Volume 6, Issue 9, September 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Study and Improvement on Video Association Mining Algorithm for Effective Confidence Support in Human Activity Dataset: An Event Perspective

K. Kumar*
Research Scholar, PRIST University,
Tamil Nadu, India

P. Sudhakar
Annamalai University,
Tamil Nadu, India

***Abstract* —Databases are to be converted into relational because of internet populating voluminous images and videos to its part. The process of data mining produces various patterns, correlations and associations from a given video data source. Numerous efficient algorithms have been proposed to do the above processes. The recent advancement of huge video databases pave way for Association Rule mining (ARM) which is one of the current data mining techniques designed to group objects together from large databases aiming to extract the interesting correlation and relation among huge amount of image or video data.**

***Keywords* — video data mining, association, dataset, events, PARTICIPATOR_KEY**

I. INTRODUCTION

Data mining [1] is an important tool for knowledge mining which includes many techniques: Association, Sequential Mining, Clustering and Deviation. Data mining activities can be classified as two categories: *Descriptive mining and Predictive mining.* Descriptive mining refers to the method in which the required characteristics of the data in the database are described. Clustering, Association and Sequential mining are the main tasks involved in the descriptive mining techniques tasks. Predictive mining derives patterns from the data in a similar manner as predictions. Predictive mining techniques include tasks like Classification, Regression and Deviation detection.

Association rule mining [12] is one of the complex problem treated in KDD and can be defined as extracting the interesting correlation and relation among huge amount of transactions.

This paper attempts to study and improvise a video association mining in the human activity dataset from event perspective.

II. VIDEO MINING

Video mining can be defined as the unsupervised discovery of patterns in an audio-visual content [1]. The temporal (motion) and spatial (color, texture, shapes and text regions) features of the video can be used for the mining task.

Oh and Bandi [2] proposed a framework for real time video data mining of raw videos which is shown in Fig. 1. In the first stage input frames are grouped to a set of basic units. In the second stage, it extracts some of the features from each segment. In the third stage, the decomposed segments are clustered into similar groups. The next two are the actual mining of the raw video sequences and the video data compression for the storage of these raw videos.

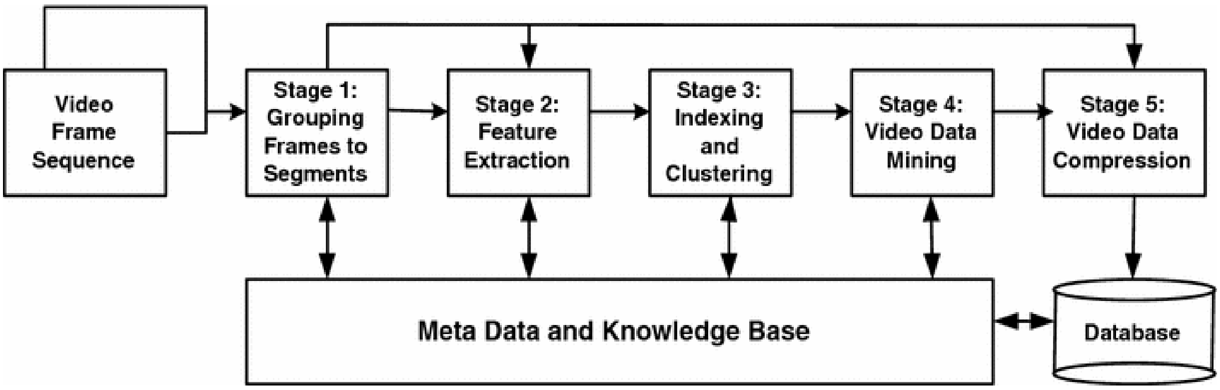


Fig 1. Stages in Video Mining

Video association mining is explored in two stages, the first being the video content processing in which the video clip is segmented into certain analysis units extracting their representative features and the second extracts the knowledge from the feature descriptors.

2.1 Related Works

Mongy et al. [3] presented a framework for video usage mining in the context of movie production that analyse the user behaviours on a set of video data to create suitable tools to help people in browsing and searching a large amount of video data.

Sivaselvan et al. [4] presented a video association mining that consist of two key phases. First, the transformation phase converts the original input video into an alternate transactional format namely a cluster sequence. Second, the frequent temporal pattern mining phase that is concerned with the generation of the patterns subject to the temporal distance and support thresholds as specified or chosen.

Lin et al. [5] developed a video semantic concept discovery framework that utilizes multimodal content analysis and association rule mining technique to discover the semantic concepts from input video data.

Chen and Shyu [6] proposed a hierarchical temporal association mining approach that integrates the association rule mining and the sequential pattern discovery to systematically determine the temporal patterns for target events.

Goyani et al. [7] proposed an A-priori algorithm to detect the semantic concepts from the cricket video.

Maheshkumar [8] proposed a method that automatically extracts silent events from the video and classifying each event sequence into a concept methodology with help of sequential association mining.

Lin and Shyu [9] proposed a weighted association rule mining algorithm to capture different significant degrees of the items (feature-value pairs) and generating the association rules for video semantic concept detection.

Kea et al. [10] developed a method based on the frequent pattern tree (FPTree) for mining association rules in video retrieval.

III. PROBLEM DEFINITON

Zhu et al. [11] proposed a multilevel sequential association mining algorithm to explore the associations between the audio and visual cues and classified the associations by assigning each of them with a class label using their appearances in the video to construct video indices. Generally, two measures (support and confidence) have been used to evaluate the quality of an association. The smaller the temporal distance between neighbour items the larger is their correlation. The associations with a large temporal distance between neighbour items suffer with weaker correlation and therefore can imply almost no knowledge. This algorithm is presented, analysed and suggested for improvement when applied on proposed human activity dataset, in the further sections.

```
Procedure VAMining ()
Input: (1) Hybrid data steam D; (2) max. association level max_level; and (3) TDT and minimal support and confidence in
       selecting associations at different levels minSup[l], minConf[l], l=1,…, max_level.
Output: Mined multi-level video associations
For (l=1; l ≤ max_level; l++)                                   // mine associations at various levels
  D[l] = Filter_Dataset (D, l);                                 // process items that are no larger than level l
  I [l, 1]= Get_1_ItemSet (D [l], l)                            // find 1-ItemSet at level l
  L[l, 1]=Get_1_LItemSet (D[l], I[l,1], minSup[l])              // find 1-LItemSet at level l
  For ( k = 2; L[l, k-1] ≠ Ø; k ++ )                            // mine associations with different lengths
    I [l, k]=CandidateGeneration(L[l, k-1])                     //generate candidates, see Fig. 16
    L[l, k] ← Φ                                                 // initialize k-LItemSet
    For each k-ItemAssociation {X} in I [l, k], {X}∈ I [l, k]   //evaluate each generated candidate
      TS{X}_TDT= Calculate_TS (D [l], TDT)                      // calculate temporal support of X
      Conf{X}_TDT= TS{X}_TDT / Min (TS{X_1},.., TS{X_k})        // calculate confidence of X
      L[l, k]←L[l, k] ∪ {X}| {X}∈I [l, k],  TS{X}_TDT ≥ minSup[l] & Conf{X}_TDT ≥ minConf[l]
              //select associations with their temporal support and confidence larger than the given thresholds
Endfor
```

Fig.2 Pseudocode for multilevel video association mining

1. A video item is a basic unit in association mining. A hybrid video stream is constructed based on sequential or continuous shot sequence.
2. An L-ItemAssociation is an association that consists of L sequential items. For example, “XY” is a 2-ItemAssociation and “XYZ” is a 3-ItemAssociation.
3. An ItemSet is an aggregation which consists of video associations.
4. L-LItemSet is an aggregation of all L-ItemAssociations whose temporal support is above than a given threshold.
5. Given a transformed hybrid video stream, the temporal distance (TD) between two items is the temporal identification difference of the shots that contain these two items.
6. The temporal distance threshold (TDT) specifies the upper bound that the temporal distance must comply with. This must not be larger than threshold.
7. Given a temporal distance threshold (TDT) the temporal support (TS) of an association is defined as the number of times this association appears sequentially in the sequence. In addition each time this association appears the temporal distance between any two neighbour items of the association should satisfy the given TDT (i.e., no more than T shots).

The maximum possible occurrences of the association are determined by the number of occurrences of the item with minimum support which leads to the confidence of the association. The larger the confidence value the more confident the association holds,

$$Conf\{X\}_{TDT-T} = TS\{X\}_{TDT-T} / Min(TS(X_1), .., TS(X_L)). \quad (1)$$

Actually, this distance is determined by the maximum number of sequentially matched items between the associations. The larger the number, the smaller their distance is.

$$SeqAssocD\{\{X\}^1, \{X\}^2\} = 1 - \frac{|LCS\{\{X\}^1, \{X\}^2\}|}{Min(P,Q)} \quad (2)$$

3.1 Human Activity Dataset

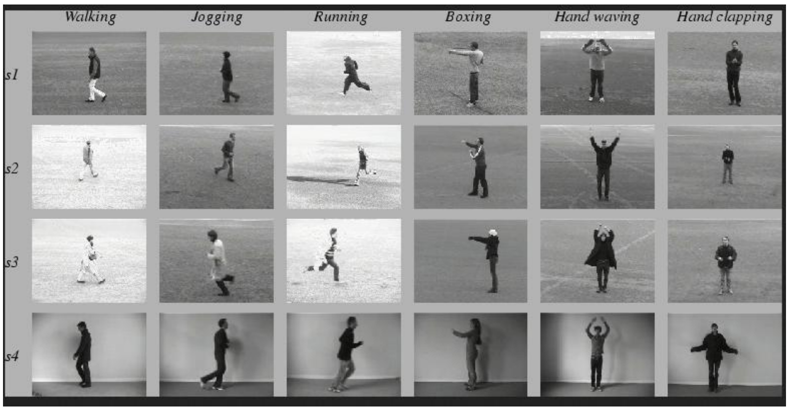


Fig 2. Human activity Dataset

Table 1: Scene_wise(Combined) hand_waving fixed at 4 [scaled 0 – 5]

	walking	jogging	running	boxing	hand clapping
s1 - s5	5	3	2	4	4

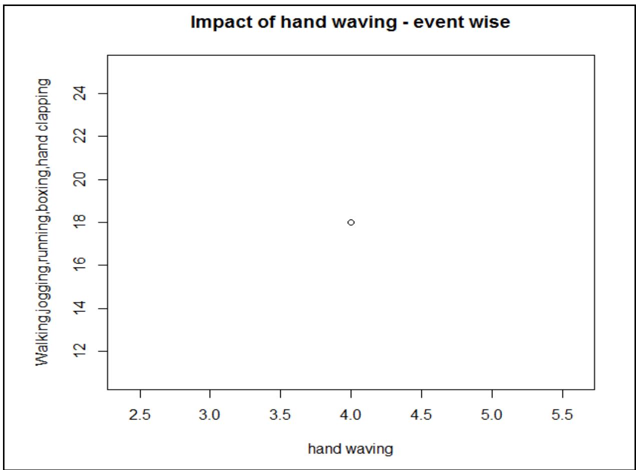
Table 2: Age_wise hand_waving fixed at 90% [scaled 0 – 100%]

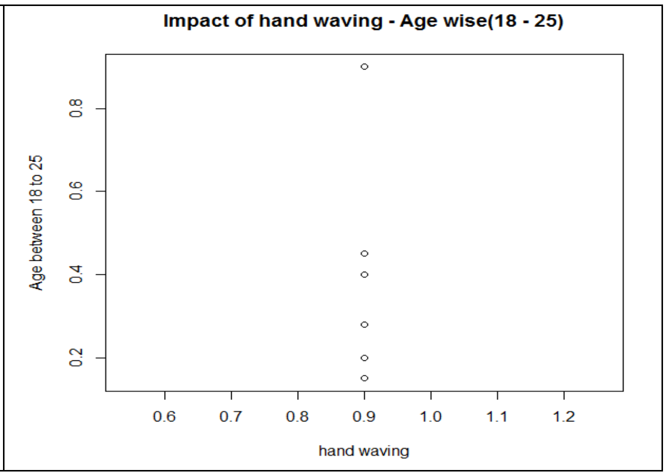
	18-25	26-35	36-45	46-60	above 60
walking	20%	26%	32%	38%	52%
jogging	28%	14%	6%	0%	0%
running	45%	34%	12%	0%	0%
boxing	15%	0%	0%	0%	0%
hand clapping	40%	28%	23%	8%	5%

Table 3: Session_wise hand_waving fixed at 1[scaled 0 – 1]

	morning	evening
walking	1	1
jogging	1	1
running	0	1
boxing	0	1
hand clapping	1	1

IV. RESULTS





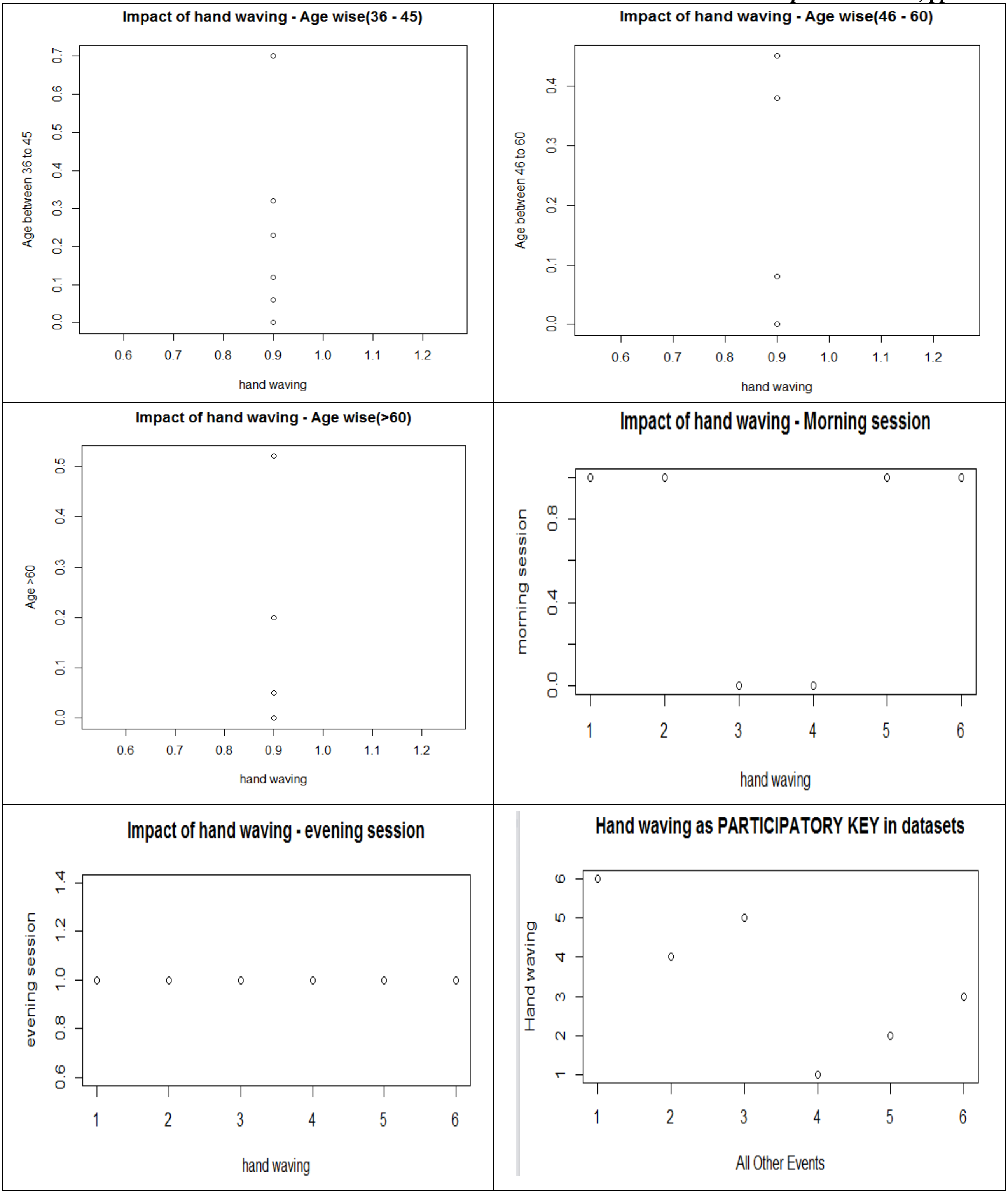


Fig 3: Experimental results based on participation of hand_waving in datasets (Tool Used - R Studio)

4.1 Improvement

An event participator occurs in majority of human activity datasets since the shot occurrences are sequential in nature.

We suggest to add a PARTICIPATOR_KEY with an upper threshold of 90% (Scales may vary according to nature of the video shots) in the datasets. In our example, *hand_waving* is assumed as the PARTICPATOR_KEY for the given datasets. Its value is fixed at 90% due to its active participation in all events like walking or jogging or boxing etc. We tend to look upon the charted values of *hand_waving* against event_wise, age_wise and session_wise. We can understand that in all the events or actions, *hand_waving* plays a major role. Hence it is assumed as a suitable PARTICIPATOR_KEY to derive the knowledge among other activities or parameters through predictions. Mathematically it can be derived as follows,

$$\text{PARTICIPATOR\_KEY}=\sum \text{seqAssocD} / \text{Max}(\text{TDT,TS}) \qquad (3)$$

V. CONCLUSION

Such a PARTICIPATOR_KEY exists in all human activity datasets. For instance a face recognition video dataset will consist of frequent, unique_event_dominated pattern in which a PARTICIPATOR_KEY can be defined. It could easily be derived to improve the support factor and indirectly the confidence part in the predictive mining algorithm that we have discussed and shown in equation (3).

REFERENCES

[1] Divakaan A, Peker K, Chang S, Radhakrishnan R, Xie L (2004) VideoMining: pattern discovery versus pattern recognition. In: Proceedings IEEE international conference on image processing (ICIP'2004). Mitsubishi Electric Research Laboratories.

[2] Oh J, Bandi B (2002) Multimedia data mining framework for raw video sequences. In: Proceedings of the third international workshop on multimedia data mining (MDM/KDD'2002) in conjunction with the eight ACM SIGKDD.

[3] Mongy S, Bouali F, Djeraba C (2005) Analyzing user's behavior on a video database. In: Proceedings of the 6th ACM international workshop on multimedia data mining: mining integrated media and complex data (MDM/KDD 2005), pp 95–100.

[4] SivaSelvan B, Gopalan NP (2007) Efficient algorithms for video association mining. In: Proceedings of the 20th conference of the Canadian Society for computational studies of intelligence on advances in artificial intelligence, vol 4509. Springer, Berlin, pp 250–260.

[5] Lin L, Ravitz G, Shyu M-L, Chen S-C (2007) Video semantic concept discovery using multimodal-based association classification.

[6] Chen M, Chen S-C, Shyu M-L (2007) Hierarchical temporal association mining for video event detection in video databases. In: The second IEEE international workshop on multimedia databases and data management (MDDM'07), in conjunction with IEEE international conference on data engineering (ICDE2007), Istanbul, Turkey.

[7] Goyanil M, Dutta S, Gohil G, Naik S (2011) Wicket fall concept mining from cricket video using a-priori algorithm. Proc Int J Multimedia Appl (IJMA).

[8] Maheshkumar HK, Palaniappan K, Sengupta S, Seetharaman G (2009) Semantic concept mining based on hierarchical event detection for soccer video indexing. J Multimedia 4(5):298–307 international conference on knowledge discovery & data mining, pp 1–10.

[9] Lin L, Shyu M-L (2010) Weighted association rule mining for video semantic detection. Int J Multimedia Data Eng Manag 1(1):37–54.

[10] Kea J, Zhana Y, Chenc X, Wanga M (2012) The retrieval of motion event by associations of temporal frequent pattern growth. Future Generation Computational Systems.

[11] Zhu X, Wu X, Elmagarmid A, Feng Z, Wu L (2005) Video data mining: semantic indexing and event detection from the association perspective. IEEE Trans Knowledge Data Engg 17(5):1–14.

[12] Su J-H, Huang Y-T, Tseng (2008) Efficient content-based video retrieval by mining temporal patterns. In: Proceedings of 9th international workshop on multimedia data mining associated with the ACM SIGKDD 2008, pp 36

