Unified Concept-based Multimedia Information Retrieval Technique

Ridwan Andi Kambau
Faculty of Computer Science
University of Indonesia
Depok, West Java, Indonesia
+62-811883717
ridwan.andi@ui.ac.id

Zainal Arifin Hasibuan
Faculty of Computer Science
University of Indonesia
Depok, West Java, Indonesia
+62-81219103349
zhasibua@cs.ui.ac.id

***Abstract*—The explosion of digital data in the last two decades followed by the development of various types of data, including text, images, audio and video known as multimedia data. Multimedia Information Retrieval is required to search various type of media. There is comprehensive information need that can not be handled by the monolithic search engine like Google, Google Image, Youtube, or FindSounds. The shortcoming of search engine today related to their format or media is the dominance of text format, while the expected information could be an image, audio or video. Hence it is necessary to present multimedia format at the same time. This paper tries to design Unified Concept-based Multimedia Information Retrieval (UCpBMIR) technique to tackle those difficulties by using unified multimedia indexing. The indexing technique transforms the various of media with their features into text representation with the concept-based algorithm and put it into the concept detector. Learning model configures the concept detector to classify the multimedia object. The result of the concept detector process is placed in unified multimedia index database and waiting for the concept-based query to be matched into the Semantic Similarities with ontology. The ontology will provide the relationship between object representation of multimedia data. Due to indexing text, image, audio, and video respectively that naturally, they are heterogeneous, but conceptually they may have the relationship among them. From the preliminary result that multimedia document retrieved can be obtained through single query any format in order to retrieve all kind of multimedia format. Unified multimedia indexing technique with ontology will unify each format of multimedia.**

Keywords— information retrieval; concept-based search; multimedia information retrieval; ontology

I. INTRODUCTION

The existence of the internet makes rapid data growth, both structured and unstructured data. The growth of unstructured data is dominant (Fig.1). Unstructured data contain various types of data such as text, image, audio, and video are called multimedia. A large amount of multimedia data will make difficult to find relevant data or information, so it is required Information Retrieval (IR) System or Multimedia Information Retrieval (MIR) to search relevant multimedia document.

Initially, IR system based on text that using a keyword as a query. The early stage of IR system development is known as the classic information retrieval, IR system divided among three model [1]; The first model is Set Theory model that represents documents and queries through sets of keywords and similarities are derived from the set-theoretic operation on those set. The second model is Vector Space model uses term weighting to rank retrieved document and performing a partial match, this model represents documents and queries usually as vectors, matrices or tuples. The similarity of the query vector and document vector is represented as a scalar value and the third model is Probabilistic model treats the process of document retrieval as probabilistic inference. Similarities are computed as probabilities that a document is relevant for given query.

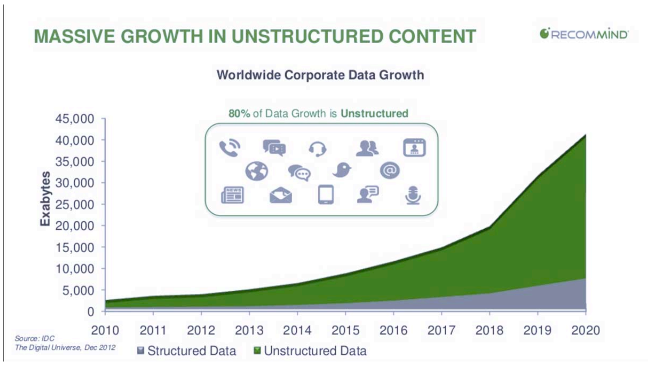

Fig. 1. The Massive Growth in Unstructured Data (Source: IDC The Digital Universe. Dec. 2012)

The various type of media need IR system not only text retrieval but also image retrieval, audio retrieval, and video retrieval. At first IR system with the various type of media using text annotation that performs searching various type of media or multimedia data automatically using text query. Like automatic image annotation in large scale [2], multi-label video annotation [3] and audio annotation [4]. The annotation gives the multimedia data textual description of the multimedia content. Text annotation makes retrieval process of multimedia data easy and powerful, even though subjective and incomplete.

978-1-5386-0549-3/17/$31.00 ©2017 IEEE

The weakness of automatic text annotation cannot recognize the content of multimedia data is handled by extracting feature content of the image, video, and audio that are called content-based MIR. Content-based MIR is MIR that search and analyzes the content of the image, audio, and video in this context might refer to color, shape, texture, motion, audio signal, a mel frequency cepstral coefficient (mfcc) that we call as feature extraction. Content-based retrieval comprises of content-based image retrieval (CBIR) for image searching [5] , content-based audio retrieval (CBAR) for audio searching [6] and content-based video retrieval (CBVR) for video searching [7].

Accurate and relevant information is not only dependent on the set of query or content from multimedia data but also determined by the context (user, time, location, document, environment, and event). Context based MIR improve effectivity of content based MIR, especially in the accuracy of the multimedia document retrieved, by adding context to retrieval technique [8].

Using a keyword in text-based IR, content of media in content-based IR and Context-based IR can not optimize the result of retrieval, still inaccurate and incomplete when the different keyword is used to describe the same concept or it does not mention query term explicitly on the label of multimedia data (annotation). Concept-based MIR is built to address the issues. Concept-based MIR can overcome the problems above with certain techniques one with such semantic processing of controlled vocabulary and ontology. And then this technique was implemented in multimedia data with more complex features and algorithm.

The shortcoming of existing Concept-based MIR is the result of retrieval only in one media, Concept-based Text Retrieval only for text, Concept-based Image Retrieval only for image, Concept-based Audio Retrieval only for audio and Concept-based Video Retrieval only for video. Whereas, the sometimes user need information from all kind of media only in one search technique.

This paper proposed an idea of Unified Concept-based Multimedia Indexing that integrating all kind of media (text, image, audio, and video) in one indexing technique. We use the indexing technique in Unified Concept-Based Multimedia Information Retrieval Technique.

The use of ontology and learning model as key enabling technology for unified concept-based MIR framework. Learning model creating the auto annotation and concepts automatically from multimedia dataset to classify the multimedia object. Ontology is used as the main component for representing the knowledge in the Semantic Similarities [9]. For this paper, we use an ontology to organize the concepts and creating the relationship between concepts. The ontology will also manage annotation, metadata and semantic features from multimedia into a concept. For experiments need we use specific ontology (cultural heritage from UNESCO) to create a data constraint in this research. [10]

In the second part of this paper discusses more related to IR System and MIR System, included Ontology and Concept Detector. While the third part describes the design Unified Concept-Based MIR technique, included the interface design and architectural design that we have proposed with its explanation, and the last, part four is the conclusion.

II. Related Paper

This paper focuses on the design of Unified Concept-based MIR, which in this technique, there are several major parts of the research including the IR system as the foundation of research, then MIR is divided into three parts: content, context, and concept-based, which emphasize the use of wide variety of media. For supporting part, there is ontology as knowledge-based and creating the concept and its relation.

A. Information Retrieval System (IRS)

First, Rapid data growth affects the development of IR System. IR System evolves constantly improve themselves. The weakness of old IR system will be rectified in the new one. The first and simplest IR system is the Boolean model [11] that stand long enough in its time as a search system. The Boolean model that included in the Set-Theoretic model using binary index term weight, it predicts the result only relevant or non-relevant, there is no ranking, which might lead to the retrieval of too few or too many documents. Vector model overcomes the shortcoming of the Boolean model with term weighting with considering how important this term for describing a document. The most popular term weighting is tf-idf (term frequency-inverse document frequency) based on frequency level [12].

Because of too many models in IRS, we have to select some model as a representation of all IRS model. First, text-based IRS with keyword has started with index term technique that using the term as a reference for indexing. Term Indexing [13] that perform indexing automatically was one of the early IRS, but the system had a very high computing cost and can not recognize synonymy and polysemy words. The issue of polysemy and synonymy is researched [14] with Latent Semantic Indexing(LSI). IRS with LSI had used Bag of Words (BoW) concept could reduce computational cost and recognize some synonymy and polysemy words, but in the experiment, many synonymy and polysemy still are not detected. This weakness is repaired by probabilistic Latent Semantic Indexing (pLSI) [15] that could improve the ability to recognize the words that have multiple meanings (polysemy). The next step of IRS development using three layers of Bayesian probability technique that is called Latent Dirichlet Allocation (LDA) [16] is used to increase the effectivity of IRS, particularly to handle synonymy and polysemy problems. However, LDA can not realize difficulties of semantic knowledge problems. The improvement of LDA is Tag-LDA that could fix semantic knowledge problems with using corpus and lexical database [17]. The use of lexical database or ontology and corpus become the latest trend in text-based IRS and emerging the new IRS is called Concept-based IRS. One of the early concept-based text retrievals [18] with Explicit Semantic Analysis (ESA). Concept-based text retrieval needs many resources and has to develop document corpus and concept detector.

B. *Multimedia Information Retrieval (MIR)*

MIR system is an information retrieval system that aims to extract information from multimedia data sources. We restrict the use of multimedia in a text, images, audio and video on the search process utilizes feature extraction of media content.

The main issue in MIR system is how to bridge the "Semantic Gap" or how to translate the easily computable low-level content-based media features to high-level concepts or terms which would be intuitive to the user [19].

The development of MIR system is divided in three major part, Content-based MIR, Context-based MIR and Concept-based MIR. We can see the explanation respectively.

1) *Content-based MIR*

Content-based MIR is divided in Content-based Image Retrieval (CBIR), Content-based Audio Retrieval (CBAR), and Content-based Video Retrieval (CBVR) and we will explain their characters and then give sample representation for each retrieval technique.

CBIR system can obtain images based on content or features such as color, texture, and shape. Based on research, CBIR is one of the active research fields from the last decade. Research about various technique for extracting colors like Color Histogram, Color Correlogram, Dominant Color Descriptor and Color Co-occurrence Matrix. The texture features like Tamura Texture Feature, Steerable Pyramid, Wavelet Transform Feature and Gabor Filter Features. The last is shaped features like Moment Invariants and Fourier Descriptor [5].

Content-based video retrieval (CBVR) can be regarded as an extension of CBIR by adding motion features because the video is a moving image. Video indexing scheme generally consists of three processes; video parsing, content analysis, and abstraction that is very different and more complex than the image content indexing. Abstraction is the process of extracting or constructing a subset of video data from the original video, such as keyframe or input of shots, scenes or stories. The outcome of abstraction from CBVR process is a keyframe. [7]

The performance of a comprehensive study on the different components of the system CBVR to understand the trade-offs between accuracy and speed on each component. Investigation directions including exploring the difference between low-level features and semantic based features, examine the factors of different compression and estimates during the search process and understand the trade-off between time and accuracy of the ranking.

There are many features that can be used to characterize the audio signal. Generally, they can be divided into two categories: acoustic features and semantic features. Acoustic features can describe an audio in terms of commonly understood acoustical characteristics like Loudness, Pitch, and Cepstrum. When Semantic features describe sound using personal descriptive language, like Timbre, Rhythm, and Events. [20]

2) *Context-based MIR*

Contextual Retrieval is defined as 'combine search technologies and knowledge about query and user context into a single technique in order to provide the most appropriate answer for user's information need'. [21]

Research on the contextual information retrieval field had proven that the state when the user conducts a search had a perceptible effect on the user's search behavior. The search context may include several dimensions such as time, location, user, current task etc. In MIRS field, it had taken a very important part of research aim to improve the relevance of the search result.

Here, some of Context-based MIR with context user, time & location, document and environment & event. In MIR with Context Document [22], contain two part in one system, the first part was CBIR and another was a context document. In CBIR using HSV Color and Gabor Filter while context document using index term and LSI Algorithm. The result was Combination text and image retrieval outperforms from single information retrieval. MIR with Context User was very popular than another context, social media often use this MIR with user context. [23]. With cluster algorithm, this MIRS with user context was better than a naïve model. CBIR with Context Time & Location often used in the gadget or device with an assortment of features this MIRS [24], can improve retrieval image & context location performance, with reducing computational Cost for checking location. MIR with Context Event using Hapori Search as sample paper to test its performance, compared with Mobile Bing Local and the result was the performance of Hapori Search. Evaluation using precision-recall denoted Hapori Search had good performance.

3) *Concept-based MIR*

Content-based MIR and Context-based MIR are difficult to describe semantic visual features or semantic audio features of MIR. Concept-based MIR has attempted to tackle this these difficulties by using manually built thesauri or by extracting latent word relationship and concept from the corpus. For multimedia data, it needs classifier to build concept detector model by gathering a large pool of multimedia data and using machine learning to select training set and testing set so that we catch the semantic visual feature or semantic audio feature in concept terms. Concept based MIR was divided in Concept-based Image Retrieval (CpBIR), Concept-based Audio Retrieval (CpBAR) and Concept-based Video Retrieval (CpBVR).

CpBIR systems rely on a technique Image Retrieval annotation text in the image as the basis for indexing and obtain image data. This technique has two shortcomings that completeness and consistency problems. The issue has been resolved with Semantic technologies [25] like ontology offers a promising approach to image retrieval as it tries to map the low-level image features to high-level ontology concept. The user can give concept/keyword as text input or can input the image itself. Semantic Image Retrieval is based on hybrid approach and uses shape, color, texture based approaches for classification purpose. We can see the illustration below in Figure.2.

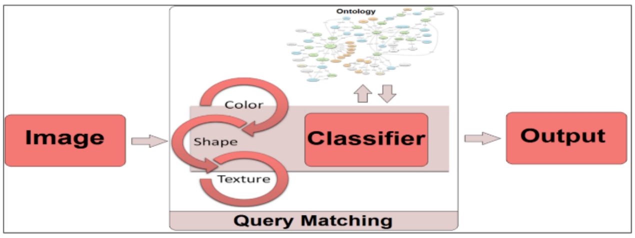

Fig. 2. Illustration of Semantic (Concept-based) Image Retrieval

And the more recent approach is the concept of image training a classifier for detection using tags and variants of Support Vector Machine which allows the use of weight training per sample. [26]. Combined with weighting appropriate tag mechanism, sample more relevant that plays a more important role in calibrating the final model concept detector. This technique can be applied to public collections online by collecting a diverse collection of images in bulk and using a probability calculation to select a training set that is associated with weight training. From this experiment was concluded using a weight derived from the tag significantly increase the strength of the resulting concepts detector.

CpBVR is a viable technique for video search and query what is behind the video. In order to obtain the desired video shot by a query should be defined based on user needs. Apart from the fact that the query can be any object, motion, texture, color and so on, with the concept of semantic queries more intuitive and realistic for the end user. An integrated semantic-based approach for similarity computation is proposed with respect to enhancing the retrieval effectiveness in CpBVR. In the [27] method is the based on the integration of knowledge-based like ontology and corpus-based semantic word (Wikipedia) similarity measures in order to retrieve video shot for the concept whose annotation is not available for the system.

C. Ontology

Ontology is a collection of concepts and relationships between concepts, which provides an abstract view of the application domain. In this study creating the ontology is proposed a new mechanism that can produce ontology automatically in order to make this approach easy to measure. For the modification of self-organizing tree algorithm (SOTA) that no one will build a hierarchy from top to bottom. And to find the right concept for each node in the hierarchy, the proposed choice of algorithm automatically from WordNet concept called linguistic ontology. By using Reuters21578 document corpus can be stated that the performance exceeds the SOTA modified algorithm Agglomerative Hierarchical Clustering (HAC) [28].

The initial purpose of Information Retrieval based ontology is to improve the ability to search the repository very large documents with the approach of the ontology-based scheme for annotation semi-automatic and retrieval system [9], where Ontology is the ability to build the concept hierarchy and supports logical inference and express semantic by connecting between concept.

In the article [10] proposed an ontology-based approach in managing cultural heritage and retrieval of multimedia collections, (Figure.3) where the objective is to allow the integration of different types of media and retrieve relevant heritage media of various types given query. For that built an ontology based on the cultural heritage of the UNESCO World Heritage Classification / WHC. Ontology here including media classification to accommodate multimedia collections are very heterogeneous but in the specific domain.

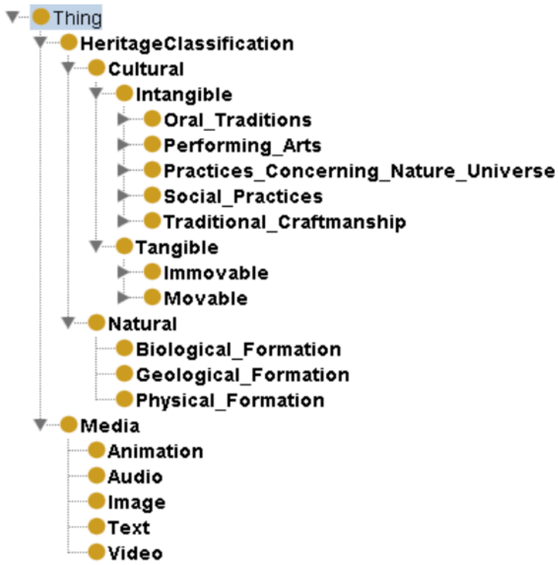

Fig. 3. Ontology design of World Heritage Classification by Portege [10]

III. Proposed Unified Concept-based Multimedia Information Retrieval Technique

In this section, we propose UCpBMIR stand for Unified Concept-Based Multimedia Information Retrieval technique. The design refers to is the previous paper, A Multidimensional Approach in Content-based Multimedia Information Retrieval System [29] that focus on the integration of feature extraction from various types of media (text, image, audio, and video) with Multidimensional indexing approach based on the content of any kind of media. The aim of this paper is how to retrieve a document in all media form, only in one-time search with one technique.

We develop the system to strengthen and improve the effectiveness previous paper with unified multimedia indexing in order to retrieve all kind media document in one-time search and using the UCpBMIR technique. The idea to unify all kind of media from [30] with data grid technique. The unified technique will enrich the meaning of the multimedia query. Creating concept and the relationship between concepts that are assisted by ontology scheme in the concept-based using semantic similarity of words need knowledge-based and corpus-based. Knowledge-based can be built from creating the ontology that can be built from a specific domain. (in this case using UNESCO natural heritage Ontology). Then Corpus-based can be generated from the multimedia dataset using concept detector.

The design of UCpBMIR technique consists of three main parts, the first part is user interface design, the second part is ontology design and the third part is architecture design of UCpBMIR.

A. User Interface Design

User interface design in information retrieval systems should be able to show the place to put an input query and select the type of the query (text, image, audio and video) and User Interface for the result of the search. The place for inputting query should have media option, in this case, there is the place to input text query, image query, audio query, and video query.

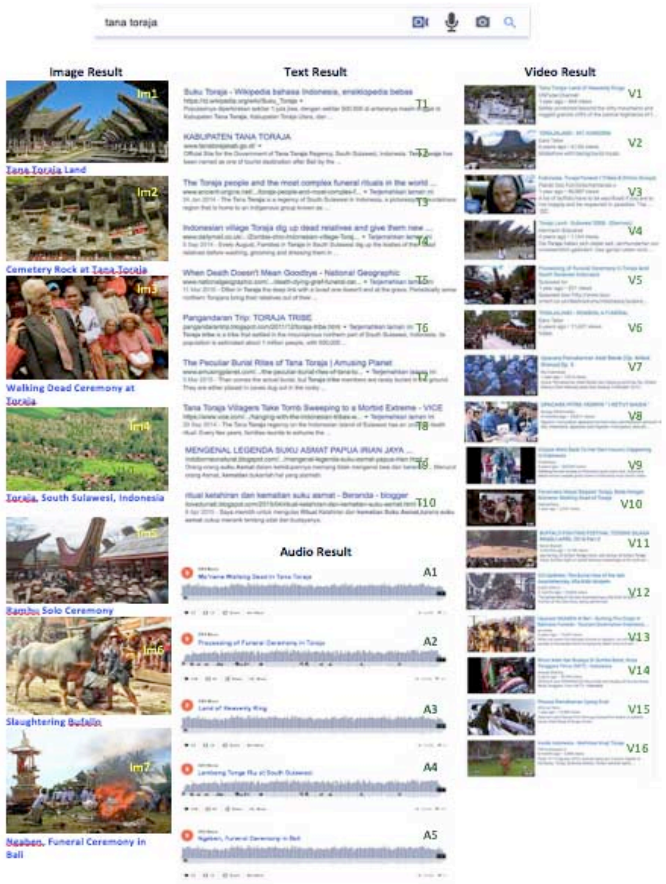

Fig. 4. User Interface Design of Concept-based Multimedia Information Retrieval in Multiplatform

For the Interface of result places to see the search results, we have to put Text Result, Image Result, Audio Result and Video Result in one page so the user can see all result from any kind of media. UCpBMIR should be designed to be able to catch multimedia queries from all types of media. For that design should be simple but able to accommodate all the needs of the query. We offer a design that can be seen in Figure 4, where we can put various types of queries to the form.

B. Ontology Design

Figure 5 shows extended ontology from Ontology Design of Cultural Heritage in Figure 3. Tribes is added as the part (sub class) of Intangible Cultural. Besides object properties of Media, there are an object properties addition: *relatedTo, hasEvent, partOf, similarTo*, and *historyOf* is used to create relationship among the cultural heritage.

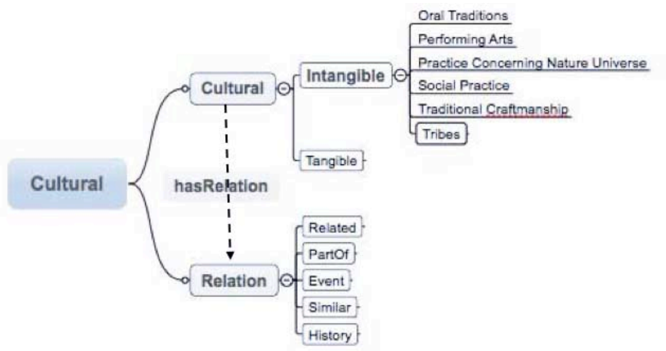

Fig. 5. Extended Ontology from Fig. 3

C. Architecture Design

The architectural design of UCpBMIR is quite complex and referring previous studies which have been mentioned above. This complexity is the result of research in the many areas that are integrated into this Architecture of UCpBMIR. The main areas are Indexing Task and Searching Task.

Our proposed architecture design can be seen in Figure 6. The Architecture Design of UCpBMIR that is divided into two major components; Indexing Task with some components and Searching Task with some components.

Indexing Task consists of Information Entities, Content Acquisition, Logical View Building and Concept Detector with Learning Process. Further, **Searching Task** contains Multimedia Query, Query Analysis, Semantic Similarity with Ontology function. Rank Result List.

1) Indexing Task

Information Entities is the repository of multimedia data collection with their metadata. Inputting into the repository is performed by **Content Acquisition**. Content Acquisition gets the data from administrator manually or web crawler automatically. The system will identify the format of the data (text, image, audio or video) and determines the physical and logical location for their storage or repository based on the file format. All multimedia data with the same format will be placed on the same logical collection. In **Logical View Building**, we process the data based on their format with some algorithm in order to produce object description. For text format is used Explicit Semantic Analysis Algorithm [31], for image format is used learning process (Support Vector Machine-SVM Classifier) [26], for video format is used integrated semantic approach [27] and for audio format is used Query by Semantic Example (QBSE) algorithm [32]. Logical View Building produces **Object Description** (D1,D2,D3, and D4) that represented in text form with metadata. **Concept Detector** with **Learning model** using Convolutional Neural Network [33] performs training and testing specific multimedia dataset (Indonesia Cultural Heritage) that [10] will identify and classify object description to become Concept terms. The result of Concept Detector is concept term that is stored in Unified Multimedia Index database and waiting for the matching process with the multimedia query in the searching task process.

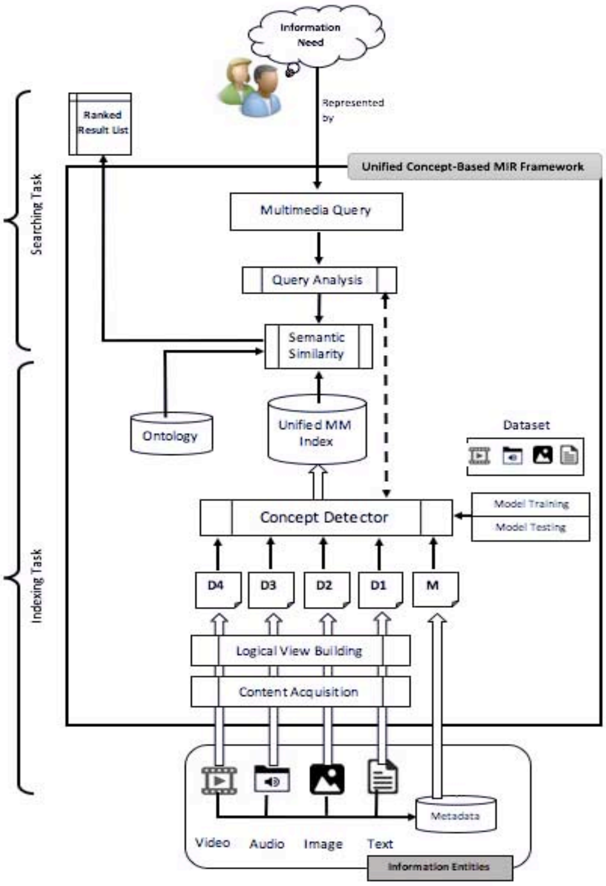

Fig. 6. Architecture Design of Unified Concept-based Multimedia Information Retrieval (UCpBMIR) Technique.

2) Searching Task

When user intent to find relevant information in the multimedia collection, they will put the **Multimedia Query** in the interface of UCpBMIR technique. **Query Analysis** using ***Concept Detector model*** will determine what kind of media that is inputted on the Query and try to translate it into Multimedia Concept Terms. The Multimedia Concept Terms as result of Query Analysis is conveyed to the **Semantic Similarities** with **Ontology** [34] creating the relation between concept terms and perform Matching Process with the Unified Multimedia Indexing database and simultaneously calculate the weight of concept terms. The output from Semantic Similarities is the **Rank Result List** based on the weight of terms.

Furthermore, the result of the Semantic Similarity is Multimedia data is ranked base on concept terms weight calculated. In this case with 'tana toraja' text query, we can see the technique result in Table 1, related with assigned number in Figure 5.

Table 1 describes the result of the technique from 'tana toraja' query, where we can see the media relation (textMedia, imageMedia, audioMedia and video Media) with object and concept relation with multimedia results.

We only use four concept relation (relatedTo, hasEvent, part of and similarTo) as the sample even though we can use more than four like historyOf or UniqueOf depend on the ontology relationship.

TABLE 1. The result of UCpBMIR with 'tana toraja' query on their ontology relation.

Object CH	relatedTo	hasEvent	partOf	similarTo
has TextMedia	T1, T7	T3,T5	T2,T6	T9,T10
has ImageMedia	Im1	Im2,Im3 Im6	Im4	Im7
has AudioMedia	A3	A1,A2	A4	A5
has VideoMedia	V1,V3	V3,V10, V11	V2, V4	V7,V8, V13,V14

D. Comparing UCpBMIR Technique with Existing IR System

Besides the advantages to retrieve the various type of media, UCpBMIR technique can retrieve conceptual relationship of the query based on the ontology. In this case, we use Indonesia Cultural Heritage as ontology domain.

Here, we try to compare between UCpBMIR technique with the existing IR system, Google and Yahoo for the text search, Google Image and CamFind for the image search, Youtube and Metacafe for the Video search and Findsounds and 4shared for the Audio search in Indonesia Cultural Heritage domain. We add object property *historyOf* relation that is not included in Table 1. The retrieval result can be seen in Table 2.

The retrieval result of UCpBMIR technique with 'tana toraja' query can retrieve all object with conceptual relation (*relatedOf, hasEvent, partOf, similarTo and historyOf*). Google and Youtube can not retrieve the object with *similarTo* relation, Google Image can not retrieve the object with *similarTo* and *historyOf* relation, and FindSound can not retrieve all the object because the technique only supports for sound and can not support audio news or music. Another IR system like CamFind can retrieve the object with *relatedTo* and *partOf* relation; Metacafe can not retrieve the object with *similarTo* and *historyOf* and 4shared can only retrieve the object with *relatedTo* relation.

TABLE 2. Comparing Retrieval Result between UCpBMIR technique with Existing IR system

Concept Relation / IR System	MEDIA	Relations				
		relatedTo	hasEvent	partOf	similarTo	historyOf
UCpBMIR	Text, Image, Audio & Video	Yes	Yes	Yes	Yes	Yes
Google	Text, Image	Yes	Yes	Yes	No	Yes
Yahoo	Text	Yes	Yes	Yes	No	No
Google Image	Image	Yes	Yes	Yes	No	No
CamFind	Image	Yes	No	Yes	No	No
Youtube	Video	Yes	Yes	Yes	No	Yes
Metacafe	Video	Yes	Yes	Yes	No	No
FindSounds	Audio	No	No	No	No	No
4shared	Audio	Yes	No	No	No	No

IV. EXPERIMENT

Prior to conducting experiments to identify the problem and analysis through the study of literature and modeling tools that generate a list of problems and the system model of Unified concept-based multimedia information retrieval technique. Then performing experimental setup, experimental implementation, and experimental evaluation. For more detail, we can see the following explanation.

A. Experimental Setup

In this section, we collect multimedia data from Indonesia Cultural Heritage that contain 500 data texts, 500 data images, 500 data audios and 500 data videos as the dataset and then put them in the multimedia repository. Further, define the metadata of multimedia data to add information in the object.

TABLE 3. Experimental Setup

Activity	Methods/Tools	Deliverables
Developing Dataset	Software Multimedia Processing	Dataset Multimedia
Defining Metadata	Software Multimedia Processing	List of Metadata
Collecting Application of Text, Image, Audio & Video Retrieval	Resource from GitHub and paid application	MIR Collection

B. Experimental Implementation

Experimental Implementation includes the construction of multidimensional indexing, creating learning model and develop the specific ontology.

TABLE 4: Experimental Implementation

Activity	Methods/Tools	Deliverables
Constructing Unified Multimedia Indexing	Java/Matlab/PHP etc	Unified Multimedia Indexing Algorithm/Model
Creating Learning Model	Convolutional Neural Network	Labels & concept terms of Multimedia Data
Develop Specific Ontology	Portege	Indonesia Cultural Heritage Domain

C. Experimental Testing and Evaluation

In this section, we will measure and evaluate effectivity of developed system with testing and compare the new system with unified concept-based multimedia information retrieval technique

TABLE 5. Experimental Testing

Activity	Method/Tools	Deliverables
Testing Concept-based (text, image, audio & video) Information Retrieval	Explicit Semantic Analysis, Support Vector Machine (SVM), Query by Semantic Example (QBSE) & Integrated Semantic Approach	Percentage of effectivity concept-based multimedia retrieval based on Precision and Recall and MAP (Mean Average Precision)
Comparing new System and Existing System	Unified Concept-based Multimedia Information Retrieval and Existing System	The new system more effective than each of concept-based multimedia retrieval

For experimental evaluation, we use Precision and Recall Approach to the measured effectivity of old and new system and use F-measure for unordered documents. For Ranking result we use Mean Average Precision (MAP) for system performance.

V. CONCLUSION AND FUTURE WORKS

The purpose of this paper is to design a unified multimedia information retrieval technique based on the concept. We want to develop unified multimedia indexing that can extract the various type of media and create concepts and its relations with ontology.

Our contribution in this paper is User Interface design that shows the place of inputting various types of queries format. The architecture of Unified Concept-based MIR technique with Unified Multimedia Indexing that can process any kinds of multimedia query format and generate various type of document retrieved (text, image, audio, and video). Furthermore, retrieval scenario for unified concept-based MIR and extended of ontology design with Indonesia's Cultural Heritage domain are the new contribution in this paper. It shows in Fig.4 as the preliminary results of unified concept-based MIR with tana toraja query and the results have proven conceptually of this research.

Our future work is semantic search with the emphasis on the detection of unified concepts-based in multimedia and development of retrieval algorithm from unified multimedia indexing technique. Beside that we want to put this application as web retrieval system, so we need more experiment with a big data, particularly multimedia data that is processed in machine learning and global ontology before we publish on the Internet. And we hope we can develop unified concept-based multimedia retrieval based on Indonesia Knowledge Domain like BahasaNet or Indonesian Wikipedia with the big multimedia corpus of Indonesia.

REFERENCES

[1] R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval: The Concept and Technology Behind Search, Essex: Addison Wesley, 2011.

[2] X. Li, L. Chen, L. Zhang, F. Lin and W.-Y. Ma, "Image annotation by large-scale content-based image retrieval," *Proceedings of the 14th ACM international conference on Multimedia,* pp. 607-610, October 2006.

[3] G.-J. Qi, X.-S. Hua, Y. Rui, J. Tang, T. Mei and H.-J. Zhang, "Correlative multi-label video annotation," *Proceedings of the 15th ACM international conference on Multimedia,* pp. 17-26, September 2007.

[4] D. Turnbull, L. Barrington, D. Torres and G. Lanckriet, "Semantic Annotation and Retrieval of Music and Sound Effects," *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING,* vol. 6, no. 2, pp. 467-476, 2008.

[5] F. Long, H. Zhang and D. D. Feng, "Fundamental of Content-based Image Retrieval," *Springer, Sign and Communication Technology,* 2003.

[6] G. Guo and S. Z. Li, "Content-Based Audio Classification and Retrieval by Support Vector Machines," *IEEE TRANSACTIONS ON NEURAL NETWORKS,* vol. 14, no. 1, pp. 209-215, January 2003.

[7] H. Zhang, "Content-based Video Analysis, Retrieval and Browsing," *Springer Singnal and Communication Technology,* pp. 27-56, 2003.

[8] D. Boughareb and N. Farah, "Context in Information Retrieval," in *International Conference on Control, Decision and Information Technologies (CoDIT)*, 2014.

[9] S. Kara and et al, "An Ontology-Based Retrieval System Using Semantic Indexing," *ICDE Workshop,* pp. 197-202, 2010.

[10] A. Rubhasy, A. Y. Paramartha, I. Budi and Z. A. Hasibuan, "Management and Retrieval of Cultural Heritage Multimedia Collection Using Ontology," *IEEE Explore,* 2014.

[11] S. Warwick, W. Frakes and R. Baeza, "Boolean Operations," *Prentice Hall,* vol. Information Retrieval, pp. 264-292, 1992.

[12] G. Salton and C. S. Yang, "On the Specification of Term Values in Automatic Indexing," *Journal of Documentation,* vol. 29, pp. 351-371, 1973.

[13] G. Salton and C. Buckley, "Term Weighting Approaches in Automatic Text Retrieval," *Information Processing and Management,*p.322, 1988.

[14] S. Deewester, S. T. Dumais, G. W. Furnas and et al, "Indexing by Latent Semantic Analysis," *Journal of American Society for Information Science,* p. 391, 1990.

[15] T. Hofmann, "Probabilistic Latent Semantic Indexing," in *Proceedings of the 22nd Annual International ACM SIGIR Conference* , 1999.

[16] D. M. Blei, A. Y. Ng and I. M. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning,* pp. 993-1022, 2003.

[17] F. LI, H. SHEN and T. HE, "Tag-Topic Model for Semantic Knowledge Acquisition from Blogs," in *7th International Conference NLP-KE*, Tokushima, Japan, 2011.

[18] O. Egozi, S. Markovitch and E. Gabrilovich, "Concept-based Information Retrieval Using Explicit Semantic Analysis," in *ACM Transactions on Information Systems*, 2011.

[19] M. S. Lew, N. Sebe, C. Djeraba and R. Jain, "Content-based Multimedia Information Retrieval: State of the Art and Challenges," *ACM Transaction on Multimedia Computing, Communication and Application,* vol. 2, no. 1, pp. 1-19, 2006.

[20] C. Xu, D. D. Feng and Q. Tian, "Content-based Retrieval for Digital Audio and Music," *Signal and Communication Technology,* 2003.

[21] J. Allan, "Challenges in Information Retrieval ang Language Modelling," *Report of a workshop held at the Center fo Intelligent Information Retrieval,* 2002.

[22] T. Westerveld, "Image Retrieval: Content versus Context," in *Proceeding Content-Based Multimedia Information Access*, Paris, 2000.

[23] W. Zemirli, L. Tamine-Lechani and M. Boughanem, "A Personalized Retrieval Model based on Influence Diagrams," in *Conference on Modeling and Using Context* , Roskilde, 2007.

[24] S. S. Tsai, D. Chen, G. Takacs, V. Chandrasekhar and et al, "Location Coding for Mobile Image Retrieval," in *Proceedings of the 5th International ICST Mobile Multimedia Communications Conference*, London, 2009.

[25] U. Manzoor and et al, "Semantic Image Retrieval : An Ontology Based Approach," *International Journal of Advanced Reseach in Artificial Intelligence,* vol. 4, no. 4, pp. 1-8, 2015.

[26] V. Papapanagiotou, C. Diou and A. Delopoulos, "Improving Concept-based Image Retrieval with Training Weight Computed from Tags," *ACM Trans. Multimedia Computing and Communication,* vol. 12, no. 2, pp. 32:1 - 32-22, 2015.

[27] S. Memar, L. S. Affendey, N. Mustapha and M. Ektefa, "Concept-based Video Retrieval Model based On The Combination of Semantic Similarity Measures," in *13th International Conference on Intelligent System Design and Application* , Malaysia, 2013.

[28] L. Khan and F. Luo, "Ontology Construction for Information Selection," in *IEEE International Conference on Tools with Artificial Intelligence*, Washington, 2002.

[29] A. Rubhasy, I. Budi, Z. A. Hasibuan and G. P. Mindara, "A Multidimensional Approach in Content-based Multimedia Information Retrieval System," *E-Indonesiana,* 2011.

[30] Z. A. Hasibuan, A. Kurniawan and R. Budiarto, "Multi-format Concept-based Information Retrieval Using Data Grid," *Journal of Advance Computing and Application,* vol. 1, no. 1, pp. 40-51, 2010.

[31] O. Egozi, S. Markovitch and E. Gabrilovich, "Concept-Based Information Retrieval Using Explicit Semantic Analysis," *ACM Transactions on Information Systems,* vol. 29, no. 2, 2011.

[32] L. Barrington, A. Chan, D. Turnbull and G. Lanckriet, "Audio Information Retrieval Using Semantic Similarity," *IEEE,* 2007.

[33] H. Xu and F. Su, "A robust hierarchical detection method for scene text based on convolutional neural networks," in *IEEE International Conference on Multimedia and Expo (ICME)*, 2015.

[34] S. Bin, F. Liying, Y. Jianzhuo and Z. Zhongcheng, "Ontology-Based Measure of Semantic Similarity between Concepts," in *WRI World Congress on Software Engineering*, 2009.