Practical Elimination of Near-Duplicates from Web Video Search

Xiao Wu[+#]
wuxiao@cs.cityu.edu.hk

Alexander G. Hauptmann[#]
alex@cs.cmu.edu

Chong-Wah Ngo[+]
cwngo@cs.cityu.edu.hk

[+]Department of Computer Science
City University of Hong Kong
83 Tat Chee Avenue, Kowloon, Hong Kong

[#]School of Computer Science
Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, USA

ABSTRACT

Current web video search results rely exclusively on text keywords or user-supplied tags. A search on typical popular video often returns many duplicate and near-duplicate videos in the top results. This paper outlines ways to cluster and filter out the near-duplicate video using a hierarchical approach. Initial triage is performed using fast signatures derived from color histograms. Only when a video cannot be clearly classified as novel or near-duplicate using global signatures, we apply a more expensive local feature based near-duplicate detection which provides very accurate duplicate analysis through more costly computation. The results of 24 queries in a data set of 12,790 videos retrieved from Google, Yahoo! and YouTube show that this hierarchical approach can dramatically reduce redundant video displayed to the user in the top result set, at relatively small computational cost.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *Information filtering, Search process*; I.2.10 [**Artificial Intelligence**]: Vision and Scene Understanding – *Video analysis;*

General Terms

Algorithms, Design, Experimentation, Performance.

Keywords

Similarity Measure, Novelty and Redundancy Detection, Filtering, Multimodality, Near-Duplicates, Copy Detection, Web Video

1. INTRODUCTION

As bandwidth accessible to average users is increasing, video is becoming one of the fastest growing types of data on the Internet. Especially with the popularity of social media in Web 2.0, there has been exponential growth in videos available on the net. Users can obtain web videos easily, and distribute them again with some modifications. For example, users upload 65,000 new videos each day on video sharing website YouTube and the daily video views are over 100 million [29]. Among these huge volumes of videos, there exist large numbers of duplicate and near-duplicate videos. It becomes important to manage these videos in an automatic and efficient way. To avoid getting swamped by almost identical copies of the same video in any search, efficient near-duplicate video detection and elimination is essential for effective search, retrieval, and browsing.

Current web video search engines tend to provide a list of search results ranked according to their relevance scores given a text query. While some users' information needs may be satisfied with the relevant items ranked at the very top, the topmost search results usually contain a vast amount of redundant videos. Based on a sample of 24 popular queries from YouTube [34], Google Video [10] and Yahoo! Video [32] (see Table 1), on average there are 27% redundant videos that duplicate or nearly duplicate to the most popular version of a video in the search results. Figure 1 shows actual search results from three currently popular web video search engines, with redundancy fairly obvious in this case. As a consequence, users need to spend significant amount of time to find the videos they need and are subjected to repeatedly watching similar copies of videos which have been viewed previously. This process is extremely time-consuming particularly for web videos, where the users need to watch different versions of duplicate or near-duplicate videos streamed over the Internet. An ideal solution would be to return a list which not only maximizes precision with respect to the query, but also novelty (or diversity) of the query topic. This problem is generally referred to as novelty ranking (or sub-topic retrieval) in information retrieval (IR) [5, 36, 37]. Unfortunately, the text-based techniques from IR cannot be directly applied to discover video novelty. For instance, text keywords and user-supplied tags attached to web videos are usually abbreviated and imprecise. Second, most videos lack the web link structure typical in HTML documents which can be exploited for finding sub-topic relatedness. Finding novelty (or conversely, eliminating duplicates) among the relevant web videos must largely rely on the power of content analysis.

Due to the large variety of near-duplicate web videos ranging from simple formatting to complex editing, near-duplicate detection remains a challenging problem. Accurate detection generally comes at the cost of time complexity [20] particularly in a large video corpus. On the other hand, timely response to user queries is one important factor that fuels the popularity of Web 2.0. To balance the speed and the accuracy aspects, in this paper, we propose a hierarchical approach combining global signatures and local feature based pairwise comparison to detect near-duplicate web videos. The tool of near-duplicate detection can be

MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.
Copyright 2007 ACM 978-1-59593-701-8/07/0009...$5.00.

used in several ways: As a filter to remove redundant videos in the listing of retrieval results, as a tool for finding similar videos in different variations (e.g. to prevent copyright infringement), or as a way to discover the essential version of content appearing in different presentations. We show that the approach is practical for near-duplicate retrieval and novelty re-ranking of web videos where the majority of duplicates can be detected and removed from the top rankings.

The rest of this paper is organized as follows. In section 2 we give a brief overview of related work. A characterization of different types of near-duplicate web videos is provided in section 3. The proposed framework for efficient near-duplicate detection is introduced in section 4. Section 5 describes the data set used. Section 6 presents experiments and results for the two tasks a) web result novelty re-ranking and b) finding similar videos. Finally, we conclude the paper with a summary.

2. RELATED WORK

2.1 Novelty Detection and Re-Ranking

Novelty/redundancy detection has been explored in text information retrieval from the event level [4, 33] to the document/sentence level [3, 39]. It is closely related to the New Event Detection (NED) [4] or First Story Detection (FSD) in Topic Detection and Tracking (TDT) [2] that investigates several aspects for the automatic organization of news stories in text area. The NED task is to detect the first story that discusses a previously unknown event. A common solution to NED is to compare news stories to clusters of stories from previously identified events. The novelty detection approaches for documents and sentences mainly focus on vector space models and statistical language models to measure the degree of novelty expressed in words. The idea of novelty detection has also been applied to web search to improve the search results [36]. Query relevance and information novelty have been combined to re-rank the documents/pages by using Maximal Marginal Relevance [5], Affinity Graph [37] and language models [36]. However, these approaches are mainly based on textual information.

Recently, multimedia based novelty/redundancy detection has also been applied to cross-lingual news video similarity measure [30] and video re-ranking [13] by utilizing both textual and visual modalities. Hsu [13] used an information bottleneck method to re-rank video search results. For web videos, the textual information is usually limited and inaccurate. Therefore, applying text analysis to web videos makes little sense. To the best of our knowledge, there is little research on near-duplicate video detection and re-ranking for large scale web video search.

2.2 Video Copy and Similarity Detection

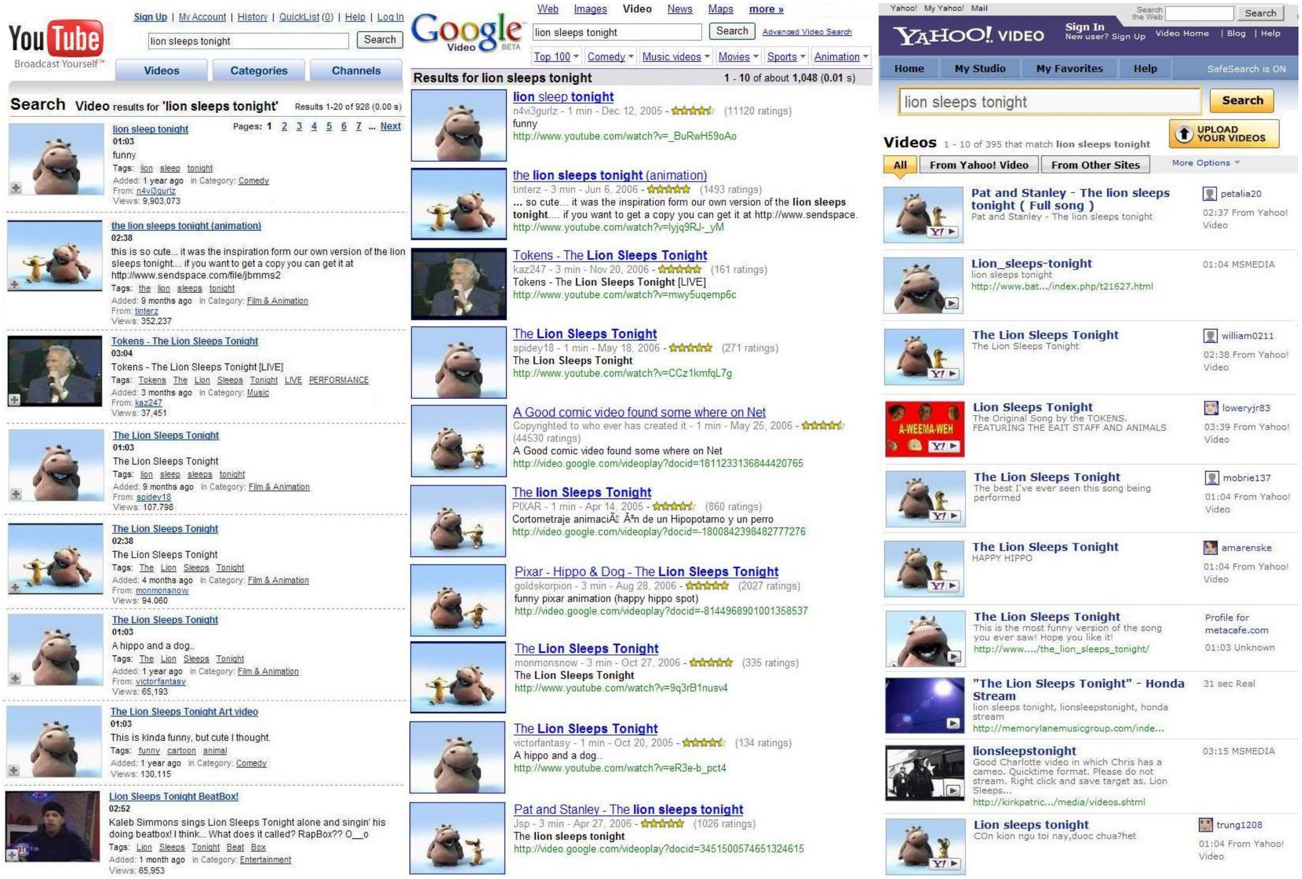
Video copy and similarity detection has been actively studied for its potential in search [6], topic tracking [31] and copyright protection [19]. Various approaches, using different features and matching algorithms have been proposed. Generally speaking, global features are suitable for identifying the majority of copies in formatting modifications such as coding and frame resolution changes [7, 8, 11, 12, 18, 35], while segment or shot-level features can detect some of copies with simple to moderate level of editing [35]. More sophisticated approaches normally involve the intensive use of feature matching at the image region level [20]. Thus an associated issue is the computation and scalability problem [17, 19, 20].

Table 1. 24 Video Queries Collected from YouTube, Google Video and Yahoo! Video (#: number of videos)

Queries			Near-Duplicate	
ID	**Query**	**#**	**#**	**%**
1	The lion sleeps tonight	792	334	42 %
2	Evolution of dance	483	122	25 %
3	Fold shirt	436	183	42 %
4	Cat massage	344	161	47 %
5	Ok go here it goes again	396	89	22 %
6	Urban ninja	771	45	6 %
7	Real life Simpsons	365	154	42 %
8	Free hugs	539	37	7 %
9	Where the hell is Matt	235	23	10 %
10	U2 and green day	297	52	18 %
11	Little superstar	377	59	16 %
12	Napoleon dynamite dance	881	146	17 %
13	I will survive Jesus	416	387	93 %
14	Ronaldinho ping pong	107	72	67 %
15	White and Nerdy	1771	696	39 %
16	Korean karaoke	205	20	10 %
17	Panic at the disco I write sins not tragedies	647	201	31 %
18	Bus uncle (巴士阿叔)	488	80	16 %
19	Sony Bravia	566	202	36 %
20	Changes Tupac	194	72	37 %
21	Afternoon delight	449	54	12 %
22	Numa Gary	422	32	8 %
23	Shakira hips don't lie	1322	234	18 %
24	India driving	287	26	9 %
Total		12790	3481	**27 %**

Among existing approaches, many emphasize the rapid identification of duplicate videos with global but compact and reliable features. These features are generally referred to as signatures or fingerprints which summarize the global statistic of low-level features. Typical features include color, motion and ordinal signature [11, 35] and prototype-based signature [7, 8, 22]. The matching between signatures is usually through bin-to-bin distance measures, probably with intelligent frame skipping [18, 35] and randomization [7, 8] so as to minimize the number of feature comparisons. These approaches are suitable for identifying almost identical videos, and can detect minor editing in the spatial and temporal domain. Another branch of approaches derive low-level features at the segment or shot level to facilitate local matching [1, 21, 23, 28]. Typically the granularity of the segment-level matching, the changes in temporal order, and the insertion/deletion of frames all contribute to the similarity score of videos. The emphasis of these approaches is mostly on variants of matching algorithms such as dynamic time warping [1], as well as maximal and optimal bipartite graph matching [28]. Compared to signature based methods, these approaches are slower but capable of retrieving approximate copies that have undergone a substantial degree of editing.

Duplicates with changes in background, color, and lighting, make serious demands for stable and reliable features at region-level details. Differing from global features, local features can be extracted after segmenting an image into regions and computing a set of color, texture and shape features for each region. A simpler approach merely segments the image into NxN blocks, and extracts features for each block. Promising approaches, which



Search Video results for 'lion sleeps tonight' Results 1-20 of 928 (0.00 s)

Pages: 1 2 3 4 5 6 7 ... Next

lion sleep tonight
01:03
funny
Tags: lion sleep tonight
Added: 1 year ago in Category: Comedy
From: n4vi3gurlz
Views: 9,903,073

the lion sleeps tonight (animation)
02:38
this is so cute... it was the inspiration form our own version of the lion sleeps tonight.... if you want to get a copy you can get it at http://www.sendspace.com/file/jbmms2
Tags: the lion sleeps tonight
Added: 9 months ago in Category: Film & Animation
From: tinterz
Views: 352,237

Tokens - The Lion Sleeps Tonight
03:04
Tokens - The Lion Sleeps Tonight [LIVE]
Tags: Tokens The Lion Sleeps Tonight LIVE PERFORMANCE
Added: 3 months ago in Category: Music
From: kaz247
Views: 37,451

The Lion Sleeps Tonight
01:03
The Lion Sleeps Tonight
Tags: lion sleep sleeps tonight
Added: 9 months ago in Category: Film & Animation
From: spidey18
Views: 107.798

The Lion Sleeps Tonight
02:38
The Lion Sleeps Tonight
Tags: The Lion Sleeps Tonight
Added: 4 months ago in Category: Film & Animation
From: monmonsnow
Views: 94.060

The Lion Sleeps Tonight
01:03
A hippo and a dog..
Tags: The Lion Sleeps Tonight
Added: 1 year ago in Category: Film & Animation
From: victorfantasy
Views: 65,193

The Lion Sleeps Tonight Art video
01:03
This is kinda funny, but cute I thought.
Tags: funny cartoon animal
Added: 1 year ago in Category: Comedy
Views: 130,115

Lion Sleeps Tonight BeatBox!
02:52
Kaleb Simmons sings Lion Sleeps Tonight alone and singin' his doing beatbox! I think... What does it called? RapBox?? O__o
Tags: Lion Sleeps Tonight Beat Box
Added: 1 month ago in Category: Entertainment
Views: 65,953

Results for lion sleeps tonight 1 - 10 of about 1,048 (0.01 s)

lion sleep tonight
n4vi3gurlz - 1 min - Dec 12, 2005 - ★★★★½ (11120 ratings)
funny
http://www.youtube.com/watch?v=_BuRwH59oAo

the lion sleeps tonight (animation)
tinterz - 3 min - Jun 6, 2006 - ★★★★★ (1493 ratings)
... so cute... it was the inspiration form our own version of the lion sleeps tonight.... if you want to get a copy you can get it at http://www.sendspace.
http://www.youtube.com/watch?v=lyjq9RJ-_yM

Tokens - The Lion Sleeps Tonight
kaz247 - 3 min - Nov 20, 2006 - ★★★★★ (161 ratings)
Tokens - The Lion Sleeps Tonight [LIVE]
http://www.youtube.com/watch?v=mwy5uqemp6c

The Lion Sleeps Tonight
spidey18 - 1 min - May 18, 2006 - ★★★★★ (271 ratings)
The Lion Sleeps Tonight
http://www.youtube.com/watch?v=CCz1kmfqL7g

A Good comic video found some where on Net
Copyrighted to who ever has created it - 1 min - May 25, 2006 - ★★★★½ (44530 ratings)
A Good comic video found some where on Net
http://video.google.com/videoplay?docid=1811233136844420765

The lion Sleeps Tonight
PIXAR - 1 min - Apr 14, 2005 - ★★★★½ (860 ratings)
Cortometraje animaciÃ Âªn de un Hipopotamo y un perro
http://video.google.com/videoplay?docid=-1800842398482777276

Pixar - Hippo & Dog - The Lion Sleeps Tonight
goldskorpion - 3 min - Aug 28, 2006 - ★★★★★ (2027 ratings)
funny pixar animation (happy hippo spot)
http://video.google.com/videoplay?docid=-8144968901001358537

The Lion Sleeps Tonight
monmonsnow - 3 min - Oct 27, 2006 - ★★★★★ (335 ratings)
The Lion Sleeps Tonight
http://www.youtube.com/watch?v=9q3rB1nusv4

The Lion Sleeps Tonight
victorfantasy - 1 min - Oct 20, 2005 - ★★★★½ (134 ratings)
A hippo and a dog..
http://www.youtube.com/watch?v=eR3e-b_pct4

Pat and Stanley - The lion sleeps tonight
Jsp - 3 min - Apr 27, 2006 - ★★★★★ (1026 ratings)
The lion sleeps tonight
http://video.google.com/videoplay?docid=3451500574651324615

Videos 1 - 10 of 395 that match **lion sleeps tonight**

Pat and Stanley - The lion sleeps tonight (Full song)
Pat and Stanley - The lion sleeps tonight
petalia20
02:37 From Yahoo! Video

Lion_sleeps-tonight
lion sleeps tonight
http://www.bat.../index.php/t21627.html
01:04 MSMEDIA

The Lion Sleeps Tonight
The Lion Sleeps Tonight
william0211
02:38 From Yahoo! Video

Lion Sleeps Tonight
The Original Song by the TOKENS. FEATURING THE EAIT STAFF AND ANIMALS
loweryjr83
03:39 From Yahoo! Video

The Lion Sleeps Tonight
The best I've ever seen this song being performed
mobrie137
01:04 From Yahoo! Video

The Lion Sleeps Tonight
HAPPY HIPPO
amarenske
01:04 From Yahoo! Video

The Lion Sleeps Tonight
This is the most funny version of the song you ever saw! Hope you like it!
http://www...../the_lion_sleeps_tonight/
Profile for metacafe.com
01:03 Unknown

"The Lion Sleeps Tonight" - Honda Stream
lion sleeps tonight, lionsleepstonight, honda stream
http://memorylanemusicgroup.com/inde...
31 sec Real

lionsleepstonight
Good Charlotte video in which Chris has a cameo. Quicktime format. Please do not stream. Right click and save target as. Lion Sleeps...
http://kirkpatric.../media/videos.shtml
03:15 MSMEDIA

Lion sleeps tonight
COn kion ngu toi nay,duoc chua?het
trung1208
01:04 From Yahoo! Video

Figure 1. Search results from different video search engines for the query "The lion sleeps tonight" demonstrate that there are a large number of near-duplicate videos in the topmost results.

have received a lot of attention recently, are to extract local feature points [15, 17, 19, 20, 27, 38]. These local points are salient local regions (e.g. corners) detected over images scales, which locate local regions that are tolerant to geometric and photometric variations [24]. While local points appear as promising features, a real challenge concerns the matching and scalability issues, since there simply exist too many local points for efficient, exhaustive comparison even between two frames. As a consequence, a major emphasis of these approaches is in exploring indexing structures [17, 19] and fast tracking with heuristics [27]. Most approaches indeed focus on keyframe-level duplicate detection [15, 27, 38]. Recent work in [20] shows how to perform video-level copy detection with a novel keypoint-against-trajectory search.

In web video search [8, 22], the duplicates can be of any variation from different formats to mixtures of complex modifications. Thus the right choice of features and matching algorithms cannot be pre-determined. This issue has not been seriously addressed, while the popularity of Web 2.0 has indeed made the problem timely and critical. In this paper, we explore a practical approach for near-duplicate web video filtering and retrieval.

3. NEAR-DUPLICATE WEB VIDEOS

3.1 Definition of Near-Duplicate Videos

Definition: *Near-duplicate web videos* are identical or approximately identical videos close to the exact duplicate of each other, but different in file formats, encoding parameters, photometric variations (color, lighting changes), editing operations (caption, logo and border insertion), different lengths, and certain modifications (frames add/remove). A user would clearly identify the videos as "essentially the same".

A video is a duplicate of another, if it looks the same, corresponds to approximately the same scene, and does not contain new and important information. Two videos do not have to be pixel identical to be considered duplicates − whether two videos are duplicates depends entirely on the type of differences between them and the purpose of the comparison. Copyright law might consider even a portion of a single frame within a full-length motion picture video as a duplicate, if that frame was copied and cropped from another video source. A user searching for entertaining video content on the web, might not care about individual frames, but the overall content and subjective impression when filtering near-duplicate videos for more effective search.

Exact duplicate videos are a special case of near-duplicate videos. In this paper, we include exact duplicates in our definition of near-duplicate videos, as these videos are also frequently returned by video search services.

3.2 Categories of Near-Duplicate Videos

To facilitate our further discussion, we classify near-duplicate web videos as the following categories:

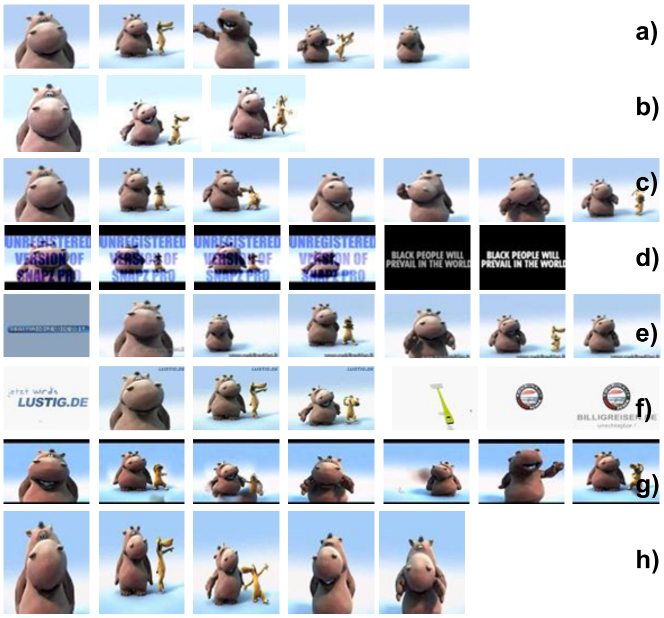


Figure 2. Keyframe sequence of near-duplicate videos with different variations (each row corresponds to one video). (a) is the standard version (b) brightness and resolution change (c) frame rate change (d) adding overlay text, borders and content modification at the end (e, f) content modification at beginning and end (g) longer version with borders (h) resolution differences

Formatting differences

- Encoding format: flv, wmv, avi, mpg, mp4, ram and so on.
- Frame rate: 15fps, 25fps, 29.97fps …
- Bit rate: 529kbps, 819kbps …
- Frame resolution: 174x144, 320x240, 240x320 …

Content differences

- Photometric variations: color change, lighting change.
- Editing: logo insertion, adding borders around frames, superposition of overlay text.
- Content modification: adding unrelated frames with different content at the beginning, end, or in the middle.
- Versions: same content in different lengths for different releases.

Furthermore, to avoid performing duplicate comparison on all frames, a video is usually viewed as a list of shots represented by representative keyframes, which will cause near-duplicate videos having different keyframe sequences. A web video is a sequence of consecutive frames to describe a meaningful scene. Commonly, a video is first partitioned into a set of shots based on editing cuts and transitions between frames, and then a representative keyframe is extracted to represent each shot. Extracting a representative keyframe from the middle of a shot therefore is relatively reliable for extracting basically similar keyframes from different near-duplicates. This mapping of video to keyframes reduces the number of frames that need to be analyzed by a factor of 100 - 5000 depending on the type of video. Although methods for detecting shots are overall quite robust for finding identical videos with the same format, when applied to near-duplicate videos with different frame rates, they could generate different keyframe sequences. It potentially induces the problem of viewpoint changes, zooming and so on, which causes the near-duplicate detection more complex.

Figure 2 shows examples of near-duplicate web videos for the query "The lion sleeps tonight" with simple scenes. We can see that the extracted keyframes are slightly different and near-duplicate variations. The overall scene is relatively simple because there are some common things throughout the videos (brown object and blue background). Figure 3 demonstrates another query "White and Nerdy" with complex scenes in which the content in the keyframes changes dramatically. Both simple and extensive changes are frequently mixed together to form more complicated transformations, making near-duplicate video detection a challenging problem.

4. HIERARCHICAL NEAR-DUPLICATE VIDEO DETECTION

In this section, we introduce the proposed hierarchical approach for near-duplicate web video detection. The framework combining global signatures and pairwise comparison is first presented in section 4.1, followed by the detailed description of global signatures with color histogram (SIG_CH) as a fast filter in section 4.2, and a more accurate but expensive local feature based pairwise comparison among keyframes (SET_NDK) in section 4.3. Finally, we summarize global signatures and pairwise comparison for near-duplicate video detection in section 4.4.

4.1 Hierarchical Framework

Our analysis of a diverse set of popular web videos shows that there are around 20% exact duplicate videos among all near-duplicate web videos. It is common for web users to upload exact duplicate videos with minimal change. This demands an approach for fast detection of duplicate videos. A global signature from color histograms (SIG_CH) is just this kind of fast measures suitable for matching videos with identical and almost identical content with only minor changes. The global signatures are basically the global statistics or summaries of low-level color features in videos. The similarity of videos is measured by the distance between signatures [35].

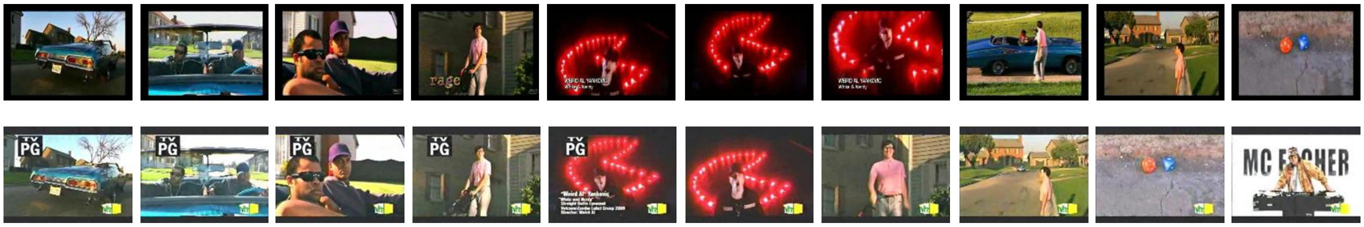


Figure 3. Two videos of complex scene query "White and Nerdy" with complex transformations (only the first ten keyframes are displayed): logo insertion, geometric and photometric variations (lighting change, black border), and keyframes added/removed

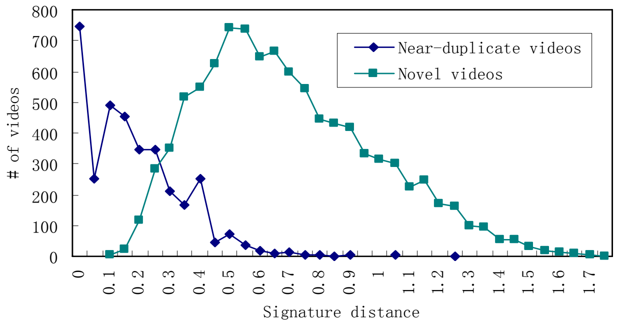


Figure 4. Signature distance distribution of near-duplicate and novel videos

However, for videos with major editing, content modification, dramatic photometric and geometric transformations, global signatures tend to be inadequate. Especially, when multiple variations are mixed together, the near-duplicate detection becomes even harder. Furthermore, due to different frame rates, and content modifications such as the insertion of commercials or title frames at the beginning and credits at the end, the extracted keyframe sequence could be different. And even non-duplicate videos could have similar color distribution as duplicate videos, which will be falsely detected as similar videos. In contrast to global signatures, pairwise keyframe comparison treats each keyframe as an independent node and two videos are compared by measuring the pairwise similarity among these nodes. Local feature based methods can accurately capture the mapping among keypoints. Pairwise comparison among keyframes can further measure the degree of overlapping between two videos. Therefore local feature based pairwise comparison (SET_NDK) has great potential in detecting near-duplicate keyframes and ultimately providing a reliable measurement for videos that have been non-trivially modified. However, the computation of local points is more expensive than mere color histograms, and the keyframes have to be compared pairwise.

To guarantee effective near-duplicate detection while meeting the speed requirements for Google-scale video collections, we propose a hierarchical method which utilizes both global signatures and local keypoints for detecting near-duplicate web videos. A global signature from color histograms is first used to detect the near-duplicate videos with high confidence and filter out very dissimilar videos. Figure 4 shows the signature distance distributions of near-duplicate and novel videos from our test set. Some videos can be directly identified as near-duplicate videos, for example, the ones with distance less than 0.2. While other videos with large distance can safely be labeled as novel ones, for example, those with distance greater than 0.7. With this filtering, a large portion of videos can be successfully identified, which reduces the computation for more expensive pairwise comparison. For videos that cannot be clearly classified as either novel or near-duplicate using global signatures (at distances between 0.2 and 0.7), we apply local feature based near-duplicate detection which provides very accurate duplicate analysis, at higher cost. The combination of global signature and pairwise comparison can balance performance and cost.

4.2 Global Signature on Color Histograms

A color histogram is calculated for each keyframe of the video, which is represented as: $H_i = (h_1, h_2, \ldots, h_m)$. As a typical feature here, we use the HSV color space. A histogram is concatenated with 18 bins for Hue, 3 bins for Saturation, and 3 bins for Value, hence $m = 24$.

A *video signature* (*VS)* is defined as an m-dimensional vector of a normalized color histogram over all keyframes in the video.

$$VS = (s_1, s_2 \cdots s_m), \qquad where \qquad s_i = \frac{1}{n}\sum_{j=1}^{n} h_{ij}$$

where n is the number of keyframes in the video, and h_{ij} is the *ith* bin of the color histogram at keyframe *j*.

We compute the distance of two signatures VS_i and VS_j based on the *Euclidean* distance:

$$\boldsymbol{R(V_i \mid V_j) = d(VS_i, VS_j) = \sqrt{\sum_{k=1}^{m}(x_k - y_k)^2}}$$

where $VS_i = (x_1, \ldots, y_m)$, and $VS_j = (y_1, \ldots, y_m)$. Two videos are regarded as near-duplicate if their distance is considered close.

The signatures of videos can be indexed and then searched without accessing the original videos. So the retrieval speed is rather fast with efficient mechanisms available for searching distance between moderately sized feature vectors [8].

4.3 Pairwise Comparison among Keyframes

For web videos that cannot be determined novel or near-duplicate using global signature, local features based method (SET_NDK) is used to measure the similarity of keyframes by pairwise comparison of keyframes from two videos, and then the redundancy of these two videos can be determined by comparing the ratio of the number of similar keyframes. In this section we will first introduce the local feature based technique to detect the near-duplicate keyframes (NDK) in videos with a sliding window, followed by the measure (set difference) of video redundancy with the information of keyframe similarity.

4.3.1 Near-duplicate Keyframe Detection with Local Features

In contrast to global features, features derived from local points can recognize various transformations from editing, viewpoint, and photometric changes. Salient regions in each keyframe can be extracted with local point detectors (e.g. DOG [24], Hessian-Affine [26]) and their descriptors (e.g., SIFT [25]) are mostly invariant to local transformations. Keypoint based local feature detection approach avoids the shortcoming of global features and therefore is particularly suitable for detecting near-duplicate web videos having complex variations.

To detect near-duplicate keyframes, the local points of each keyframe were located by Hessian-Affine detector [26]. The local points were then described by PCA-SIFT [19], which is a 36 dimensional vector for each local point. With a fast indexing structure, local points were matched based on a point-to-point symmetric matching scheme [27]. In our experiments, we will treat two keyframes as similar if the number of local point matching pairs between two keyframes is above a certain threshold.

4.3.2 Keyframe Matching Window

To find all near-duplicate/similar keyframes in two videos, the traditional method is to exhaustively compare each keyframe pair, in which the time complexity is the production of the numbers of keyframes in two videos. When videos consist of a large number

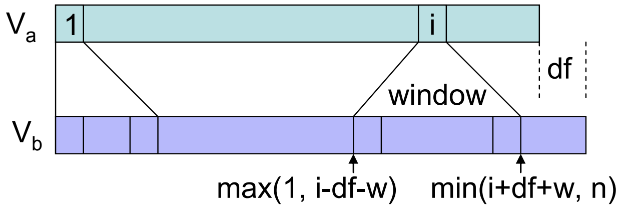


Figure 5. Matching window for keyframes between two videos

of keyframes, it is expensive and not feasible for large scale web video collections.

To reduce the computation, each keyframe was only compared to the corresponding keyframes in another video within a certain sliding window. For near-duplicate web videos, there exists certain mapping among keyframes. For example, the corresponding near-duplicate keyframes of one video in Figure 3 are within a certain distance in another video. To avoid unnecessary comparison and guarantee minimal miss detection, we utilize a sliding window policy to effectively reduce the computation. For the *ith* keyframe in one video, it is only compared with the keyframes of another video within the following range:

$$Range = [max(1, i - df - w), min(i + df + w, n)]$$

where *n* is the length of another video, i.e. the number of keyframes, *df* is the length difference between two videos, *w* is the window size. In our experiments, the window size *w* is fixed to 5. Figure 5 gives an example of matching window between two videos. The whole near-duplicate keyframe list is generated by transitive closure based on the information of each two keyframes, which forms a set of NDK groups [27].

This scheme is especially useful for complex scene videos with a large number of keyframes, such as queries 15, 17 and 23 in Table 1. These videos are represented by as many as 100 keyframes, where this scheme can greatly diminish the number of necessary comparisons. Although the sliding window scheme might miss part of near-duplicate keyframes for a single keyframe in videos of simple scenes, these missed near-duplicate keyframes will be eventually included by transitive closure considering the fact that keyframes for simple scene videos are usually very similar.

4.3.3 Set Difference of Keyframes

Once the similar keyframes have been identified, we use normalized set difference as the metric to evaluate the similarity between two videos. The set difference measure represents each video as a set of keyframes, either near-duplicate keyframes (NDK) or non-near-duplicate keyframes (non-NDK). It calculates the ratio of the number of duplicate keyframes to the total number of keyframe in a video.

It is measured by the following formulation:

$$\boldsymbol{R}(\boldsymbol{V}_i \mid \boldsymbol{V}_j) = \left(\frac{|\boldsymbol{KF}_i \cap \boldsymbol{KF}_j|}{|\boldsymbol{KF}_i|} + \frac{|\boldsymbol{KF}_i \cap \boldsymbol{KF}_j|}{|\boldsymbol{KF}_j|}\right)/2$$

KF_i is the set of keyframes contained in video V_i. This measure counts the ratio of intersected near-duplicate keyframes. The higher the rate, the more redundant the video.

Table 2. Comparison of Near-Duplicate Detection Capability for Global Color Histogram Signatures (SIG_CH) and Pairwise Comparison among Keyframes (SET_NDK)

Typical Near-Duplicate Categories		Freq %	SIG_ CH	SET_ NDK
Exactly duplicate		20%	√	√
Photometric variations		20%	X	√
Editing (inserting logo, text)		15%	P	√
Resolution		2%	√	√
Border (Zoom)		8%	P	√
Simple scene	Content modification	20%	X	P
	Different lengths	10%	√	√
Complex scene	Content modification	25%	P	√
	Different lengths	5%	X	√
Other		5%	X	P

√: able to detect X: unable to detect P: partially able to detect

4.4 Signature vs. Pairwise Comparison

The categories of web video variations and the capability of global signature based on color histograms (SIG_CH) and local feature based pairwise comparison of keyframes (SET_NDK) are listed in Table 2. The table categorizes different types of near-duplicates, and provides estimates of how frequently this category appeared in our web video test collection of 12,790 videos (Freq %). It also identifies which of the two approaches, SIG_CH and SET_NDK, is suitable for each type of near-duplicate detection.

The color histograms based global signature is able to detect duplicate and near-duplicate videos with certain minor variations (e.g. small logo insertion). Furthermore, the detection capability for simple scenes and complex scenes is different. For the simple scene video like "The lion sleeps tonight" in Figure 2, the key aspect (theme) of the extracted keyframes is a brown lion with a blue background. Dropping/inserting a couple of similar keyframes will not seriously affect the color distribution. A global signature using color histograms potentially can detect certain kinds of near-duplicate videos. But for complex scenes, such as "White and Nerdy" in Figure 3, the insertion and removal of keyframes will cause extensive changes in the global color signatures. The global signature is unable to recognize near-duplicates with different lengths because the color and ordinal distributions have changed quite dramatically. Generally, computing global signatures is fast, but their potential to detect the near-duplicate videos is limited.

On the other hand, local points are effective for finding duplicates with photometric and geometric variations, complex editing and zooming. Moreover, the local mapping among keyframes is especially suitable for detecting duplicate videos with different versions, insertion/deletion keyframes and various keyframe sequences caused by shot boundary detection algorithms. However, the matching process is naturally slow due to the large numbers of keypoints and the high dimensionality of the keypoint descriptors. Typically there are hundreds to thousands of keypoints identified in one keyframe. Although fast indexing structure (e.g. LSH [19], LIP-IS [27]) can filter out comparison among feature points and the matching window strategy reduces the comparison among keyframes, the matching (nearest neighbor search) is computationally expensive and not scalable to very large video databases.

The hierarchical approach combing the global signature and pairwise comparison is a reasonable solution to provide effective and efficient near-duplicate web video detection. Even though our experiments were done with one specific set of global features and local point descriptors, the basic principles of the approach, and its cost/effectiveness analysis, would easily apply to other sets of global features and other spatial or local point descriptors.

5. DATASET

To test our approach, we selected 24 queries designed to retrieve the most viewed and top favorite videos from YouTube. Each text query was issued to YouTube, Google Video, and Yahoo! Video respectively and we collected all retrieved videos as our dataset. The videos were collected in November, 2006. Videos with time duration over 10 minutes were removed from the dataset since they were usually documentaries or TV programs retrieved from Google, and were only minimally related to the queries. The final data set consists of 12,790 videos. Tables 3 and 4 summarize the formats and sources of web videos respectively. The query information and the number of near-duplicates to the dominant version (the video most frequently appearing in the results) are listed in Table 1. For example, there are 1,771 videos in query 15 "White and Nerdy", and among them there are 696 near-duplicates of the most common version in the result lists. Shot boundaries were detected using tools from CMU [14] and each shot was represented by a keyframe. In total there are 398,015 keyframes in the set.

To analyze the performance of the novelty re-ranking and near-duplicate video retrieval, two non-expert assessors were asked to watch videos one query at a time. The videos were ordered according to the sequence returned by the video search engines. For near-duplicate video retrieval, the most popular video was selected as the seed video for each query. The assessors were requested to label the videos with a judgment (redundant or novel) and to form the ground truth. To evaluate the re-ranking results, the assessors were also requested to identify the near-duplicate clusters in an incremental way and the final ranking list was formed based on the original relevance ranking after removing near-duplicate videos.

5.1 Performance Metric

To evaluate the performance, we use measures: *precision* and *recall*, and *novelty mean average precision (NMAP)*. The former measure is to assess the performance of near-duplicate detection, while the latter measures the ability to re-rank relevant web videos according to their novelty. Let $\boldsymbol{G}$ be the ground truth set of redundant videos and $\boldsymbol{D}$ be the detected one.

$$\boldsymbol{Recall} = |\boldsymbol{G} \cap \boldsymbol{D}| / |\boldsymbol{G}| \qquad \boldsymbol{Precision} = |\boldsymbol{G} \cap \boldsymbol{D}| / |\boldsymbol{D}|$$

The *novelty mean average precision (NMAP)* measures the mean average precision of all tested queries, considering only novel and relevant videos as the ground truth set. In other words, if two videos are relevant to a query but near-duplicate to each other, only the first video is considered as a correct match. For a given query, there are total of N videos in the collection that are relevant to the query. Assume that the system only retrieves the top k candidate novel videos where r_i is the number of novel videos seen so far from rank 1 to i. The NMAP is computed as:

$$\boldsymbol{NMAP} = (\textstyle\sum_{i=1}^{k} \boldsymbol{i} / \boldsymbol{r_i}) / \boldsymbol{N}$$

Table 3. Video Format Information

Formats	No. Videos	Percentage
FLV	10925	85.4 %
MPG	45	0.3 %
AVI	1714	13.4 %
WMV	98	0.7 %
MP4	8	0.1 %

Table 4. Video Source Information

Sources	YouTube	Google	Yahoo!
No. videos	10720	1428	642
Percentage	83.8 %	11.2 %	5 %
Total	12790		

6. EXPERIMENTS

In this paper, we discuss two experimental tasks: search result novelty re-ranking and near-duplicate web video retrieval. Search result novelty re-ranking aims to provide novel videos based on relevance ranking by eliminating all near-duplicate videos. Near-duplicate web video retrieval seeks to find all videos that are near-duplicates to a query (seed) video. Potentially the first scenario is a more challenging task since the number of possible near-duplicate videos increases quadratically.

6.1 Task 1: Novelty Re-Ranking

The objective of search results novelty re-ranking is to list all the novel videos while maintaining the relevance order. To combine query relevance and novelty, each video V_i is computed through a pairwise comparison between V_i and every previously ranked novel video V_j, which is calculated by:

$$\boldsymbol{R}(\boldsymbol{V_i} \mid \boldsymbol{V}_1, ..., \boldsymbol{V}_{i-1}) = \max_{1 \le j \le i-1} \boldsymbol{R}(\boldsymbol{V_i} \mid \boldsymbol{V_j})$$

The precede ranked video that most similar to V_i determines the redundancy of V_i. The ranked list after removing all near-duplicate videos will be presented to the user.

To evaluate the performance of novelty re-ranking, we compared the re-ranking results based on time duration, global signatures and the hierarchical method. The original ranking from the search engine acts as the baseline. Given the intuition that duplicate videos usually have similar time durations, the re-ranking based on time duration was also tested. In addition to the most popular version in the results, there are other subordinate versions different from the dominant one. Figure 6 illustrates the time duration distribution of videos in query "Sony Bravia", which potentially indicates a couple of subsidiary versions (e.g. version of 147 second) in the results differing from the most popular one (version of 70 second). If the time difference between two videos is within an interval (e.g. 3 seconds), they will be treated as redundant. Similarly, two videos were regarded as duplicate when their signature difference is close enough (e.g. less than 0.15). In this experiment, we tested different intervals (e.g. 0, 3, 5 seconds) and signature thresholds (e.g. 0.15, 0.2, 0.3), and the one with the best performance is reported.

Usually, the top search results receive the most attention for users. The performance comparison up to top 30 search results is illustrated in Figure 7 and the average NMAP over all top k levels is listed in Table 5. It is obvious that the performance for original search results is not good because duplicate videos are commonly appeared in the top list. The time duration information can distinguish novel videos at the beginning, however different web videos could have the same duration, especially for videos queries accompanied with background music or music videos, e.g. queries

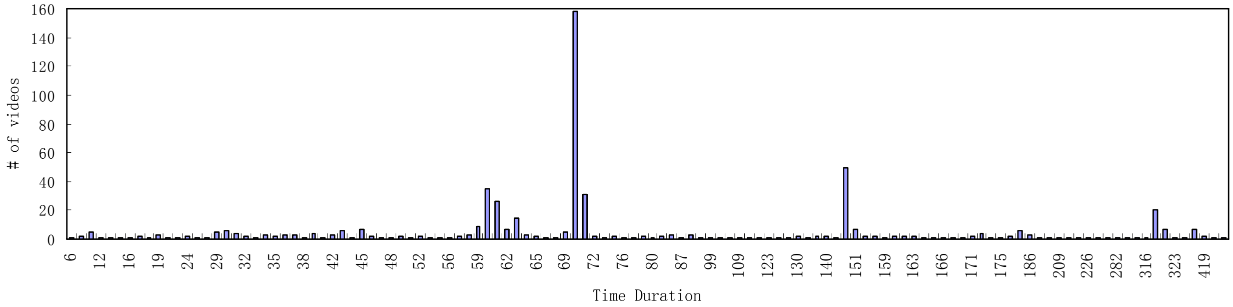


Figure 6. The time duration distribution for the query "Sony Bravia" (query 19) indicates that there might be multiple sets of duplicate videos different from the most popular video in the search results

1, 10, 23. As the number of videos increases, the information of time duration is inadequate, therefore the performance drops a lot. Although the global signature method can identify duplicate videos to some extent, the ability for duplicate videos is limited. A lot of near-duplicate videos cannot be correctly detected. Therefore the re-ranking list still consists of some duplicate videos and some novel videos were falsely removed. Overall, our hierarchical method effectively eliminates duplicate videos, which improves the diversity in the search results. So it achieves a good and stable performance across all top k levels.

Table 5. Overall Novelty Re-Ranking Performance

Solutions	Average NMAP
Original Ranking	0.76
Re-Ranking by Time Duration	0.74
Re-Ranking by Global Signature	0.84
Re-Ranking by Hierarchical Method	0.94

As search engines demands for quick response, the computation time is an important factor for consideration. The average number of keyframe pair comparison for top k re-ranking over 24 queries is listed in Table 6. Compared to fast re-ranking with global signatures and time duration, the hierarchical method is more expensive. However, using the global signature filtering and the sliding window, the hierarchical method has greatly reduced the computation compared to the exhaustive comparison among keyframes, which makes the novelty re-ranking feasible. Depending on the complexity of keyframes, the time for keyframe pair comparison ranges from 0.01 to 0.1 second for a Pentium-4 machine with 3.4G Hz CPU and 1G main memory. The average time to re-rank the top-10 results is around a couple of minutes. With the fast development of computer and parallel processing, especially for platform like Google parallel architecture, it is not a problem to response the queries quickly with our hierarchical approach.

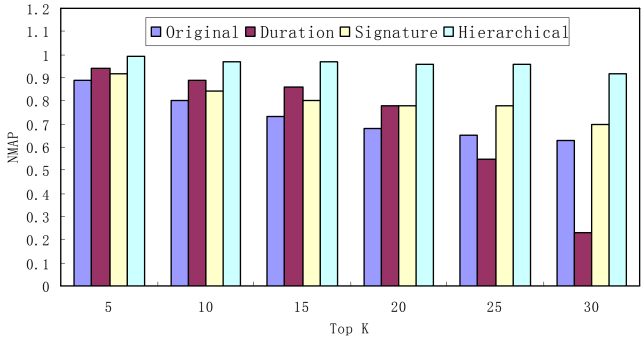


Figure 7. Performance comparison of novelty re-ranking

Table 6. Average number of keyframe pair comparison for top k ranking over all queries with the hierarchical method

Top k	5	10	15	20	25	30
Pairs	2858	8305	12944	19633	29765	43730

6.2 Task 2: Near-Duplicate Video Retrieval

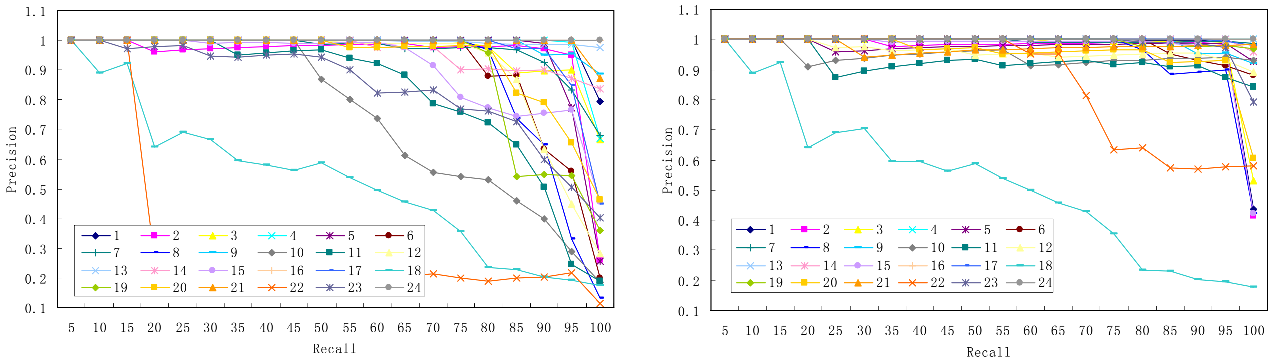
In addition to the novelty re-ranking, the users can also retrieve all videos that are near-duplicate to a query video. Given a seed (query) video V_s, all relevant videos are compared with the seed video to see if they are near-duplicates. It is computed by:

$$\boldsymbol{R}(\boldsymbol{V}_i) = \boldsymbol{R}(\boldsymbol{V}_i \mid \boldsymbol{V}_s)$$

Here, the redundancy measure is based on the proposed hierarchical method that combines the global signature and pairwise measure. The videos having small signature distance are directly labeled as near-duplicate while the dissimilar ones are filtered out as novel videos. For the uncertain videos, local features are further used to measure the redundancy of videos.

In this task, we retrieve the most popular video in each query. The seed (query) video can be determined automatically or manually according to the time duration distribution of the videos in the rank list, the relevance ranking and the global signature. The popular video in the top most list with the dominant time duration was picked as the seed video, and other videos were compared with it to see if they are near-duplicate to it.

The detailed and general performance comparison for near-duplicate retrieval is shown in Figure 8 and 9 respectively. As seen from Figure 8(a), global signature on color histogram (SIG_CH) achieves good performance for queries with simple scene or complex scene with minor editing and variations, e.g. queries 13 and 24. These near-duplicate videos have minor changes, so signature alone can detect most of the near-duplicate videos and filter out dissimilar videos. But for queries with complex scene (e.g. queries 10, 15, 22, 23), the signature based method is insufficient. Dissimilar videos can have similar color distribution to the seed video. Especially in videos with major variations, and insertion/removal of keyframes, this will cause remarkable difference of color distributions. However, the pairwise comparison method based on local features can effectively identify the near-duplicate keyframe mapping and eliminate the dissimilar videos with similar color signatures. Compared to Figure 8(a), the precision-recall curves using hierarchical method (HIRACH, Figure 8(b)) has prominent improvement. Most of the queries have high precision, especially at high recall levels. The pairwise comparison is especially useful for queries of complex scenes (e.g. queries 10, 15, 22, 23). The



(a) SIG_CH **(b) HIRACH**

Figure 8. Performance of near-duplicate video retrieval

queries having relatively low precision and recall by HIRACH are queries 18 and 22. For query 18 ("Bus uncle"), it was originally captured by a cell phone in the bus, so the scene is a little vague and the quality is bad. Furthermore, near-duplicate videos are undergone extensive editing and content modification (e.g. overlay text, frame insertion), while the query video clip consists of only two keyframes, which makes this detection a difficult task. So the precision and recall are low. For query 22 ("Numa Gary"), a lot of unrelated frames were inserted at the beginning and end for some near-duplicate videos, which induces low similarity scores. Therefore, the performance of query 22 is not good enough at high recall. Overall, the hierarchical method achieves satisfactory results.

Figure 9 demonstrates the average precision over 24 queries. It is easy to see that HIRACH improves the performance extensively, which successfully detects the near-duplicate videos with complex transformations and filter out dissimilar ones. The average precision over all recall levels (0.05–1.0) is shown in Table 7 and the last column of Figure 9 (denoted as AVG). The average precision is improved from 0.892 (SIG_CH) to 0.952 (HIRACH).

Table 7. Average precision of all queries over all recall levels

Methods	SIG_CH	HIRACH
Average	0.892	0.952

7. CONCLUSION

With the exponential growth of web videos, especially the coming of the Web 2.0 era, a huge number of near-duplicate videos are commonly returned from current video search engines. The diversity of near-duplicate videos ranges from simple formatting to complex mixture of different editing effects, which causes the near-duplicate video detection a challenging task. To tradeoff the performance and speed requirements, we proposed a hierarchical method to combine global signatures and local pairwise measure. Global signatures on color histogram were first used to detect clear near-duplicate videos with high confidence and filter out obviously dissimilar ones. For videos that cannot be clearly classified as novel or near-duplicate using global signatures, we applied the local feature based near-duplicate detection which provides very accurate duplicate analysis with a higher cost. Experiments on a data set of 12,790 videos retrieved from YouTube, Google Video, and Yahoo! Video show that the hierarchical approach can effectively detect a large diversity of near-duplicate videos and dramatically reduce redundant video displayed to the user in the top result set, at relatively small computational cost.

Our current research can be further extended to find the essential content that is frequently appeared across relevant videos. It could act as a good tool for gleaning a quick summary of the most important clips from the returned videos. This approach could also be used to develop customized web video crawlers that tailored to recognize users' interests and send out on autonomous search missions. Furthermore, we will build classifiers to automatically partition video into simple and complex scenes and then apply different strategies to each in the future.

8. ACKNOWLEDGEMENT

The work described in this paper was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (CityU 118905).

We'd like to thank Rong Yan for the web video crawler and Wan-Lei Zhao for the NDK detection.

9. REFERENCES

[1] D. A. Adjeroh, M. C. Lee, and I. King. A Distance Measure for Video Sequences. *CVIU*, pp. 25–45, 1999.

[2] J. Allan, editor. Topic Detection and Tracking: Event-based Information Organization. *Kluwer Academic Publishers,* 2002.

[3] J. Allan, C. Wade, and A. Bolivar. Retrieval and Novelty Detection at the Sentence Level. *ACM SIGIR'03*.

[4] T. Brants, F. Chen, and A. Farahat. A System for New Event Detection. *ACM SIGIR'03*, Canada, Jul. 2003.

[5] J. Carbonell and J. Goldstein. The Use of MMR, Diversity-based Reranking for Reordering Documents and Producing Summaries. *ACM SIGIR'98*.

[6] S-F. Chang, W. Hsu, L. Kennedy, L. Xie and et al. Columbia University TRECVID-2005 Video Search and High-Level Feature Extraction. *TRECVID 2005*, Washington DC, 2005.

[7] S. C. Cheung and A. Zakhor. Efficient Video Similarity Measurement with Video Signature. *IEEE Trans. on CSVT*, vol. 13, no. 1, pp. 59–74, Jan. 2003.

[8] S. C. Cheung and A. Zakhor. Fast Similarity Search and Clustering of Video Sequences on the World-Wide-Web. *IEEE Trans. on CSVT*, vol. 7, no. 3, pp. 524–537, June 2005.

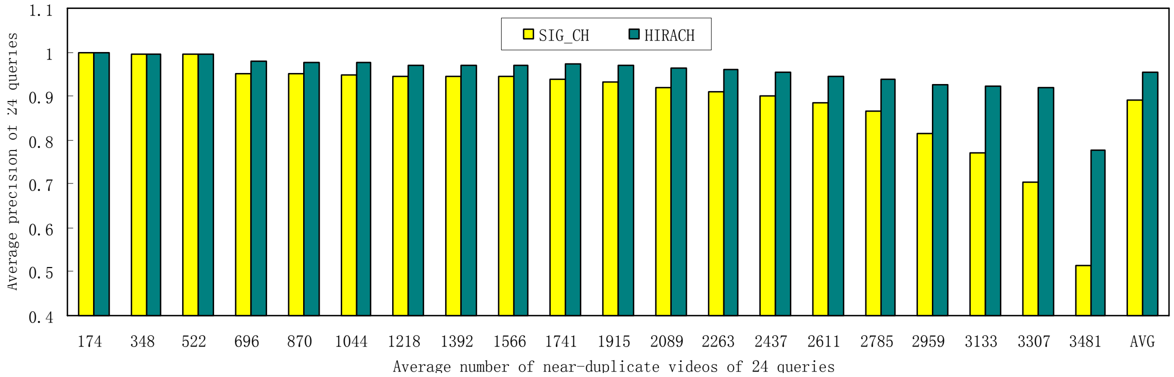


Figure 9. Average near-duplicate retrieval performance comparison for different approaches over all queries

[9] E. Gabrilovich, S. Dumais, and E. Horvitz. Newsjunkie: Providing Personalized Newsfeeds via Analysis of Information Novelty. *WWW'04*, USA, 2004, pp. 482-490.

[10] Google Video. Available: http://video.google.com.

[11] A. Hampapur and R. Bolle. Comparison of Sequence Matching Techniques for Video Copy Detection. *Conf. on Storage and Retrieval for Media Databases*, 2002.

[12] T. C. Hoad and J. Zobel. Fast Video Matching with Signature Alignment. *MIR'03*, pp. 262- 269, USA, 2003.

[13] W. H. Hsu, L. S. Kennedy and S-F. Chang. Video Search Reranking via Information Bottleneck Principle. *ACM MM'06*, USA, pp. 35-44, 2006.

[14] Informedia. Available: http://www.informedia.cs.cmu.edu.

[15] A. Jaimes. Conceptual Structures and Computational Methods for Indexing and Organization of Visual Information. *Ph.D. Thesis*, 2003.

[16] A. K. Jain, A. Vailaya, and W. Xiong. Query by Video Clip. *ACM Multimedia Syst. J.*, vol. 7, pp. 369–384, 1999.

[17] A. Joly, O. Buisson and C. Frelicot. Content-Based Copy Retrieval Using Distortion-based Probabilistic Similarity Search. *IEEE Trans. on MM*, vol. 9, no. 2, Feb. 2007.

[18] K. Kashino, Takayuki, and H. Murase. A Quick Search Method for Audio and Video Signals Based on Histogram Pruning. *IEEE Trans. on MM*, vol. 5, no. 3, 2003.

[19] Y. Ke, R. Sukthankar, and L. Huston. Efficient Near-Duplicate Detection and Sub-Image Retrieval. *ACM MM'04*.

[20] J. Law-To, B. Olivier, V. Gouet-Brunet and B. Nozha. Robust Voting Algorithm Based on Labels of Behavior for Video Copy Detection. *ACM MM'06*, pp. 835-844, 2006.

[21] R. Lienhart and W. Effelsberg. VisualGREP: A Systematic Method to Compare and Retrieve Video Sequences. *Multimedia Tools Appl.*, vol. 10, no. 1, pp. 47–72, Jan. 2000.

[22] L. Liu, W. Lai, X.-S. Hua, and S.-Q. Yang. Video Histogram: A Novel Video Signature for Efficient Web Video Duplicate Detection. *MMM'07*.

[23] X. Liu, Y. Zhuang, and Y. Pan. A New Approach to Retrieve Video by Example Video Clip. *ACM MM'99*, 1999.

[24] D. Lowe. Distinctive Image Features from Scale-Invariant Key Points. *IJCV*, vol. 60, pp. 91-110, 2004.

[25] K. Mikolajczyk and C. Schmid. A Performance Evaluation of Local Descriptors. *CVPR'03*, pp. 257-263.

[26] K. Mikolajczyk and C. Schmid. Scale and Affine Invariant Interest Point Detectors. *IJCV*, 60 (2004), pp. 63-86.

[27] C-W. Ngo, W-L. Zhao, Y-G. Jiang. Fast Tracking of Near-Duplicate Keyframes in Broadcast Domain with Transitivity Propagation. *ACM MM'06*, pp. 845-854, USA, Oct. 2006.

[28] Y. Peng and C-W. Ngo. Clip-based Similarity Measure for Query-Dependent Clip Retrieval and Video Summarization. *IEEE Trans. on CSVT*, vol. 16, no. 5, May 2006.

[29] Wikipedia. http://en.wikipedia.org/wiki/Youtube.

[30] X. Wu, A. G. Hauptmann, and C.-W. Ngo. Novelty Detection for Cross-Lingual News Stories with Visual Duplicates and Speech Transcripts. *ACM MM'07*.

[31] X. Wu, C-W. Ngo, and Q. Li. Threading and Autodocumenting News Videos. *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 59-68, March 2006.

[32] Yahoo! Video. Available: http://video.yahoo.com.

[33] Y. Yang, J. Zhang, J. Carbonell and C. Jin. Topic-conditioned Novelty Detection. *ACM SIGKDD'02*, Canada.

[34] YouTube. Available: http://www.youtube.com.

[35] J. Yuan, L.–Y. Duan, Q. Tian, S. Ranganath and C. Xu. Fast and Robust Short Video Clip Search for Copy Detection. *Pacific Rim Conf. on Multimedia (PCM)*, 2004.

[36] C. Zhai, W. Cohen and J. Lafferty. Beyond Independent Relevance: Methods and Evaluation Metrics for Subtopic Retrieval. *ACM SIGIR'03*.

[37] B. Zhang et. al. Improving Web Search Results Using Affinity Graph. *ACM SIGIR'05*.

[38] D-Q. Zhang and S-F. Chang. Detecting Image Near-Duplicate by Stochastic Attributed Relational Graph Matching with Learning. *ACM MM'04*, USA, Oct. 2004.

[39] Y. Zhang, J. Callan, and T. Minka. Novelty and Redundancy Detection in Adaptive Filtering. *ACM SIGIR'02*, 2002.