Singaporean Journal of Scientific Research(SJSR)
Journal of Technology and Engineering System(IJTES)
Vol.8.No.3 2016 Pp.190-198
available at :www.iaaet.org/sjsr
Paper Received : 08-03-2016
Paper Accepted: 19-04-2016
Paper Reviewed by: 1.Prof. Cheng Yu 2. Dr.M. Akshay Kumar
Editor : Dr. Chu Lio

AN INTERACTIVE AND SYNERGISTIC SYNTHESIS OF MULTIMEDIA DATA EXTRACTION ON VARIOUS PATTERNS

S.Hariharasudhan[*1] and Dr.B.Raghu[2]
Research Scholar[1], Professor[2]
Bharath University, Tamilnadu, India[1]
Sri Ramanujar engineering College, Tamilnadu, India[2]

ABSTRACT-----In this Modern globe advanced technology in multi-media acquiring and retention technology have led to enormous development in large and elaborated multi-media information database.. If these multi-media data files are taken for analysis, multipurpose useful and fantastic information's to end users can be discovered. Multi-Media data extraction agrees with the extractions of implicit and underlying knowledge, Multi-media data collection relationships, or other forms and patterns which are not explicitly and expressively stored in multi-media data files. Multi-media data mining and production is just an expansion, extension, enlargement of data mining, as it is considered an knowledge domain and interdisciplinary task that brings upon skillfulness and expertness in computer digitalised vision, multi-media processing, multimedia retrieval, data- mining, machine organization acquisition and learning , database ,unreal and artificial intelligence. This paper concisely accounts the multimedia excavation and mining , while remarks and comments referenced cover the leading theoretical and hypothetical problems.

KEYWORDS: - text data extraction; image data extraction; audio sound data extraction; video data extraction.

1 INTRODUCTION

In digital data acquiring and retention technology, the speedy development has tracked to the accelerated growing tremendously and quantity of information stored in databases. Though precious content may be concealing behind the data file ,the irresistible accumulation volume makes it challenging for humans to extract them without omnipotent tools. Multi-media data extraction groups that can mechanically extract or pull out semantically purposeful information from multi-media data files which are progressively in demand. Because of this, a various techniques have been projected scoping from simple measures to more intelligent systems like speaker emotion acknowledgement in audio [1], automatic summarisation of Television programs [2]. Unremarkably, multi-media file database systems stocks or stores and negociate a large aggregation of multi-media goals, such as Texts,Videos,Images, audios and hypertext data [3]. Thus, in multimedia representations, knowledge disclosure deals with non-structured collection of information. Due to this , we demand implementations for observing relationships between targets or portions within multi-media representation components, such as categorizing images founded on their content, data extracting patterns in sounds, categorizing speeches and sound, and acknowledging and following objects in video streams. In common, the multi-media dat files from a databases must be firstly preprocessed to modify their quality. Afterwards, these multimedia dat files experiences various transformations and characteristics extraction to create the most important salient features from the multi-media dat files.

With the created meaures, data extraction can be moved out with the help of data mining proficienciesor techniques to detect important patterns. These resultant subsequent patterns are then appraised and taken in order to get the final practical application cognition. The forth coming section 2 elaborates the Operations of practical application of Multi-media data extraction. Section 3 the characteristics data extraction from multi-media data is examined.Section 4 presents data pre-processing technique that includes data cleaning or improvement , normalization or Standardisation, transformation or Modification and characteristic option. In Section 5, the supervised and unsupervised possibilities, which are considered for multimedia mining are depicted. In Section 6, few practical application of multimedia data mining are mentioned, while some polite issues are commented in the finalsection.

2 OPERATIONS OF PRACTICAL APPLICATION OF MULTI-MEDIA DATA EXTRACTION

In Figure 1, we list the framework of employing multimedia extraction in various multimedia types.

Data collection and Grouping is the initial point of a learning process, as the attributes of unprocessed data regulates the overall manageable presentation. Then, the end of data pre-processing is to find essential features from unprocessed data. Data pre-processing regards data cleaning, standardization, modification, feature option, etc. Studying can be unambiguous, if enlightening features can be known at pre-processing level. Elaborated operation depends extremely on the nature of unprocessed data and trouble's domain. In few proceedings, prior cognition can be exceedingly valuable. For numerous systems, this stage is still mainly conducted by domain proficients.

The product of data information pre-processing is the upbringing set. Given a upbringing set, a acquisition model has to be selected to study from it. It must be noted that the steps of multi-media data extraction are often repetitious. The expert can also transition back and forth between leading works in order to ameliorate the outputs. When Multi-media data mining operation is Analyzed with ordinary data mining, multi-media data extraction reaches much more greater quality resulting from: a) the large volume of data information, b) the changeability and homogeneity of the multimedia data and c) the Multi-media complacent meaning is unobjective. The high spatial property of the property spaces and the size of the multi-media data sets make the attribute extraction a provocative problem. In the pursuing Section, we study the characteristic extraction process for multi-media data information.

3 CHARACTERISTIC DATA EXTRACTION

There are many kinds of characteristics data extraction. Here lets us proceed with two types.

1. Description-based and 2. content-based. The Description-based method uses metadata, such as caption,keywords, time of creation and size . The content-based method is based on the content of the object itself [3].

3.1 Characteristic extraction from passage

Text compartmentalization is a conventional categorization problem applied to the textual environment. It resolves the issue of assigning text data to predefined collections. In the studying stage, the tagged upbringing data are initially pre-processed to take away discarded information and to "normalize" the data [4]. For example, in text written document punctuation representational process and non-alphanumeric characters are discarded, because they do not help in grouping. Furthermore, all characters are regenerated to smaller case to modify matters. The second step is to work out the properties which are useful to differentiate one class from another class. For a text written document, this remarkably means distinguishing the keywords that resumes the table of contents of the document. These keywords may be learned is to look for words that occur oftentimes in the written document. These words taken care to be what the written document is about to intimate. The words that happen too often, such as "is", "the" are no use , since are prevailing in every written document.

These communal oral communication may be abstracted using a "Ban-list" of oral communication words during the pre-processing level. From the left oral communication words, a good algorithmic is to look for oral communication words that happen often in written document of the same class, but often in written document of another classes. To make out with written document of assorted extents, proportional frequency is preferable over absolute frequency [4]. Some writers used phrases, rather than individualistic words, as indexing terms [5], but the observational results found to have not been consistently encouraging effects. Another issue of text is variant. Variant refers to different forms of the same word, e.g. "keep", "keeps", "kept", "keeping". This may be cleared by rooting, means exchange of entire variants of a linguistic unit by a standard one [4].

3.2 Characteristic extraction from images

Image sorting categorizes images into semantic information databases that are pre-categorized. In the same semantic information databases, images will have big variances with unlike visual forms (e.g. images of living beings, images of Buildings). In addition images from different semantic databases the information might share a joint background. In [6], the authors distinguish three types of feature vectors for image description: 1) pixel-level characteristics, 2) region level characteristics, and 3) tile level characteristics. Pixel level characteristics store textural and spectral information about each and every pixel of the picture. For example, the fraction of the end members, such as water or concrete, can depict the content of the pixels. Region level characteristics explains groupings of pixels. Following the partitioning process, each region is represented by its boundary and amount of attributes, which gives information about the content of the region in terms of the end members and shape , size ,texture, fractal scale etc [6].

Tile level for picture characteristics gives information about whole images using texture, fractal scale .Moreover, other researchers proposed an information-driven framework that aims to highlight the role of information at various levels of representation [7]. This model adds a level of information: the Knowledge Level and Pattern and that incorporates environment, accompanying alpha-numeric data information and the semantic relationships revealed from the Picture.

3.3 Characteristic extraction from Audio sounds

Auditory communication data play an essential part in multi-media utilization. Music information has two main branches: 1.Audio information and 2. symbolical. Duration, volume, Attack, and instrument type of every single note are forthcoming collection of data. It is manageable to access statistical standards such as mean key and tempo for each sound item. Furthermore, it is accomplishable to attach to each item high-level signifier, such as device kind and figure. On the other way, auditory communication information agrees with real signals and any characteristics need to be mined through signal- analysis. The researchers of [8] utilized only perceptual features such as loudness, luminosity, pitch etc. On the other way, researchers has choosen only perceptual features to represent sound clips [9]. Another researcher team used 12 cepstral features, as well [10]. However, some of the most frequently used features for audio classification are [11], [12]:

1. Overall power: The temporal energy of an audio frame is termed by the rms of the audio signal magnitude within each frame.
2. Frequency Centroid (FC): It indicates the weighted average of frequency ingredients of a frame.
3. Zero Crossing Rate (ZCR): ZCR is also a remarkably used impermanent feature. ZCR counts the number of times that an audio signal crosses its 0 axis of rotation.
4. Bandwidth (BW): Bandwidth is the weighted average of the squared deviations between frequency component and its relative frequency centroid.

3.4 Characteristic extraction from Video Images

In video data extraction, there are different types of videos: a) the made (e.g. films, running videos),
b)the Unprocessed video Images (e.g. critical traffic videos, Investigation videos etc), and c) the surgical videos (e.g. ultra sound videos such as echocardiograms). Higher-level data information from video regards:

• Finding trigger effects
• Explaining typical and abnormal structures of activity, creating object centric views of an activity
• Sorting actions into titled categories .Clustering and determinant interactions between entities [13].The first stage for extracting raw video data information is classifying input exposures to a set of primary units are applicable to the structure of the video information. In developed videos information, the most used primary unit is a shot defined as a grouping of frames transcribed from a individual camera operation. Shot detection models can be categorized into multi collections: statistics based, pixel based, feature based ,transform basedand histogram based [14]. Color or grayscale histograms (such as in image mining) can also be used [15]. To segment video, color histograms, as well as motion and texture features can be used [16]. Commonly, if the difference between the two sequential frames is bigger than a definite threshold measure, then a shot boundary is reasoned between two comparable frames. The deviation can be discovered by examining the corresponding pixels of two images [17]. A set of other procedural characteristics copied from medium editing effects, motion and colours in videos images are presented in [18].

4 DATA PRE-PROCESSING TECHNIQUE

In a multimedia information database, there are many objectives which have many antithetical concepts of involvements. For example, only the colour concept can have 256 properties, with each investigating the frequency of a given colour in Pictures. The picture may still have other properties. Choosing a subset of Charactristic features is a method of reducing the problem size [21]. This bring down the dimensionality of the data information and enables basic cognitive process algorithms to operate faster and more impressively. The problem of characteristic interaction can also be self-addressed by building new charactristic features from the basic charateristics features set. This wonderful technique is called feature construction/transformation [22].

Sampling is also accepted by the statistics world organisation that argues “a powerful computationally intense procedure operating on a sub-sample of the data might in fact provide superior accuracy than a less sophisticated one using the entire data base” [19]. Furthermore, discretization can importantly trim down the number of possible values of the continuous characterictic feature, as big definite quantity of possible characteristic feature values contributes to draw-out and ineffective process of machine learning [20]. Moreover, normalization is also advantageous since there is frequently a large difference between the minimum and maximum values of the properties.

5 MODELS FOR MULTIMEDIA MINING

Multimedia classification and clustering are the supervised and unsupervised classification of multimedia data files are then grouped together.

5.1. Multimedia Classification models

Machine learning (ML) and significant information extraction can only be accomplished, when some objectives have been known and acknowledged by the System. The object identification problem can be adverted as a supervised description problem. Opening with the supervised models, we like to mention the determination trees. An overview of existing works in decision trees is furnished in [23]. Decision trees can be interpreted into a set of rules by making a separate rule for each path from root to a leaf in the full tree. Nevertheless, rules can also be immediately induced from training information data using a assortment of rule-based algorithms. An excellent overview of present works in rule based methods is given in [24]. Artificial Neural Networks (ANNs) are performing of inductive learning, based on computational procedure models of biological networks and neurons . A modern overview of present works in ANNs is given in [25].

A Bayesian network [26] is a graphical model for probabilistic relationships among a set of Charactristic features.Instance-based learning algorithms are lazy-learning algorithms [27] as they delay the induction or generalization process until categorization is performed. During the preparation phase, the lazy-learning algorithms postulates less computation time than eager-learning algorithms. Nevertheless, during the categorization process, they need more computation time. The Support Vector Machines (SVMs) is the hottest technique that regards the notion of a “margin”. Maximising the margin and thereby creating the largest possible distance between the separating hyperplane and the instances on either side of it, is proven to reduce upper bound on the expected generalisation error [28].

5.2 Multi-media Clustering Models

In unsupervised categorization, the difficulty is to group a acknowledged aggregation of unlabeled multimedia files into purposeful clusters according to the multi-media data values without a priori cognition. Clustering algorithms can be classified into hierarchical methods, partitioning methods, grid-based methods,,density-based methods and model-based methods. An superior survey of clustering techniques can be found in [29]. Partitioning methods are bifurcated into two subcategories, the centroid and the medoids algorithms**.**

The centroid algorithms symbolize each cluster by using the gravity centre of the occurrences. The medoid algorithms symbolize each cluster by means of the instances nearest to the gravity centre. The hierarchical methods group data occurrences into a tree of clusters [30]. Density-based clustering algorithms attempts to find clusters based on density of data points on a region. The main idea of density-based clustering is , for every instance of a cluster, the neighbor-hood of a granted radius has to contain at least a minimal number of occurrences [31]. Grid-based clustering algorithms prototypically quantize the clustering space into a limited number of cells (hyper-rectangles) and then execute the obligatory operations on the quantized space. Cells that contain certain number of components are treated as compact and the compact cells are connected to form the clusters [32].

5.3 Association rules in Multi-media Objects

The most association rules surveys have been concentrating on the corporate data in alphanumeric databases [33]. There are three standards of the association priciple: confidence ,support and interest. The support factor points the relative happening of both X and Y within the total information data set of transactions. It is termed as the ratio of the number of occurrences gratifying both X and Y over the total number of occurrences. The confidence factor is the measure of Y given to X and is defined as the ratio of the number of occurrences gratifying both X and Y over the number of occurrences satisfying X. The support factor points the frequencies of the happening patterns in the rule, and the confidence factor refers the strength of significance of the rule. The interest factor is a measure of manlike interest in the rule. For example, a full interest means that if a transaction contains X, then it is much more possible to have Y than the other components. Comparatively small research has been conducted on extracting multimedia data [34].

There are various types of associations principles: union between image content and non image content features. Association data extraction in multimedia data information can be changed into issues of association data extraction in transactional information databases. The picture can be modeled as a transaction, allotted with an ImageID, and the characteristics of the images are the items contained in the transaction. Therefore, mining the frequently happening patterns among different images gets mining the prevailing patterns in a set of transactions. In [35], the authors broaden the concept of content-based multimedia association concepts using feature charactristic localization. They presented the concept of innovative refinement in discovery of patterns in image representations.

6.APPLICATIONS OF MULTI-MEDIA OBJECTS

Satellite data is utilized in numerous different areas compassing from agriculture, forestry, and environmental studies.The applications using satellite data information includes activities of crop and timber surface area, statementing crop yields and forest harvest, observing urban growth, corresponding of ice for shipping, Possibilities of pollution, acknowledgement of definite rock types.

For information, the CONQUEST system [36] combines satellite data information with geo-physical data to discover patterns in global climate change. The SKICAT system [37] integrates techniques for image processing and data categorization in order to identify 'sky objects' captured in a very large satellite picture set. Moreover, the analysis and data extraction of traffic video images sequences in order to find information provide an economic formulation for every day traffic operations. There are some multimedia data extraction frameworks [39], [40] for traffic monitoring systems. Moreover, various respective methods for the spotting of faces in images and image series are reported in [41]. Detection of general sport video documents appears almost unfeasible due to the sizeable variety in sports.

7.CONCLUSIONS

This paper depicts cured illustrious techniques for multimedia data extraction. In text mining there are two open problems: lexical ambiguity and synonymity. lexical ambiguity refers to the fact that a word can have multiple significances. Characteristic Synonymity means that contrastive words can have the same or analogous message. In audio and video data extraction, an important open problem remains due to the accumulation of information across multiple media is combining video image and audio sound information into one comprehensive valuation.In image data extraction an unlock problem remains due to the the accumulation of different types of image information data.

REFERENCES:

[1] Petrushin, V.A., Emotion Recognition in Speech Signal: Experimental Study, Development, and Application, *Proc. 6th International Conference on Spoken Language Processing (ICSLP 2000)*, Beijing, 2000. Vol. IV, pp 222–228.

[2]Maybury M.T. (Ed.) *Intelligent Multimedia Information Retrieval*, AAAI Press/MIT Press,Menlo Park, CA / Cambridge, MA, 1997.

[3]Yoshitaka A. and Ichikawa T., A survey on content-based retrieval for multimedia databases.*IEEE Trans.on Knowledge and Data Engineering*, Vol 11, 1999, pp. 81-93.

[4]Sebastiani Fabrizio, Machine learning in automated text categorization. *ACM Computing Surveys*, Vol 34, 2002, pp. 1-47.

[5]Schutze, H., Automatic word sense discrimination. Computational Ling., Vol 24,1998, pp. 97–124.

[6]Zhang Ji, Wynne Hsu, Mong Li Lee. Image Mining: Issues, Frameworks and Techniques, in *Proc. of the Second International Workshop on Multimedia Data Mining (MDM/KDD'2001)*, San Francisco, CA, USA, 2001, pp. 13-20.

[7]Zhang J., W. Hsu and M. L. Lee. An Information-driven Framework for Image Mining, in *Proc. of 12th International Conference on Database and Expert Systems Applications*, Munich, 2001, pp. 232 - 242 .

[8]Wold E., T. Blum, D. Keislar, and J. Wheaton, Content-based classification, search and retrieval of audio, *IEEE Multimedia Magazine*, vol. 3, 1996, pp. 27-36.

[9]Liu Z., J. Huang, Y. Wang, and T. Chen, Audio Feature Extraction and Analysis for Scene Segmentation and Classification, *Journal of VLSI Signal Processing*, Vol. 20, 1998, pp.61-79.

[10]Foote J. et al, Content-based retrieval of music and audio, *Multimedia Storage Archiving Syst. II*, vol. 3229, 1997, pp. 138-147.

[11]Wang Y., Z. Liu, and J.-C. Huang, Multimedia Content Analysis, *IEEE Signal Processing Magazine*, Nov. 2000, pp. 12-36.

[12]Uitdenbogerd A. L., van Schyndel R. G., A review of factors affecting music recommender success, *3rd International Conference on Music Information Retrieval*, Paris, 2002, pp. 204-208.

[13]Rosenfeld A., D. Doermann, D. DeMenthon, Eds., *Video Mining*, Kluwer, 2003.

[14]Boresczky J. S. and L. A. Rowe, A comparison of video shot boundary detection techniques, *Storage & Retrieval for Image and Video Databases IV, Proc. SPIE 2670, 1996, pp.*170-179.

[15]Ardizzone E. and M. Cascia. Automatic video database indexing and retrieval.

Multimedia Tools and Applications, Vol. 4, 1997, pp. 29–56.

[16]Yu H. and W.Wolf. A visual search system for video and image databases. In *Proc. IEEE Int'l Conf. On Multimedia Computing and Systems*, Ottawa, Canada, June 1997, pp. 517–524.

[17] Zhang, H.J., Low, C.Y., Smoliar, S.W. and Wu, J.H., Video parsing, retrieval and browsing: an integrated and content-based solution, *Proc. ACM Multimedia '95*, pp. 15-24.

[18]Truong B.T. and C. Dorai, Automatic genre identification for content-based video categorization, *Proc. of ICPR' 2000*, pp 230-33.

[19]Friedman, J.H., Data mining and statistics: What's the connection?, Computing Science and Statistics, Vol. 29, No. 1, 1998, pp. 3-9.

[20]Berka, P. & Bruha, I., Discretization and grouping: Preprocessing steps for data mining.*Lecture Notes in AI*, Springer, 1998, pp. 239-245.

[21]Liu, H. & Motoda H., *Feature Extraction, Construction and Selection: A Data Mining Perspective*, Kluwer, 1998.

[22]Markovitch S. & Rosenstein D., Feature Generation Using General Construction Functions, *Machine Learning*, Vol 49, 2002, pp. 59-98.

[23]Murthy, Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey, *Data Mining and Knowledge Discovery,* Vol. 2, 1998, pp. 345–389.

[24]Furnkranz, J., Separate-and-Conquer Rule Learning. *Artificial Intelligence Review*, Vol. 13, 1999, pp. 3- 54.

[25]Neocleous, C. & Schizas, C., Artificial Neural Network Learning: A Comparative Review, *LNAI* 2308, Springer, 2002, pp. 300–313.

[26]Jensen, F.. *An Introduction to Bayesian Networks*. Springer, 1996.

[27]Aha, D., *Lazy Learning*. Dordrecht: Kluwer Academic Publishers, 1997.

[28]Cristianini, N. & Shawe-Taylor, J., *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*.Cambridge University Press, Cambridge, 2000.

[29]Jain A.K., Murty M.N., Flyn, P.J. (1999), Data Clustering: A Review, *ACM Computing Surveys*, Vol. 31, No. 3, 1999, pp. 264 - 323.

[30]Zhang, T., Ramakrishnan, R., and Linvy, M., BIRCH: An efficient data clustering method for very large data sets. *Data Mining and Knowledge Discovery*, Vol. 1, 1997, pp. 141–182.

[31]Ester, M., Kriegel, H.-P., Sander, J., and Xu X., A density-based algorithm for discovering clusters in large spatial data sets with noise. *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining*. Portland, 1996, pp. 226–231. [32]Wang W., Yang J. and Muntz.R. (1997), STING: A Statistical Information Grid Approach to Spatial Data Mining, *Proc. of the 23rd VLDB Conference*, Athens, Greece, 1997, pp. 186-195. [33]Agrawal R., R. Srikant, Fast Algorithms for Mining Association Rules, *Proc. of the 20th Int'l Conference on Very Large Databases*, Santiago, Chile, 1994, pp. 487-499.

[34]Shyu Mei-Ling, Shu-Ching Chen and R. L.Kashyap, Generalized Affinity-BasedAssociation Rule Mining for Multimedia Database Queries, *Knowledge and Information Systems*, Vol, 3, 2001, pp. 319-337.

[35]Za O., a Han, H. Zhu, In mining recurrent items in multimedia with progressive resolution refinement. *Proc. of the IEEE International Conference on Data Engineering*, 2000, pp.461-470.

[36]Stolorz P., H. Nakamura, E. Mesrobian, R. Muntz, E. Shek, J. Santos, J Yi, K Ng, S. Chien, Mechoso C., and J. Farrara. Fast spatio-temporal data mining of large geophysical datasets. In *Proc. of Int'l Conf. on KDD*, 1995, pp. 300–305.

[37]Fayyad U., S. Djorgovski, and N. Weir. Automating the analysis and cataloging of sky surveys. *Advances in Knowledge Discovery with Data Mining*, 1996, pp. 471–493.

[38]Wijesekera D. and D. Barbara. Mining cinematic knowledge: Work in progress. In

P*roc. Of International Workshop on Multimedia Data Mining (MDM/KDD'2000)*, Boston, pp. 98–103.

[39]Dailey D., F. Cathey, and S. Pumrin. An algorithm to estimate mean traffic speed using uncalibrated cameras. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, 2000, pp. 98–107.

[40]Cucchiara R., M. Piccardi, and P. Mello. Image analysis and rule-based reasoning for a traffic monitoring system. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, June 2000, pp. 119–130.

[41]Snoek C.G.M. and M. Worring, Multimodal Video Indexing: A Review of the State-of-the-art *Multimedia Tools and Applications*, Vol. 25, No. 1, 2005, pp. 5-35.

[42]Kobla V., D. DeMenthon, and D. Doermann.Identification of sports videos using replay, text, and camera motion features. *In SPIE Conference on Storage and Retrieval for Media Databases*,volume 3972, 2000, pp. 332-343.