithm. If a matching leaf node is not found, create a new leaf to insert the data.
Repeat steps 3,4 and 2 until the training dataset is completed

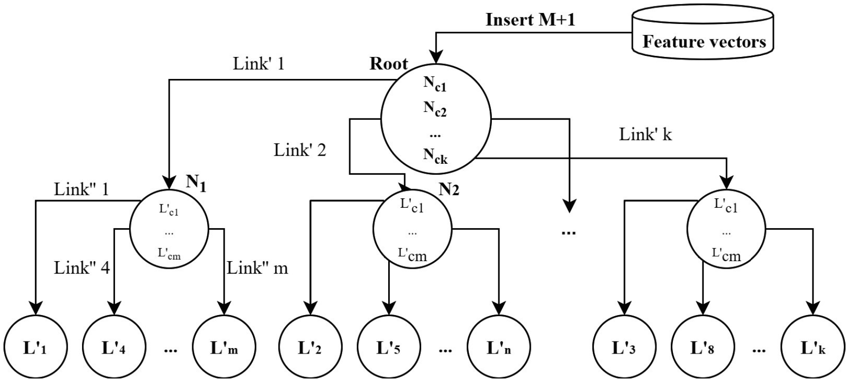

Figure 6. Step 2, 3, 4 of the process creating iC-Tree

Figure 4, 5, 6 describe the process creating iC-Tree with the combine of K-means and KNN algorithms.

4. The proposed model based on an improved self-balanced clustering tree for semantic-based image retrieval SBIR_iCT

4.1. Image dataset

In this paper, each image is segmented into different objects. Our experiment is evaluated on ImageCLEF dataset, which used a method of Hugo Jair Escalante [8] for image segmentation. ImageCLEF dataset contains 41 folders (from 0-th folder to 40-th folder) with 20,000 annotated and segmented images with diverse themes. Each region is assigned to a label, which is mapped with a semantic category, which has 276 classes.

Figure 7 shows an original image and five regions belonging to the classes including child-boy, cloth, hat, face-of-person, wall. When a region is segmented, that region is classified into a semantic category and assigned a label as a vocabulary. Each region is extracted a feature vector including features [12]: Region area, width and height; Features of locations including mean and standard deviation in the x and y-axis; Features of shape including boundary/area, convexity; Features of colors in RGB and CIE-Lab space including average, standard deviation and skewness, etc.

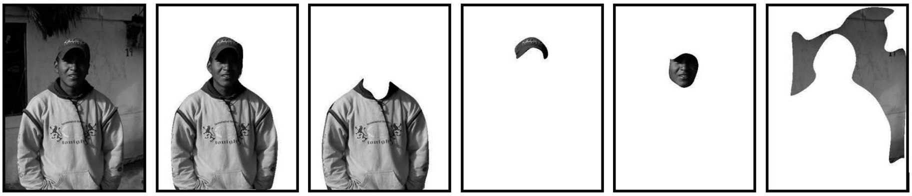

Figure 7. Origin image and segmented images (1004.jpg)

4.2. The proposed model of SBIR_iCT

The general architecture model of the SBIR_iCT system is described in Figure 8. The SBIR_iCT is implemented by classifying images into visual word vectors from the original clusters based on the improved self-balanced clustering tree iC-Tree. The result of this system is the set of similar images and its semantic conceptual subclasses and visual word vector. The SBIR_CT query system consists of two phases: (1) preprocessing phase: extracting feature vectors of each region image to generate inputs for training automatically clustering and classification data on an improved self-balancing clustering iC-Tree; (2) Image query phase: for each image query, we extracted feature vectors and retrieve on iC-Tree. The result of the querying process is a set of similar images and its semantic, visual word vector. Then, the SPARQL command is generated automatically from visual word vector to query on ontology.

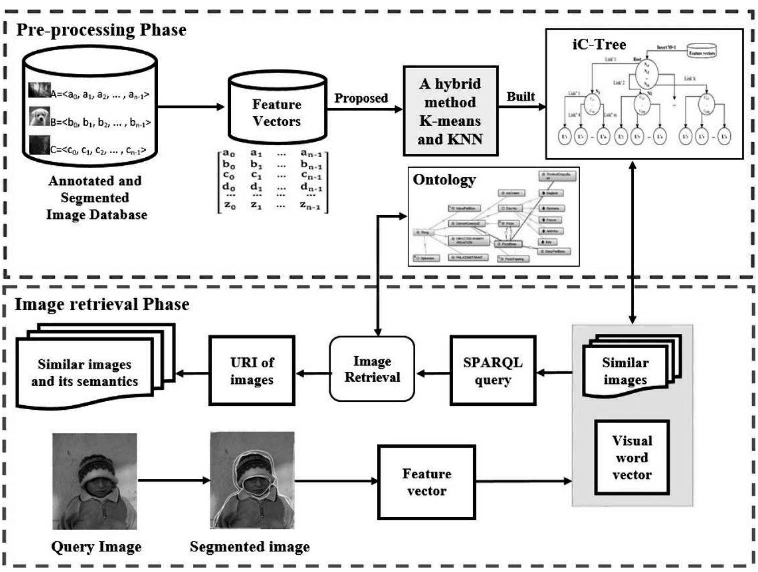

Figure 8. Model of semantic-based image retrieval SBIR_iCT

4.2.1. Pre-processing phase of SBIR_iCT

(i) **Step 1:** Extract data sample included feature vectors f and semantic category w of each region corresponding to each image in dataset;

(ii) **Step 2:** Train an improved self-balanced clustering tree structure iC-Tree to store data samples based on a hybrid method K-means and KNN and Euclidean measure;

(iii) **Step 3:** Build ontology as RDF triple language to describe semantics for image dataset. Ontology is described on Protégé in Figure 9.

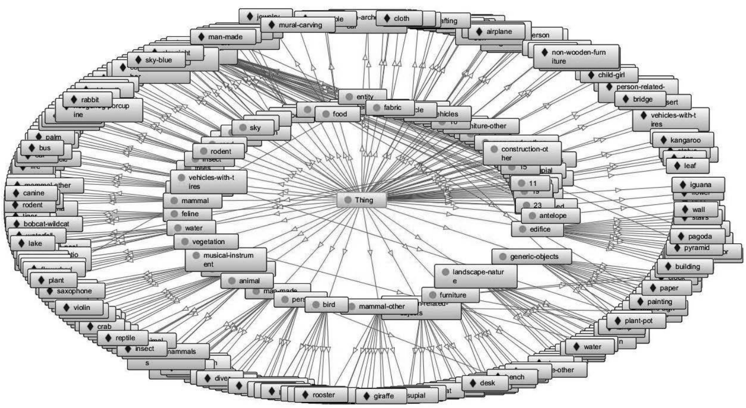

Figure 9: Ontology of ImageCLEF dataset on Protégé

4.2.2. Query phase of SBIR_iCT

(i) **Step 1:** For each query image IQ, the feature vectors of regions image are extracted and retrieved on iC-Tree; the result is a set of similar image, its semantic concepts and visual word vector.

(ii) **Step 2:** Create a SPARQL query based on the visual word vector and retrieve on ontology to produce a set of URIs and the metadata of images;

(iii) **Step 3:** Arrange similar images by similarity measure of the query image.

4.3. The query process based on iC-Tree

The query process based on iC-Tree is described in Figure 10. The feature vector of each query region is compared with feature vectors on the iC-Tree in the direction from the root to the leaf nodes by Euclidean distance. Each the feature vector of the query image region identifies a cluster at the leaf node of C-Tree. On the basis of C-Tree, we find matched clusters, from which we synthesize and find out similar images based on the frequency of images and Euclidean measure, which have the similar regions with the regions of the query image; these image regions belong to a semantic category, which is the basis for extracting semantics of query image. Based on the frequency of visual semantic concept classes, the visual word vector is created to perform the retrieval image in the semantic approach.

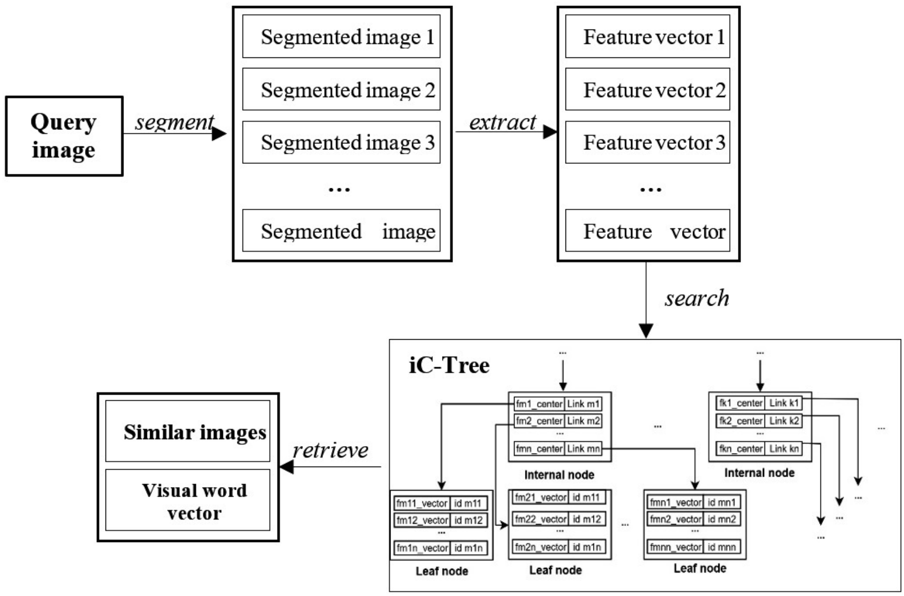

Figure 10. The model retrieval image based on the iC-Tree

Figure 11 is an illustration of the visual word vector of the set similar image, which is generated from retrieval image process 1000.jpg (in ImageCLEF dataset). This image is segmented into 5 regions with equivalent visual words for each region such as: child boy, cloth, wall, hat, and face-of-person. The retrieval images process of 1000.jpg on C-Tree produces a set of similar images and visual word vectors. Visual word vector is stored in text files with 5 vocabularies, which have the most frequency in the set of similar image: face-of-person (119), child-boy (80), cloth (67), wall (42), and hat (32).

Figure 11. Illustration of a visual word vector

Retrieval image algorithm on C-Tree is described as follows:

Algorithm IRCT

Input: feature vector f of query image I_Q, C-Tree
Output: Set of similar image *SI*
Function $IRCT(f, I_Q, v)$
Begin
    v=Root;
    If *(v is Leaf)* **then**
        $SI = v_i.E|i = 1..count$;
        Return SI;
    Else
        For $(f \in v)$ **do**
            $m = argmin\{Minkowski(f, v_i.f)|i = 1..v.count\}$;
        EndFor
        $v = v.E[m].l$;
        $IRCT(f, I_Q, v)$;
    EndIf
End

---

5. Experiments and discussions

5.1. Experimental application

We conducted an evaluation of the effectiveness of the proposed improvement, a K-means and KNN hybrid method, to cluster large data on a self-balanced clustering tree. First, we installed a self-balanced clustering tree with the original K-means, then an improved self-balanced clustering tree with a combination of K-means and KNN algorithms. In this section the results and evaluation of the proposed system SBIR_iCT on ImageCLEF dataset will be reviewed. In our experiment, the SBIR_CT system is built on the dotNET Framework 4.5 platform, the *C#* programming language. The SBIR_iCT system is performed in two phases: preprocessing phase and query phase, which are implemented on computers with Intel (R) CoreTM i7-8750H processors, CPU 2.70GHz, RAM 8GB and Windows 10 Professional operating systems. Figure 12 describes the SBIR_iCT system for semantic image retrieval.

In order to assess effectiveness of proposed method, we used the following as evaluation metrics: precision, recall, F-measure. The formulas of these values are as follows:

$$precision = \frac{|relevant\ images \cap retrieved\ images|}{|retrieved\ images|} \tag{1}$$

$$recall = \frac{|relevant\ images \cap retrieved\ images|}{|relevant\ images|} \tag{2}$$

$$F{-}measure = 2 \times \frac{(precision \times recall)}{(precision + recall)} \tag{3}$$

Our empirical data set is divided into two sections, one for training data and one for test data. Figure 11 describes the semantic image retrieval system.

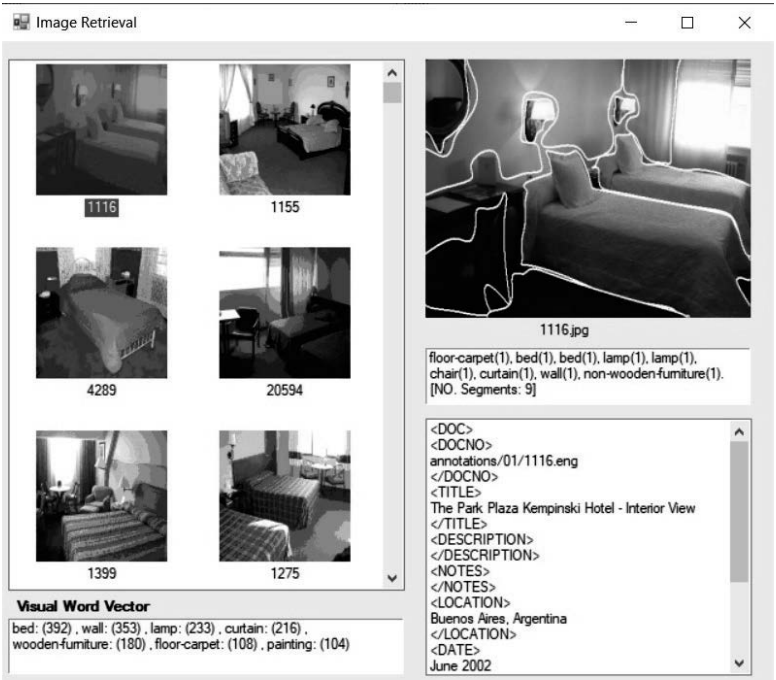

Figure 12. The semantic retrieval image system

5.2. Results and discussions

In order to assess effectiveness of proposed methods, we used the following as evaluation metrics: precision, recall, F-measure. We used 7092 images on the ImageCLEF dataset to get the image retrieval performance of the proposed methods.

5.2.1. Result of SBIR_CT

We installed a self-balanced clustering tree with the original K-means and built application of SBIR based on this data structure. Performance of image retrieval of SBIR_CT on ImageCLEF dataset is described in Table 1. The averages of performance are such as: recall 0.4403, precision 0.6510, F-measure 0.5227, and average query time 73.0605 ms.

Folders	No. images	Avg.recall	Avg.precision	Avg.F-measure	Avg.query time (ms)
00-10	2239	0.412843042	0.63972223	0.49943441	82.2642317
11-20	1820	0.459227484	0.61276569	0.52322946	76.7232867
21-30	1491	0.412109099	0.63408214	0.49720632	73.5502254
31-40	1542	0.477112611	0.71750647	0.57088284	59.7042889
AVG	7092	0.440323059	0.65101913	0.52268826	73.0605082

Table 1. Performance of image retrieval of SBIR_CT on ImageCLEF dataset

Figure 13 describes the average Precision, Recall, F-measure of 41 folders in ImageCLEF dataset. Each point on curve describes the average value of a folder. The ImageCLEF dataset consists of 41 folders (0-th to 40-th), so this figure has 41 point on each curve. This graph shows that the precision of the retrieval is at an average level, with many subjects of image dataset for high precision. The accuracy of the query system is concentrated in the 0.5 to 0.7 range. In particular, the precision of folder 39 is the largest at 0.8625 and the precision of folder 13 is lowest at 0.5137. The recall of the SBIR_CT system is quite low, so the F-measure is not high. The recall of the query system is concentrated in the 0.35 to 0.5 range, the recall of folder 20 is the largest at 0.5571 and the recall of folder 21 is lowest at 0.3154.

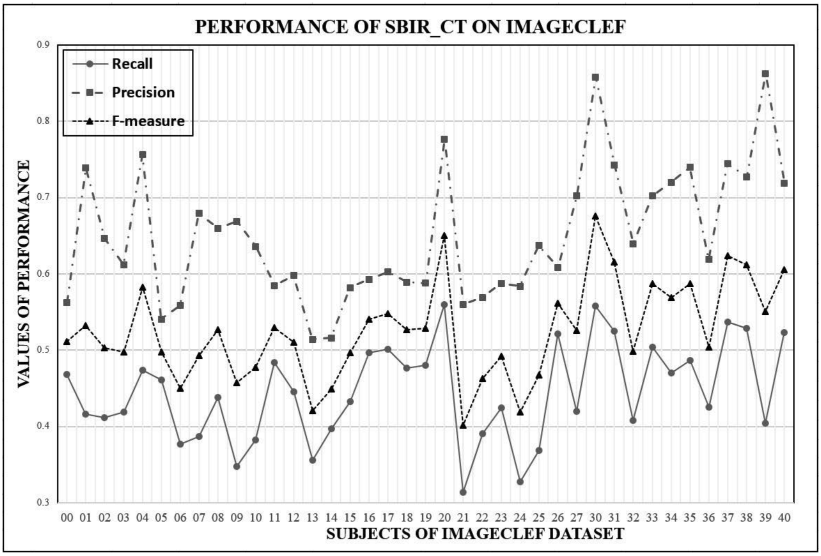

Figure 13. The performance of SBIR_CT on ImageCLEF dataset

5.2.2. Result of SBIR_iCT

The hybrid method K-means and KNN is proposed to build an improved self-balanced clustering tree iC-Tree. We build the Semantic-Based Image Retrieval SBIR_iCT based on this improved data structure. Performance of image retrieval of SBIR_CT on ImageCLEF dataset is showed in Table 2. The averages of performance are such as: recall 0.4640, precision 0.7083, F-measure 0.5588, and average query time 85.7898 ms.

Folders	No. images	Avg. recall	Avg. precision	Avg. F-measure	Avg. query time (ms)
00-10	2239	0.461833	0.692905	0.553593	91.5378976
11-20	1820	0.452912	0.666004	0.538203	87.2450577
21-30	1491	0.465221	0.687378	0.552251	92.0385671
31-40	1542	0.476163	0.786902	0.59128	72.3378652
AVG	7092	0.464032	0.708297	0.558832	85.7898469

Table 2. Performance of image retrieval of SBIR_CT on ImageCLEF dataset

Figure 14 describes the average Precision, Recall, F-measure of SBIR_iCT on ImageCLEF dataset. This graph shows that the accuracy of the query system is concentrated in the 0.6 to 0.75 range. The precision of folder 36 is the largest at 0.8730 and the precision of folder 12 is lowest at 0.6091. The recall of the query system is concentrated in the 0.45 to 0.5 range, the recall of folder 20 is the largest at 0.5980 and the recall of folder 21 is lowest at 0.3766.

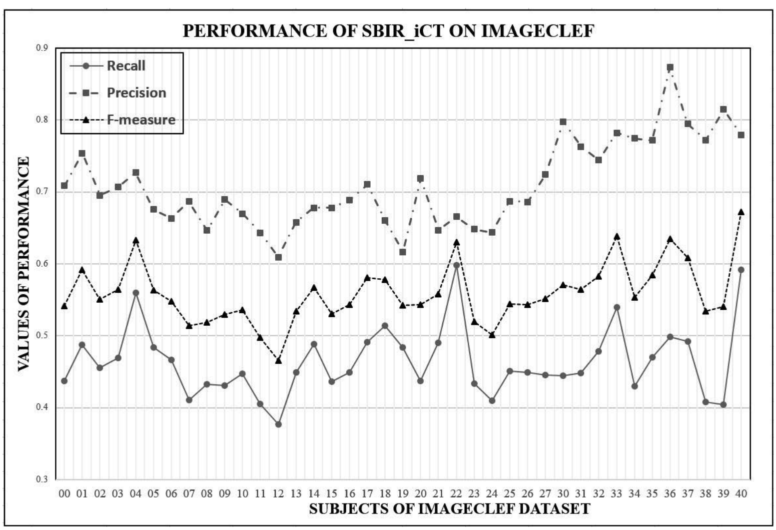

Figure 14. The performance of SBIR_iCT on ImageCLEF dataset

5.2.3. Discussions

We compare the performances obtained above to evaluate the effectiveness of proposed improvement. Figure 13 and 14 show that, the average Precision, Recall of SBIR_iCT is higher than SBIR_CT on the same dataset imgeCLEF. The accuracy of SBIR_CT and SBIR_iCT is concentrated in $[0.5, 0.7]$ and $[0.6, 0.75]$ respectively. Figure 16 describes a comparison in precision of SBIR_iCT and SBIR_CT on ImageCLEF. This graph shows that the majority of the points on the curve of SBIR_iCT are higher than the points on the curve of SBIR_CT. It proves that our improvements are effective. The result of this comparison is shown in Table 3.

Method	Mean Average Recall	Mean Average Precision (MAP)	Mean Average F-measure	Avg. query time (ms)
SBIR_CT (K-means)	0.440323	0.651019	0.522688	73.0605082
SBIR_iCT (Hydrid K-means and KNN)	0.464032	0.708297	0.558832	85.7898469

Table 3. Comparison performance of SBIR_CT and SBIR_iCT

From the above comparison results, the query time of the proposed method SBIR_iCT is more than SBIR_CT. However the Precision, Recall, F-measure of this improved method gives better. The experimental results show that proposed methods are correct and effective.

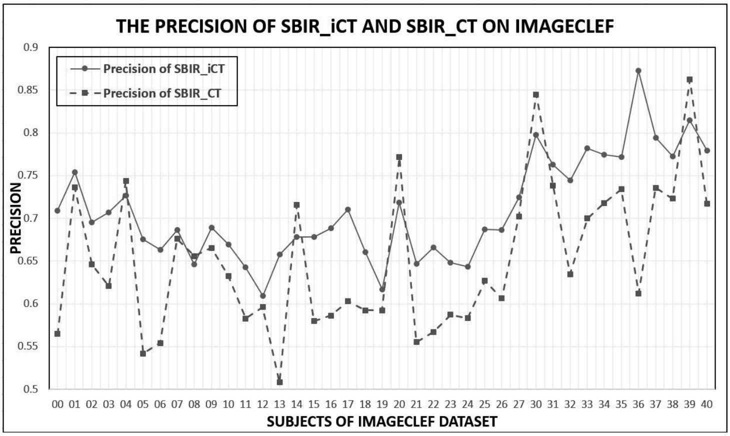

Figure 15. A comparison in precision of SBIR_iCT and SBIR_CT on ImageCLEF

At the same time, the values of Mean Average Precision (MAP) of proposed method are compared with other methods on the same dataset, which are described in Table 4. It shows that the accuracy of SBIR_iCT is higher than that of other methods. However, the MAP of Y. Cao's method [24] gives better value than the proposed method of this paper. In Y. Cao's method, the authors perform image retrieval relied on CNN. In this method, two vectors are created including the image vector and the sentence vector. This system only searches similar images and it does not create semantics of image content as well as does not query on ontology. So this method only performs the first stage of the semantic image retrieval. In our proposed method, we extracted semantics of image from low-level visual feature vectors based on iC_Tree. This process creates a set of similar images with their semantics and visual word vector and query on ontology. Then we automatically create a query based on SPARQL language and query on ontology. We compared this work to show the difference between two problems, including the image retrieval based on semantics and the semantic-based image retrieval. We intend to improve the retrieval accuracy of proposed system with different techniques in the future works.

The comparison results show the accuracy and effectiveness of the proposed model and algorithm. Therefore SBIR_iCT can be developed to improve the efficiency of semantic image retrieval systems.

Methods	**Mean Average Precision (MAP)**
H. Cevikalp 2017 [4]	0.4678
M. Jiu, 2017 [11]	0.5970
Y. Cao, 2016 [24]	0.7236
V. Vijayarajan, 2016 [23]	0.4618
Our proposed method SBIR_iCT	0.7082

Table 4. Comparison mean average precision (MAP) of methods on ImageCLEF dataset

6. Conclusions

In this paper, we implemented a semantic-based image retrieval system SBIR_iCT based on an improved self-balanced clustering C-Tree using hybrid K-means and KNN. The proposed model is based on semi-supervised learning techniques. At the same time, we developed a method for extracting semantic images on ontology. The retrieval process on iC-Tree find similar images and visual word vector; then the SPARQL command is automatically generated to query on ontology. We implemented our SBIR_iCT system based on the proposed methods, model and algorithms. The experiments are evaluated on

ImageCLEF dataset with the precision 70.82%, the recall 46.40% and the F-measure 55.88%. Experimental results are compared to other methods on the same image dataset. The experimental results show that proposed methods are correctness and effectiveness. SBIR_iCT system can be developed and improved to increase image retrieval efficiency. In a future work, we intend to improve our algorithm image classification by using deep learning techniques and building ontology from image collections on WWW.

Acknowledgment. The authors would like to thank the Faculty of Information Technology, University of Science - Hue University for their professional advice for this study. We would also like to thank HCMC University of Food Industry, University of Education, and research group SBIR HCM, which are sponsors of this research. This work has been sponsored and funded by Ho Chi Minh City University of Food Industry under Contract No. 507/ HD-DCT. We also thank anonymous reviewers for helpful comments on this article.

References

[1] **Allani O., H.B. Zghal, N. Mellouli and H. Akdag H,** Pattern graph-based image retrieval system combining semantic and visual features, *Multimedia Tools and Applications*, **76** (2017), 20287–20316.

[2] **Smeulders, A.W.M., M. Worring, S. Santini, A. Gupta and R. Jain,** Content-based image retrieval at the end of the early years, *IEEE Trans. Pattern. Anal. Mach. Intell.*, **22(12)** (2000), 1349–1380.

[3] **Alzu'bi, A, A. Amira A and N. Ramzan,** Semantic content-based image retrieval: A comprehensive study, *J. Vis. Commun. Image Represent*, **32** (2015), 20–54.

[4] **Cevikalp, H., M. Elmas and S. Ozkan,** Large-scale image retrieval using transductive support vector machines, *Computer Vision and Image Understanding*, **173** (2018), 2–12.

[5] **Crucianu, M., M. Ferecatu and N. Boujemaa,** Relevance feedback for image retrieval: a short survey, *Report of the DELOS2 European Network of Excellence* (*FP6*), 2014.

[6] **Demir, B. and L. Bruzzone,** A novel active learning method in relevance feedback for content-based remote sensing image retrieval, *IEEE Trans. Geosci. Remote. Sens.*, **53(5)** (2015), 2323–2334.

[7] **Eakins, J. and M. Graham,** *Content-Based Image Retrieval*, Technical Report, University of Northumbria at Newcastle, 1999.

[8] **Escalante, H.J., C.A. Hernández, J.A. Gonzalez, A. López-López, M. Montes, E.F. Morales and L.E. Sucar,** The segmented and annotated IAPR TC-12 benchmark, *Computer Vision and Image Understanding*, **114(4)** (2010), 419–428.

[9] **Gevers, T. and A. Smeulders,** Content-based image retrieval by viewpoint-invariant color indexing, *Image Vision Comput.*, **17** (1999), 475–488.

[10] **Gombos, G. and A. Kiss,** P-Spar(k)ql: SPARQL evaluation method on Spark GraphX, In: Muhammad, Younas; Irfan, Awan; Joyce, El Haddad (eds.) *2016 IEEE 4th International Conference on Future Internet of Things and Cloud Workshops* (*FiCloudW*) Vienna, Austria : IEEE, (2016) pp. 188–193.

[11] **Jiu, M. and H. Sahbi,** Nonlinear deep kernel learning for image annotation, *IEEE Transactions on Image Processing*, **26(4)** (2017), 1820–1832.

[12] **Pandey, H.S. and PriteeKhanna,** A semantics and image retrieval system for hierarchical image databases, *Information Processing & Management*, **5(4)**, (2016), 571–591.

[13] **Liu, Y., D. Zhang and G. Lu,** Region-based image retrieval with high-level semantics using decision tree learning, *Pattern Recognition*, **41(8)** (2008), 2554–2570.

[14] **Escalante, H.J. et al.,** The segmented and annotated IAPR TC-12 benchmark, *Computer Vision and Image Understanding*, **114(4)** (2010), 419–428.

[15] **Mehmood, Z., T. Mahmood and M.A. Javid,** Content-based image retrieval and semantic automatic image annotation based on the weighted average of triangular histograms using support vector machine, *Applied Intelligence*, **48(1)** (2018), 166–181.

[16] **Mezaris, V., I. Kompatsiaris and M.G. Strintzis,** Region-based image retrieval using an object ontology and relevance feedback, *EURASIP Journal on Advances in Signal Processing* 2004, Article number: 231946 (2004).

[17] **Minu, R.I. and K.K. Thyagharajan,** Multimodal ontology search for semantic image retrieval, Submitted to *International Journal of Computer System Science and Engineering* for February, 2012.

[18] **Ouiem Bchir, Mohamed Maher and Hadeel Aljam,** Region-based image retrieval using relevance feature weights, *International Journal of fuzzy Logic and Intelligent Systems*, **18(1)** (2018), 65–77.

[19] **Poslad, S. and K. Kesorn,** A Multi-Modal Incompleteness Ontology model (MMIO) to enhance information fusion for image retrieval, *Information Fusion*, **20** (2014), 225–241.

[20] **Pérez, J., M. Arenas and C. Gutierrez,** Semantics and complexity of SPARQL, In: *International Semantic Web Conference*, 2006, November, Springer, Berlin, Heidelberg, pp. 30–43.

[21] **Lee, M.-H., S. Rho and E.-I. Choi,** Ontology based user query interpretation for semantic multimedia contents retrieval, *Multimedia Tools and Applications*, **73(2)** (2014), 901–915.

[22] **Sethi, I.K. and I.L. Coman,** Mining association rules between low-level image features and high-level concepts, *Proceedings of the SPIE Data Mining and Knowledge Discovery*, vol. III, 2001, pp. 279–290.

[23] **Vijayarajan, V. and M. Dinakaran,** A generic framework for ontology based information retrieval and image retrieval in web data, *Human-centric Computing and Information Sciences*, **6(18)** (2016), 1–30.

[24] **Cao, Y., M. Long, J. Wang, Q. Yang and P.S. Yu,** Deep visual-semantic hashing for cross-modal retrieval, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, San Francisco, California, USA, pp. 1445–1454.

L. M. Thanh
Faculty of Information Technology
University of Science
Hue University
Hue
Vietnam
lmthanh@hueuni.edu.vn

N. T. U. Nhi
The Faculty of Statistics and Informatics
University of Economics
The University of Danang
DaNang
Vietnam
nhintu@due.edu.vn

N. T. U. Thi and P. T. A. Han
Quoc hoc Hue high school for the Gifted
Hue
Vietnam
tyz.nguyen32@gmail.com
hanpham94@gmail.com

Van The Thanh
Office of Scientific Research Management and Postgraduate Affairs
HCMC University of Food Industry
HoChiMinh city
Vietnam
thanhvt@hufi.edu.vn