Dimensionality Reduction with Random Projection and Distance Space for Video Similarity Measurement: Application with Sports Video Classification

Prisana Mutchima and Parinya Sanguansat

Faculty of Information Technology, Rangsit University, Thailand

Summary
This paper proposes the video similarity measurement approach for sports video classification by dimensionality reduction with random projection (RP) and distance space. Most video data are huge files, which vary in terms of length and amount of data, resulting in time-consuming data processing; therefore, reducing the dimensionality of the data becomes a necessity. All frames of training videos are extracted by color histogram based method. After that, all features of videos are projected onto a low-dimensional subspace by RP for reducing the dimensionality of the data. Afterwards, the clustering technique is performed to provide the centroids of each cluster, called reference vectors. Distance from each reference vector in database to the observation sequence is distance space which is the new feature space. Finally, videos will be classified by term weighting and the nearest neighbor classifier. Accordingly, the proposed approach helps enhance feature dimension reduction, resulting in faster data processing. The experimental results show that the proposed approach outperforms the other approaches significantly in sports video similarity measurement.

Key words:
Video Similarity Measurement, Random Projection, Distance Space, Sports Video Classification, Term Weighting

1. Introduction

With the fast growth of video sources, efficient video classification and management becomes increasingly important. Video similarity (or difference) measurement is the key issue in video classification, one of the essential steps for content-based video retrieval (CBVR) system [1]. Furthermore, efficient measurement of the video similarity also plays crucial role in several multimedia information systems, owing to its wide applications in many areas such as advertising, news video broadcasting and personal video archive.

Many approaches have been attempted for measuring video similarity and video classification. Following the literature review, we found that one popular video representation technique is to represent each video sequence with frames [2]. Recently, the technique for measuring the video similarity based on the percentage of visually similar frames between the two sequences has been proposed in [1],[3],[4]. One commonly used technique for video similarity measurement is the Naïve Video Similarity (NVS) [1],[3]. This technique first finds the total number of frames from each video sequence which has at least one similar frame with the other sequence. Then, the ratio of these numbers will be computed to the total numbers of frames. After that, the threshold is used to compare the difference between the frames. The efficiency of such technique depends on the effective selection of the optimal frame similarity threshold. Practical implementation would be very difficult in identifying the optimal frame similarity threshold because it is unpredictable pattern and has to be manually determined, resulting in time-consuming data processing. Moreover, the optimal threshold will depend on the training set. If the training set change, some unknowns will not work to categorize. In [5], we used expected value to average distance of video frames instead of the threshold. Accordingly, we applied the L_1 metric to measure the distance in comparing the color histograms and averaged distance of video frames by expected value, i.e. harmonic mean, geometric mean, arithmetic mean and median. In addition, categorization was performed using the nearest neighbor classifier.

Most of the features in video categorization are based on the frequency of the features. In text categorization, the Bag-of- Words (BoW) model is a popular approach [6]. It is frequently used for document categorization in a collection of documents, where each document is represented by its word frequency [7]. But individual words have different significant; therefore, term weighting is applied to measure the importance of a word for content classification [8]. The weighting function can depend upon the term frequency and collection frequency of the features [9]. Terms are frequently referred in individual documents; appear to be useful as recall-enhancing devices. This suggests that term frequency (*tf*) factor be used as part of the term-weighting system measuring the frequency of occurrence of the terms in the document or query texts. But term frequency factors alone cannot ensure acceptable retrieval performance. Specifically, when the high frequency terms are not concentrated in a few particular documents but instead are prevalent in the

Manuscript received May 5, 2011.
Manuscript revised May 20, 2011.

whole collection, all documents tend to be retrieved, and this affects the search precision. Hence, a new collection-dependent factor must be introduced that favors terms concentrated in a few documents of a collection. The well-known inverse document frequency (*idf*) (or inverse collection frequency) factor performs this function. But the tf.idf weighting is not suitable for documents that have different lengths [10]. Therefore, each document feature vector is normalized to unit length by cosine normalization, called the *tfc* weighting [11]. In [12] uses the logarithm function to reduce difference of the word frequency that appear in the document, called the *ltc* weighting [10].

Regarding the image comparison, various features such as color [13] texture [14] and shape [15] are used in several approaches [16]. Among these features, color features are the most basic features, which are widely used and prove to be highly effective for image comparison [1],[3],[17],[18]. Therefore, in this paper, we focus on color features to compare similarity of low-level visual features of images that can reflect the perceptual similarity among images [19] and propose random projection and distance space to reduce the dimensionality of the datasets for video similarity measurement. The proposed approach can be briefly described: Firstly, all frames of training videos are extracted by color histogram based method, which directly captures the probability distribution of the color [20]. Secondly, all features of videos are projected onto a low-dimensional subspace using a random projection. Thirdly, the clustering technique is performed to provide the centroids of each cluster, called *reference vectors*. These vectors are used as a set of basis to create new space, called *distance space*. For any sequence in distance space, the new feature is represented by the frequencies of similar frame comparing with each reference vector. Finally, videos will be classified by term weighting and the nearest neighbor classifier.

The remainder of this paper is organized as follows. In Section 2, techniques for video similarity measurement are described. Techniques for dimensionality reduction are proposed in Section 3. Term weighting schemes are described in Section 4. In Section 5, experimental results in measuring video similarity are presented to demonstrate the performance of our proposed approach. Finally, conclusions are discussed in Section 6.

2. Video Similarity Measurement

2.1 Naïve Video Similarity

Naïve Video Similarity (NVS) is a traditional technique to measure video similarity by finding the total number of frames from each video sequence that has at least one visually similar frame with the other sequence, and then computing the ratio of this number to the overall total number of frames. Individual frames in a video are represented by high dimensional feature vectors from a metric space. In order to be robust against editing changes in temporal domain, a video *X* is defined as a finite set of feature vectors and ignores any temporal ordering. The metric *d(x,y)* measures the visual dissimilarity between frames *x* and *y* which are visually similar to each other if and only if $d(x,y) \le \varepsilon$ for an $\varepsilon > 0$ independent of x and *y*, where *ε* is the frame similarity threshold.

This method uses the L_1 metric to measure the distance. It is defined by the sum of the absolute difference between each bin of the two histograms. This method denotes the L_1 metric between two feature vectors x and *y* as *d(x,y)* as follows:

$$d(x,y) = \sum_{i=1}^{4} d_q(x_i, y_i) \tag{1}$$

where

$$d(x,y) = \sum_{i=1}^{4} \left\| x_i[j] - y_i[j] \right\| \tag{2}$$

where x_i and y_i for $i \in \{1, 2, 3, 4\}$ represent the quadrant color histograms from the two image feature vectors, *n* is the number of histogram bins and $\|\bullet\|$ is the L_1 metric. A small $d(\bullet,\bullet)$ value usually indicates visual similarity, except when two images share the same background color.

Let *X* and *Y* are two video sequences, represented as sets of feature vectors. The numbers of frames in video *X* that have at least one visually similar frame in *Y* is represented by $\Psi_{(X,Y;\varepsilon)}$, where 1_A is the indicator function with $1_A = 1$ if *A* is not empty, and zero otherwise that *x* and *y* are two video frames, represented as feature vectors and *ε* is the frame similarity threshold. The Naïve Video Similarity between *X* and *Y*, *nvs(X,Y ; ε)*, is defined as follows:

$$nvs(X,Y;\varepsilon) = \frac{\Psi_{(x,y;\varepsilon)} + \Psi_{(Y,X;\varepsilon)}}{|X| + |Y|} \tag{3}$$

where

$$\Psi_{(X,Y;\varepsilon)} = \sum_{x \in X} 1_{\{y \in Y : d(x,y) \le \varepsilon\}} \tag{4}$$

and

$$\Psi_{(Y,X;\varepsilon)} = \sum_{y \in Y} 1_{\{x \in X : d(y,x) \le \varepsilon\}} \tag{5}$$

where $|\bullet|$ denotes the cardinality of a set or the number of frames in a given video.

If every frame in video X has a similar match in Y and vice versa, $nvs(X, Y; \varepsilon) = 1$. If X and Y share no similar frames at all, $nvs(X, Y; \varepsilon) = 0$.

2.2 Expectation-based Measuring Video Similarity

This approach can measure similarity of video efficiently by using expected value to average distance of video frames instead of the threshold. Each video sequence was represented with frame and each frame was represented with the color histogram to help enhance feature reduction. After that, categorization was performed using the nearest neighbor classifier with the L_1 metric to measure distance by comparing each sampling frame of the training videos with all sampling frames of the test videos.

Let X and Y are two video sequences, represented as frames. The metric $d(x,y)$ measures the visual similarity between frames x and y. We denote the distance metric between two feature vectors x and y as $d(x, y)$, as follows:

$$d(x_a, y_b) \triangleq \sum_{i=1}^{4} \left\| x_i[a] - y_i[b] \right\| \tag{6}$$

where x_i and y_i for $i \in \{1, 2, 3, 4\}$ represent the quadrant color histograms from the two image feature vectors to merge spatial information into the image features. Spatial information describes the physical location of objects and the relationship between objects. Let a is the a^{th} sampling frame of X and b is the b^{th} sampling frame of Y.

The measuring video similarity between two video se quences X and Y, SIM, is defined as:

$$SIM(X,Y) \triangleq \min\{D(X,Y)\} \tag{7}$$

where

$$D(X,Y) \triangleq \mathrm{E}[d(x_a, y_b)] \tag{8}$$

where $\mathrm{E}[\bullet]$ is expectation operator. The similarity between two video sequences can be measured at various by changing the number of histogram bins and the expected value. The measuring video similarity is comparison minimum of average of frame distance measures.

3. Dimensionality Reduction

To reduce the dimensionality of the datasets, this study uses feature extraction method before random projection and distance space. Each video sequence was represented with frames, and each individual frame in the video was represented with the color histograms. Besides, to incorporate spatial information into the image features, the image was partitioned into four quadrants, with each quadrant having its own color histogram. After that, all features of videos are projected onto a low-dimensional subspace by RP for reducing the dimensionality of the data. Afterwards, the clustering technique is performed to provide the centroids of each cluster, called *reference vectors*. These vectors are used as a set of basis to create new space, called *distance space*. For any sequence in distance space, the new feature is represented by the frequencies of similar frame comparing with each reference vector. Finally, videos will be classified by term weighting and the nearest neighbor classifier. An overview of our approach is illustrated in Fig. 1.

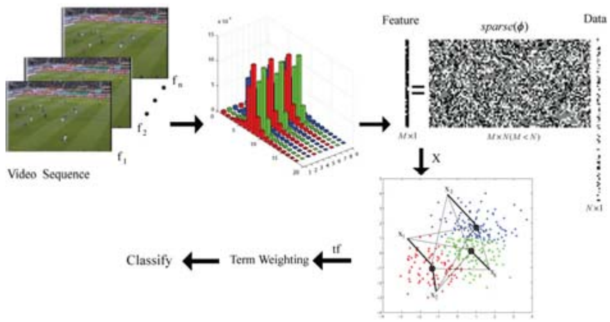

Fig. 1. Schematization of the proposed framework

3.1 Random Projection

Random projection (RP) has emerged as a powerful dimensionality reduction method. Its most important property is that it is a general data reduction method. In RP, the original high dimensional data is projected onto a low-dimensional subspace using a random matrix whose columns have unit length. In contrast to other methods, such as PCA, that compute a low-dimensional subspace by optimizing certain criteria (e.g., PCA finds a subspace that maximizes the variance in the data), RP does not use such criteria; therefore, it is data independent [21], [22],[23],[24].

In random projection, the set of points of size p in original q-dimensional Euclidean space is projected to a s-dimensional ($n \ll q$) subspace through the origin, using a random $q \times s$ matrix R whose columns have unit lengths in order to achieve dimension reduction as follows:

$$W_{p \times s} = F_{p \times q} R_{q \times s} \tag{9}$$

where $R_{q \times s}$ is the random matrix, $F_{p \times q}$ is the original obse rvations set of size p in q-dimension, and $W_{p \times s}$ is the proj ection in s-dimension subspace.

3.2 Distance Space

In this paper, we propose technique to create new feature space, called *distance space* which refers to distance from each reference vector in database to the observation sequence.

The distance space process is described in Algorithm 1. First, all frames of the training videos, X_i, are extracted by color histogram based method in Section 2.1 and projected by random matrix. After that, the clustering technique is performed to provide the centroids of each cluster, called *reference vectors*, ξ_k , using k-mean. Finally, the new feature vector, F_i, is represented by the frequencies of similar frame comparing with each reference vector.

Algorithm 1 Distance Space Algorithm

Require: $X_{i=1...N}$, $\xi_{k=1...C}$

Ensure: $G_{i=1...N}$

All frames of the training videos, X_i, are extracted by color histogram based method in Section 2.1 and projected by random matrix.
Perform clustering in this feature spaces to keep centriods of each cluster as reference vectors, ξ_k .
for i = 1 to N **do**
    $G_i \leftarrow 0$
    for j = 1 to $|X_i|$ **do**
        for k = 1 to C **do**
            $D_k \leftarrow \|X_i[j] - \xi_k\|$
        end for
        L $\leftarrow \arg\min_k(D_k)$.
        $G_i[L] \leftarrow G_i[L]+1$.
    end for
    $G_i \leftarrow G_i / |X_i|$.
end for
return $G_{i=1...N}$

---

4. Term Weighting Schemes

The perhaps most commonly used document representation is the so called vector space model [25]. In the vector space model, documents are represented by vectors of words. Usually, one has a collection of documents which is represented by a word-by-document matrix A, where each entry represents the occurrences of a word in a document [10] i.e.,

$$A = (w_{mr}) \tag{10}$$

where w_{mr} is the weight of term m in document r.

Let tf_{mr} be the frequency of term m in document r, N the number of documents in a document collection, H the number of words in the collection after stop word removal and word stemming, and n_m the number of documents containing term m.

1) *tf-weighting (Term Frequency)*: The simplest approach is to assign the weight to be equal to the number of occurrences of term m in document r. This weighting scheme is referred to as term frequency and is denoted tf_{mr} , with the subscripts denoting the term and the document in order [25].

$$w_{mr} = tf_{mr} \tag{11}$$

2) *tf.idf-weighting (Term Frequency - Inverse Document Frequency)*: The previous scheme does not take into account the frequency of the word throughout all documents in the collection. A well-known approach for computing word weights is the *tf.idf*-weighting [25] which assigns the weight to term m in document r in proportion to the number of occurrences of the term in the document, and in inverse proportion to the number of documents in the collection for which the term occurs at least once [10].

$$w_{mr} = tf_{mr} \times \log\left(\frac{N}{n_m}\right) \tag{12}$$

Observe that in this equation the value of w_{mr} decreases as n_m increases and vice versa. Since it combines both the two factors which are the distribution of a term within a certain document (term frequency tf_{mr}) and its distribution in a document collection (logarithm of the ratio of the number of documents to the number of documents containing the term) [24].

3) *tfc-weighting:* The *tf.idf* -weighting does not take into account that documents may be of different lengths. The *tfc* weighting is similar to the *tf.idf* -weighting except for the fact that length normalization is used as part of the word weighting formula [10].

$$w_{mr} = \frac{tf_{mr} \times \log\left(\frac{N}{n_m}\right)}{\sqrt{\sum_{e=1}^{H}\left[tf_{er} \times \log\left(\frac{N}{n_e}\right)\right]^2}} \quad (13)$$

Here H is the number of unique terms in a document vector space. Variable tf_{er} is defined as the frequency of term e in document m and ne is defined as the number of documents containing term e.

4) *ltc-weighting:* A slightly different approach uses the logarithm function to reduce difference of the word frequency that appears in the document [10].

$$w_{mr} = \frac{\log(tf_{mr}+1) \times \log\left(\frac{N}{n_m}\right)}{\sqrt{\sum_{e=1}^{H}\left[\log(tf_{er}+1) \times \log\left(\frac{N}{n_e}\right)\right]^2}} \quad (14)$$

5. Experimental Results

In this section, the experimental results were shown to demonstrate the performance of the video similarity measurement by dimensionality reduction with random projection and distance space. All experiments use the datasets from TV sports programs described in Section 5.1. The results of feature extraction were presented in Section 5.2. In Section 5.3, the video similarity measurement was applied for video classification. Finally, the results of performance measurement were described in Section 5.4.

5.1 Datasets

The data were obtained from 400 video sequences of TV sports program, comprising of 20 sport genres, namely basketball, beach volleyball, bicycle racing, bowl, bowling, boxing, car racing, football, hoggy, motorcycle racing, rugby, ski, snooker, squash, swimming, table tennis, tennis, volleyball, walkathon and wrestling. The datasets were divided into two groups, i.e. 200 training and 200 test video sequences. The number of frames of each video sequence is 30 frames per second in MPEG-2 format. The resolution of the datasets evaluation sequences is 480×720 pixels, and the length of each video is about 30 second.

5.2 Feature Extraction

To reduce the dimensionality of the datasets, this study uses feature extraction method. The original features were transformed into new sampling features. For image classification, the color histogram was widely used as an important color feature indicating the content of the image. Moreover, the advantage of using the color histogram is its robustness to affine transformation, especially rotation and scaling of the image content [16]. Therefore, in our experiments, each video sequence was represented with frames, and each individual frame in the video was represented with the color histograms. Besides, to incorporate spatial information into the image features, the image was partitioned into four quadrants, with each quadrant having its own color histogram.

5.3 Classification

The most straightforward and highly popular method to measure the similarity between two features is to compute the distance between them using a certain distance metric. In many fields such as content-based image retrieval, the sum of the absolute differences (L_1 metric), is widely used [27]. Accordingly, we applied the L_1 metric to measure the distance in comparing the reference vector with the observation sequence. In addition, classification was performed using the nearest neighbor classifier because it is one of the most common instance-based learning algorithms, and the simplest possible classification scheme.

5.4 Performance Measurement

For the performance, several criteria including number of sampling frames, number of histogram bins, number of reference vectors and feature dimension were set up to test in order to identify the accuracy rate of each criterion in video similarity measurement with random projection and distance space by running each criterion 10 times.

1) *Feature Dimension:* The experiment compared, the accuracy in using different the feature dimension as the criterion. The experiment set the number of sampling frames as 10, the number of histogram bins as 18 and the number of reference vectors as 40, while varying the number of feature dimension from 10 to 100. The results show in Table 1 and Fig. 2.

Table 1: Mean and standard deviation of number of feature dimension

Feature dimension	Mean (%)	S.D.
10	89.90	3.48
20	91.80	2.78
30	94.60	2.37
40	94.90	2.18
50	93.30	3.86
60	95.10	2.13
70	95.10	2.77
80	95.30	1.34
90	94.60	2.01
100	***96.50***	***1.08***

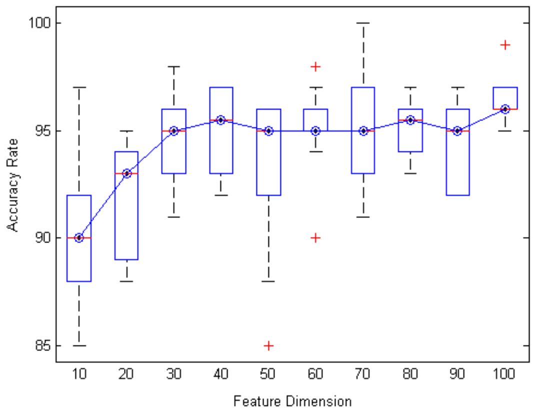

Fig. 2 Box and Whisker Diagram to Show the Spread of the Accuracy Rate of Number of Feature Dimension.

2) *Number of Sampling Frames:* The experiment compared, the accuracy in using different the number of sampling frames as the criterion. The experiment set the number of histogram bins as 18, the number of reference vectors as 40 and the feature dimension as 100, while varying the number of sampling frames from 10 to 20. The results show in Table 2 and Fig. 3.

Table 2: Mean and standard deviation of number of sampling frames

No. of sampling frames	Mean (%)	S.D.
10	96.50	1.08
11	95.10	2.23
12	95.30	1.77
13	95.80	1.03
14	96.40	1.51
15	94.40	1.84
16	96.10	0.99
17	96.90	1.52
18	96.10	2.51
19	96.80	1.40
20	96.60	1.17

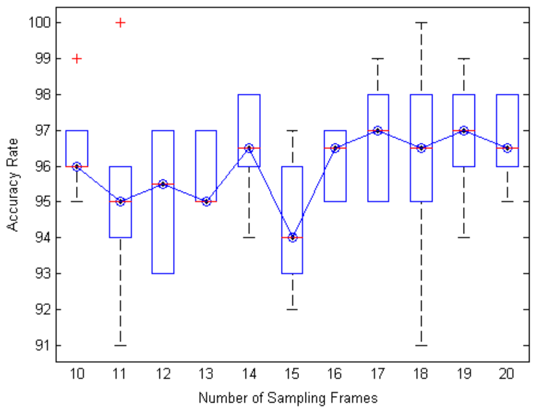

Fig. 3 Box and Whisker Diagram to Show the Spread of the Accuracy Rate of Number of Sampling Frames.

3) *Number of Histogram Bins:* The experiment compared, the accuracy in using different the number of histogram bins as the criterion. The experiment set the number of sampling frames as 10, the number of reference vectors as 40 and the feature dimension as 100, while varying the number of histogram bins from 10 to 20. The results show in Table 3 and Fig. 4.

Table 3: Mean and standard deviation of number of histogram bins

No. of histogram bins	Mean (%)	S.D.
10	96.00	1.41
11	93.90	1.60
12	94.00	2.36
13	96.70	1.57
14	95.40	1.35
15	97.10	0.32
16	95.70	1.57
17	95.50	2.55
18	96.50	1.08
19	97.30	0.92
20	96.40	1.64

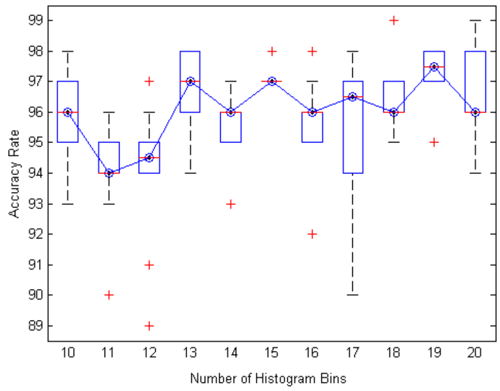

Fig. 4 Box and Whisker Diagram to Show the Spread of the Accuracy Rate of Number of Histogram Bins.

4) *Number of Reference Vectors:* The experiment compared, the accuracy in using different the number of reference vectors as the criterion. The experiment set the number of sampling frames as 10, the number of histogram bins as 18 and the feature dimension as 100, while varying the number of reference vectors from 10 to 80. The results show in Table 4 and Fig. 5.

Table 4: Mean and standard deviation of number of reference vectors

No. of reference vectors	Mean (%)	S.D.
10	87.10	4.61
20	94.90	2.28
30	95.80	2.04
40	96.50	1.08
50	95.90	1.73
60	96.00	1.33
70	94.50	1.34
80	94.40	1.90

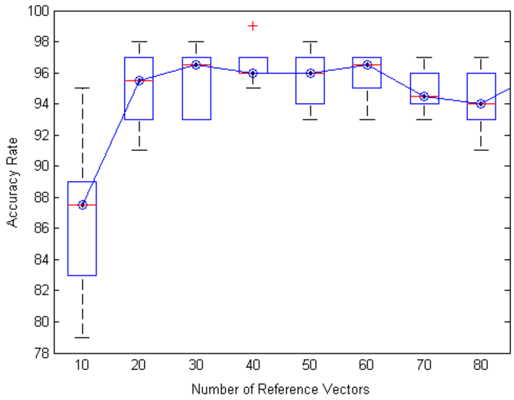

Fig. 5 Box and Whisker Diagram to Show the Spread of the Accuracy Rate of Number of Reference Vectors.

5) *Term weighting schemes comparison:* The experiment compared, accuracy rate of *tf*, *tf.idf*, *tfc* and *ltc*. The results show that the accuracy rate of the proposed method is highest, as shown in Table 5.

Table 5: Term weighting schemes comparison

Technique	Accuracy Rate (%)
tf	94.90
tf.idf	95.05
tfc	95.35
ltc	95.65

6) *Comparison with other methods:* The experiment compared, dimension and accuracy rate of NVS, expectation-based method and proposed method. The results show that dimension of the proposed method is very smaller than other methods while the accuracy rate is comparable, as shown in Table 6.

Table 6: Comparison the results of NVS, expectation-based method and proposed method

Technique	Dimension	Accuracy Rate (%)
NVS	288,000	95
Expectation-based method	288,000	97
Proposed method	100	95.65

6. Conclusions

The random projection and distance space is the efficient approach for video similarity measurement. Compared to the commonly-used Naïve Video Similarity (NVS) and expectation-based method, the approach proves to be more efficient in feature reduction and requires less data processing time, while still delivers acceptable accuracy rate. In using the random projection and distance space to measure video similarity, all frames of training videos were extracted by color histogram-based method. After that, all features of videos are projected onto a low-dimensional subspace using a *random projection* (RP). Then the clustering technique is performed to provide the centroids of each cluster, called reference vectors. Distance from each reference vector in database to the observation sequence is *distance space* which is the new feature space. Finally, videos will be classified by term weighting and the nearest neighbor classifier. Comparing to the commonly-used Naïve Video Similarity (NVS) and expectation-based method, the approaches can reduce more dimensionality of the dataset, resulting in efficient feature reduction and less data processing time. In measuring the similarity of videos, the NVS and expectation-based method took around 2,400 minutes to process the data, as each sampling frame of the testing videos had to be compared with all the sampling frames of the training videos for distance comparison, and the method also required as much as 288,000 dimensions for efficient data processing. In comparison, the random projection and distance space method took only 30 minutes to process the data (80 times less), and requires only 100 dimensions (2880 times less) to measure the similarity of the same dataset, with acceptable accuracy rate (95.65%). Therefore, the proposed approach can be used to measure video similarity efficiently and effectively, with its capability in feature reduction and data processing time outperforming the NVS and expectation-based method.

Acknowledgments

This study is supported by Rangsit Univerity, Suan Dusit Rajabhat University Foundation. Additionally, the invaluable recommendation and supervision from the anonymous reviewers are much appreciated.

References

[1] S. S. Cheung and A. Zakhor, "Efficient video similarity measurement with video signature," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 13, pp. 59-74, 2003.
[2] Z. Xiangmin, et al., "Efficient similarity search by summarization in large video database," presented at the Proceedings of the eighteenth conference on Australasian database - Volume 63, Ballarat, Victoria, Australia, 2007.
[3] S. S. Cheung and A. Zakhor, "Fast similarity search on video signatures," in Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, 2003, pp. II-1-4 vol.3.
[4] H. T. Shen, et al., "Towards effective indexing for very large video sequence database," presented at the SIGMOD '05: Proceedings of the 2005 ACM SIGMOD international conference on Management of data, Baltimore, Maryland, 2005.

[5] P. Mutchima and P. Sanguansat, "A Novel Approach for Measuring Video Similarity without Threshold and Its Application in Sports Video Categorization," in Pervasive Computing Signal Processing and Applications (PCSPA), 2010 First International Conference on, 2010, pp. 868-872.
[6] M. Sahlgren and R. Coster, "Using bag-of-concepts to improve the performance of support vector machines in text categorization," in Proceedings of the 20th international conference on Computational Linguistics, 2004.
[7] L. Ballan, et al., "Video event classification using string kernels," Multimedia Tools and Applications, vol. 48, pp. 69-87, 2010.
[8] A. Arif, M. M. Rahman, and S. Y. Mukta, "Information retrieval by modified term weighting method using random walk model with query term position ranking," in 2009 International Conference on Signal Processing Systems, 2009, pp. 526–530.
[9] D. Squire, et al., "Relevance Feedback and Term Weighting Schemes for Content-Based Image Retrieval," in Visual Information and Information Systems. vol. 1614, D. Huijsmans and A. Smeulders, Eds., ed: Springer Berlin / Heidelberg, 1999, pp. 656-656.
[10] K. Aas and L. Eikvil, "Text Categorisation: A Survey," ed, 1999.
[11] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," Information Processing & Management, vol. 24, pp. 513-523, 1988.
[12] C. Buckley, "Automatic Query Expansion Using SMART : TREC 3," in In Proceedings of The third Text REtrieval Conference (TREC-3), 1994, pp. 69--80.
[13] H. Nor Hazlyna, et al., "Comparison of acute leukemia Image segmentation using HSI and RGB color space," in Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on, 2010, pp. 749-752.
[14] A. Suruliandi and K. Ramar, "Local Texture Patterns - A Univariate Texture Model for Classification of Images," in Advanced Computing and Communications, 2008. ADCOM 2008. 16th International Conference on, 2008, pp. 32-39.
[15] L. Huitao, "Image-dependent shape coding and representation," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 15, pp. 345-354, 2005.
[16] W. Xiaoling and M. Hongyan, "Enhancing Color histogram for Image Retrieval," in Proceedings of the 2009 International Workshop on Information Security and Application (IWISA 2009), 2009, pp. 622-625.
[17] Z. Zhang, et al., "Novel color feature representation and matching technique for content-based image retrieval," in Multimedia Computing and Systems, 2009. ICMCS '09. International Conference on, 2009, pp. 118-122.
[18] A. M. Ferman, et al., "Robust color histogram descriptors for video segment retrieval and identification," Image Processing, IEEE Transactions on, vol. 11, pp. 497-508, 2002.
[19] W. Huang, et al., "A Review of Region-Based Image Retrieval," Journal of Signal Processing Systems, vol. 59, pp. 143-161, 2010.
[20] Z. Xiong, et al., A Unified Framework for Video Summarization, Browsing \& Retrieval: with Applications to Consumer and Surveillance Video. Orlando, FL, USA: Academic Press, Inc., 2005.
[21] N. Goela, et al., "Face Recognition Experiments with Random Projection," in Biometric Technology for Human Identification II, 2005, pp. 426-437.
[22] S. Deegalla and H. Bostrom, "Reducing High-Dimensional Data by Principal Component Analysis vs. Random Projection for Nearest Neighbor Classification," in Machine Learning and Applications, 2006. ICMLA '06. 5th International Conference on, 2006, pp. 245-250.
[23] W. Wu and J. Hu, "Similarity search based on random projection for high frequency time series," in Cybernetics and Intelligent Systems, 2008 IEEE Conference on, 2008, pp. 388-393.
[24] G. Li, et al., "A video retrieval algorithm using random projections," in Image Processing (ICIP), 2009 16th IEEE International Conference on, 2009, pp. 797-800.
[25] C. D. Manning, et al., Introduction to Information Retrieval: Cambridge University Press, 2008.
[26] J. Zhang and T. N. Nguyen, "A New Term Significance Weighting Approach," Journal of Intelligent Information Systems, vol. 24, pp. 61-85, 2005.
[27] J. Yu, et al., "A New Study on Distance Metrics as Similarity Measurement," in Multimedia and Expo, 2006 IEEE International Conference on, 2006, pp. 533-536.

Prisana Mutchima received the B.Sc. degrees in Environmental Science from the Thammasat University, Thailand in 1997, and the M.Sc. degrees in Technology of Information System Management from the Mahidol University, Thailand in 2000. She is currently pursuing the Ph.D. degree in the Department of Information Technology at Rangsit University, Thailand, since 2008. Her research areas are video retrieval and library science.

Parinya Sanguansat received the B.Eng., M.Eng. and Ph.D. degrees in Electrical Engineering from the Chulalongkorn University, Thailand in 2001, 2004 and 2007 respectively. He is an assistant professor in the Department of Information Technology at Rangsit University, Thailand. His research areas are digital signal processing in pattern recognition including on-line handwritten recognition, face and automatic target recognition.