A Human-Centered Computing Framework to Enable Personalized News Video Recommendation

Hangzai Luo and Jianping Fan

Abstract. In this chapter, an interactive framework is developed to enable personalized news video recommendation and allow news seekers to access large-scale news videos more effectively. First, multiple information sources (audio, video and closed captions) are seamlessly integrated and synchronized to achieve more reliable news topic detection, and the inter-topic contextual relationships are extracted automatically for characterizing the interestingness of the news topics more effectively. Second, topic network (i.e., news topics and their inter-topic contextual relationships) and hyperbolic visualization are seamlessly integrated to achieve more effective navigation and exploration of large-scale news videos at the topic level, so that news seekers can have a good global overview of large-scale collections of news videos at the first glance. Through a hyperbolic approach for interactive topic network visualization and navigation, large amounts of news topics and their contextual relationships are visible on the display screen, and thus news seekers can obtain the *news topics of interest* interactively, build up their mental search models easily and make better search decisions by selecting the visible news topics directly. Our system can also capture the search intentions of news seekers implicitly and further recommend the most relevant news videos according to their importance and representativeness scores. Our experiments on large-scale news videos (10 TV news programs for more than 3 months) have provided very positive results.

1 Introduction

According to the CIA world factbook, there are more than 30,000 television stations in the world. These stations broadcast a large number of TV news programs

Hangzai Luo
Software Engineering Institute, East China Normal University, Shanghai, China
e-mail: hzluo@sei.ecnu.edu.cn

Jianping Fan
Dept. of Computer Science, UNC-Charlott, NC 28223, USA
e-mail: jfan@uncc.ed

D. Schonfeld et al. (Eds.): Video Search and Mining, SCI 287, pp. 261–281.
springerlink.com

(news videos) every day. Different organizations and individuals utilize these broadcast news videos for different purposes, such as presidential candidates' debat for public assessment, economic performance analysis and prediction, sports and crime reports. People watch the news videos (TV news programs) to understand what is happening now and predict what might happen in the near future, so that they can make better daily decisions.

Due to the large number of broadcast channels and TV news programs, finding news videos of interest is not a trivial task: (a) Most existing content-based video retrieval (CBVR) systems assume that news seekers can formulate their information needs precisely either in terms of keywords or example videos. Unfortunately, news seekers may not be able to know what is happening now (i.e., if they know it, it is not a news), thus it is very hard for them to find the suitable keywords or example videos to formulate their news needs precisely without obtaining sufficient knowledge of the available news topics of interest. Thus there is an urgent need to develop new techniques for detecting news topics of interest from large-scale news videos to assist news seekers on finding news videos of interest more effectively. (b) Because the same news topic can be discussed in many TV channels and news programs, topic-based news search may return large amounts of news videos and thus simple news search via keyword matching of news topics may bring the serious problem of information overload to news seekers. (c) Most existing CBVR systems treat all the news seekers equally while completely ignoring the diversity and rapid change of their search interests. Besides the rapid growth of broadcast TV channels and news programs, we have also observed different scenarios of news needs from different people, thus it is very difficult to come up with a *one size fits all* approach for accessing large-scale news videos. (d) The keywords for news topic interpretation may not be expressive enough for describing the rich details of video content precisely and using only the keywords may not be able to capture the search intentions of news seekers effectively. Thus visualization is becoming a critical component of personalized news video recommendation system [1-2, 9-12]. (e) The objectives for personalized video recommendation and content-based video retrieval are very different, which make it unsuitable to directly apply the existing CBVR techniques for supporting personalized video recommendation. Thus supporting personalized news video recommendation is becoming one important feature of news services [3-4].

There are some existing approaches to support personalized video recommendation by using only the associated text terms such as the titles, tags, and comments [3-4], and the relevant videos are recommended according to the matching between the associated text terms for video content description and the users' profiles. Unfortunately, the text terms, which are associated with the videos, may not have exact correspondence with the underlying video content. In addition, a sufficient collection of users' profiles may not be available for recommendation purpose. Thus there is an urgent need to develop new frameworks for supporting personalized news video recommendation, which may not completely depend on the users' profiles and the associated texts for video content description.

Context between the news topics is also very important for people to make better search decisions, especially when they are not familiar with the available news topics and their search goals or ideas are still fuzzy. The inter-topic context can give a good approximation of the interestingness of the news topics (i.e., like PageRank for characterizing the importance of web pages [17]). Thus it is very attractive to integrate topic network (i.e., news topics and their inter-topic contextual relationships) for characterizing the interestingness of the news topics, assisting news seekers on making better search decisions and suggesting the future search directions.

To incorporate topic network for supporting user-adaptive topic recommendation, it is very important to develop new algorithm for large-scale topic network visualization, which is able to provide a good balance between the local detail and the global context. The local detail is used to help news seekers focus on the news topics of interest in current focus. The global context is needed to tell news seekers where the other news topics are and their contextual relationships with the news topics in current focus, such global context can effectively suggest the new search directions to news seekers. Thus supporting visualization and interactive navigation of the topic network is becoming a complementary and necessary component for personalized news video recommendation system and it may lead to the discovery of unexpected news videos and guide the future search directions effectively.

On the other hand, the search criteria are often poorly defined or depend on the personal preferences of news seekers. Thus supporting interactive visualization, exploration and assessment of the search results are very important for allowing news seekers to find the news videos of interest according to their personal preferences. Information retrieval community has also recognized that designing more intuitive system interface for search result display may have significant effects on assisting users to understand and assess the search results more effectively [13]. To incorporate visualization for improving news search, effective techniques for intelligemt news video analysis should be developed to discover the meaningful knowledge from large-scale news videos.

Several researchers have used the ontology (i.e., video concepts and their simple inter-topic contextual relationships such as “IS-A" and “part-of") to assistant visual content anslysis and retrieval [23-24]. Because the news content are highly dynamic, the inter-topic contextual relationships cannot simply be characterized by using “IS-A" or “part-of", which are used for ontology construction. Thus it is unacceptable to incorporate the ontology for supporting personalized news video recommendation. On the other hand, automatic video understanding is still an open problem for computer vision community [25-31].

In this chapter, an interactive approach is developed to enable personalized news video recommendation, and our approach has significant differences from other existing work: (a) Rather than performing semantic video classification for automatic news video understanding, we have integrated multiple information sources to achieve more reliable news topic detection. (b) The associations among the news topics (i.e., inter-topic contextual relationships) are determined automatically and an interestingness score is automatically assigned to each news topic via statistical analysis, and such interestingness scores are further used to select the news topics of interest and filter out the less interesting news topics automatically. (c) A hyperbolic visualization tool is incorporated to inform news seekers with a better global overview of large-scale news videos, so that they can make better search decisions and find the most relevant news videos more effectively. (d) A novel video ranking algorithm is developed for recommending the most relevant news videos according to their importance and representativeness scores.

The chapter is organized as follows. Section 2 briefly reviews some related work on news topic detection and personalized information recommendation; Section 3 introduces our work on integrating topic network and hyperbolic visualization to enable user-adaptive topic recommendation; Section 4 introduces our new scheme on news video ranking for supporting personalized news video recommendation; Section 5 summarizes our work on algorithm and system evaluation; We conclude in Section 6.

2 Related Work

To enable personalized news video recommendation, one of the most important problems is to extract *news topics of interest* automatically from large-scale news videos. This problem is becoming very critical because of the following reasons: (a) The amount of news topics could be very large; (b) Different news topics may have different importance and interestingness scores, such importance and interestingness scores may also depend on the personal preferences of news seekers. In this section, we have provided a brief review of some existing work which are critical for developing personalized news video recommendation system: (1) automatic news topic detection; (2) news visualization; (3) personalized video recommendation.

Topic extraction refers to the identification of individual stories or topics within a broadcast news video by detecting the boundaries where the topic of discussion changes. News topics may be of any length and consist of complete and cohesive news report on one particular topic. Each broadcast channel has its own peculiarities in terms of program structures and styles, which can be integrated for achieving more accurate detection of news topics and their boundaries [11-12]. News topics can also be detected by using some existing techniques for named-entity extraction [5-6].

There are two well-accepted approaches for supporting personalized information retrieval [20-22]: *content-based filtering* and *collaborative filtering*. Because the profiles for new users are not available, both the collaborative filtering approach and the content-based filtering approach cannot support new users effectively. Thus there is an urgent need to develop more effective frameworks for supporting personalized news video recommendation.

Visualization is widely used to help the users explore large amount of information and find interesting parts interactively [9-12]. Rather than recommending the most interesting news topics to news seekers, all of these existing news visualization systems disclose all the available news topics to them, and thus news seekers have to dig out the news topics of interest by themselves. When large-scale news collections come into view, the number of the available news topics could be very large and displaying all of them to news seekers may mislead them. Thus it is very important to develop new algorithms for characterizing the interestingness of news topics and reducing the number of news topics to enable more effective visualization and exploration of large-scale news videos.

3 User-Adaptive News Topic Recommendation

In this chapter, a novel scheme is developed by incorporating *topic network* and hyperbolic visualization to recommend the *news topics of interest* for assisting news seekers on accessing large-scale news videos more effectively. To do this, an automatic scheme is developed to construct the topic network for representing and interpreting large-scale news videos at the topic level. In addition, a hyperbolic visualization technique is developed to enable interactive topic network navigation and recommend the news topics of interest according to the personal preferences and timely observations of news seekers, so that they can make better search decisions.

3.1 News Topic Detection

For TV news programs, there are three major information sources (audio, video and closed captions) that can be integrated and synchronized to enable more reliable

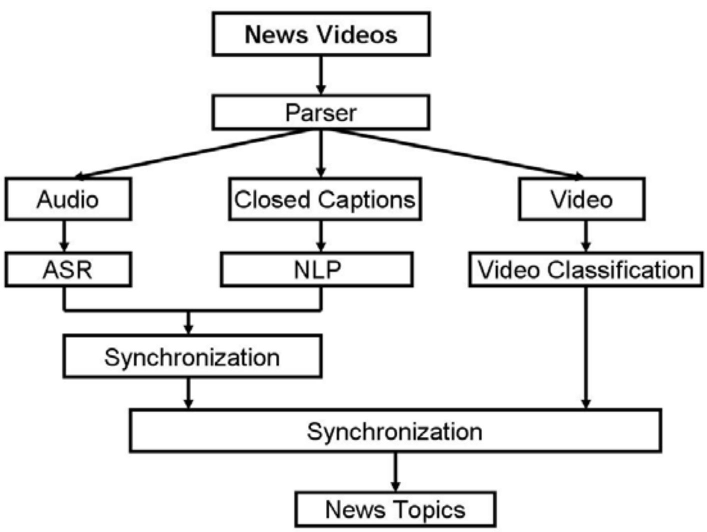

Fig. 1 The flowchart for synchronizing multiple sources for news topic detection, where automatic speech recognition (ASR), natural language processing (NLP), and semantic video classification are seamlessly integrated.

news topic detection. We have developed a new scheme for automatic news topic detection by taking the advantage of multiple information sources (cross-media) as shown in Fig. 1. First, automatic speech recognition (ASR), natural language processing (NLP), and semantic video classification are performed on these three information sources parallelly to determime the keywords for news topic description from both the audio channel and the closed captions and detect the video concepts from the video channel. Second, the audio channel is synchronized with the closed caption channel, and the video channel is further synchronized with the audio channel and the closed caption channel. Finally, the detection results of news topics from these three information sources are integrated to boost the performance of our news topic detection algorithm.

The closed captions of news videos can provide abundant information and such information can be used to detect the news topics of interest and their semantic interpretations with high accuracy. To do this, the closed captions are first segmented into a set of sentences, and each sentence is further segmented into a set of keywords. In news videos, some special text sentences, such as "*somebody*, CNN, *somewhere*" and "ABC's *somebody* reports from *somewhere*", need to be processed separately. The names for news reporters in those text sentences are generally not the content of news report. Therefore, they are not appropriate for news semantics interpretation and should be removed. Because there have some clear and fixed patterns for these specific sentences, we have designed a context-free syntax parser to detect and mark this information. By incorporating 10-15 syntax rules, our parser can detect and mark such specific sentences in high accuracy. Standard text processing techniques are used to remove the stop words automatically.

Most named entity detectors may fail in processing all-capital strings because initial capitalization is very important to achieve accurate named entity recognition. One way to resolve this problem is to train a detector with ground truth from the text documents of closed captions. However, it's very expensive to obtain the manually marked text material. Because English has relatively strict grammar, it's possible to parse the sentences and recover the most capital information by using part-of-speech (POS) and lemma information. TreeTagger [7] is used to perform the part-of-speech tagging. Capital information can be recovered automatically by using the TreeTagger parsing results.

After such specific sentences are marked and the capital information is recovered, an open source text analysis package LingPipe [8] is used to perform the named entity detection and resolve co-reference of the named entities. The named entities referring to the same entity are normalized to the most representative format to enable statistical analysis, where the news model of LingPipe is used and all the parameters are set to default value. Finally, the normalized results are parsed again by TreeTagger to extract the POS information and resolve the words to their original formats. For example, TreeTagger can resolve "better" to "well" or "good" according to its POS tag.

We have defined a set of over 4000 elemental topics, each keyword represents an elemental topic, and all these detected news stories that consist of one particular keyword are assigned to the corresponding cluster of news topic. Our multi-task

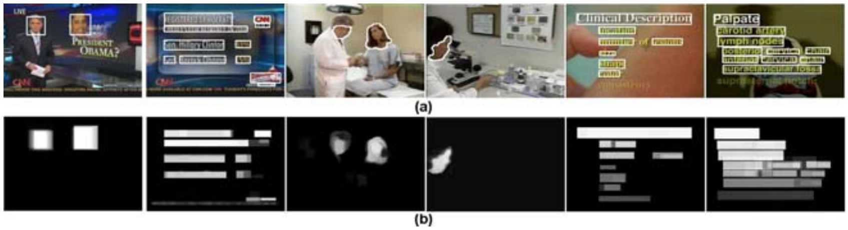

Fig. 2 Integrating confidence map for salient object detection: (a) original images and the detected salient objects; (b) confidence maps for the salient objects.

learning algorithm is performed to learn the topic detectors from a given corpus by exploiting the inter-topic correlation [25-27]. Once we have a set of topic detectors, they are used to determine the most topic-similar clusters for the new piece of news videos.

For TV news videos, the video shots are the basic units for video content representation, and thus they can be treated as one of the semantic items for news topic detection. Unlike the keywords in text documents, the re-appearance of video shots cannot be detected automatically via simple comparison of their visual properties. For news videos, video objects, such as text areas and human faces, may provide important clues about news stories of interest. Text lines and human faces in news videos can be detected automatically by using suitable computer vision techniques [28]. Obviously, these automatic detection functions may fail in some cases. Thus the results that are detected by using a single video frame may not be reliable. To address this problem, the detection results on all the video frames within the same video shot are integrated and the corresponding confidence maps for the detection results are calculated as shown in Fig. 2 [27]. The video concepts associated with the video shots can provide valuable information to enable more accurate news topic detection, and semantic video classification is one potential solution to detect such video concepts [27]. To detect the video concepts automatically, we have adopted our previous work reported in [25-28].

Unfortunately, the closed captions may not synchronize with the video channel accurately and have a delay of a few seconds in general. Thus the news topics that are detected from the closed captions cannot directly be synchronized with the video concepts that are detected from the news videos. On the other hand, the closed captions have good synchronization with the relevant audios. Therefore, they can be integrated to take advantage of cross-media to clarify the video content and remove the redundant information. Even the audio channel generally synchronizes very well with the video channel, the accuracy of most existing techniques for automatic speech recognition (ASR) is still low. By integrating the results for automatic speech recognition with the topic detection results from the closed captions, we can synchronize the closed captions with the video content in higher accuracy. After the closed captions are synchronized with the news videos, we can assign the video shots to the most relevant news topics that are accurately detected from the closed captions. Thus all the video shots, which locate between the start time and the end time of a given new topic that has been detected from the closed captions, are assigned to the given news topic automatically.

3.2 Topic Association Extraction

The contextual relationships among these significant news topics are obtained automatically, where both the semantic similarity and the co-occurrence probability for the relevant news topics are used to define a new measurement for determining the inter-topic associations effectively. The inter-topic association (i.e., inter-topic contextual relationship) $\phi(C_i, C_j)$ is determined by:

$$\phi(C_i, C_j) = -\alpha \cdot \log \frac{d(C_i, C_j)}{2L} + \beta \cdot \frac{\psi(C_i, C_j)}{\log \psi(C_i, C_j)}, \quad \alpha + \beta = 1 \qquad (1)$$

where the first part denotes the semantic similarity between the news topics C_j and C_i, the second part indicates their co-occurrence probability, α and β are the weighting parameters, $d(C_i, C_j)$ is the length of the shortest path between the news topics C_i and C_j by searching the relevant keywords for news topic interpretation from WordNet [23], L is the maximum depth of WordNet, $\psi(C_i, C_j)$ is the co-ocurrence probability between the relevant news topics. The co-occurrence probability $\psi(C_i, C_j)$, between two news topics C_j and C_i, is obtained in the news topic detection process. Obviously, the value of the inter-topic association $\phi(C_i, C_j)$ increases with the strength of the contextual relationship between the news topics C_i and C_j.

Thus each news topic is automatically linked with multiple relevant news topics with the higher values of the associations $\phi(\cdot, \cdot)$. One portion of our large-scale topic network is given in Fig. 3, where the news topics are connected and organized according to the strength of their associations, $\phi(\cdot, \cdot)$. One can observe that such a topic network can provide a good global overview of large-scale news videos and

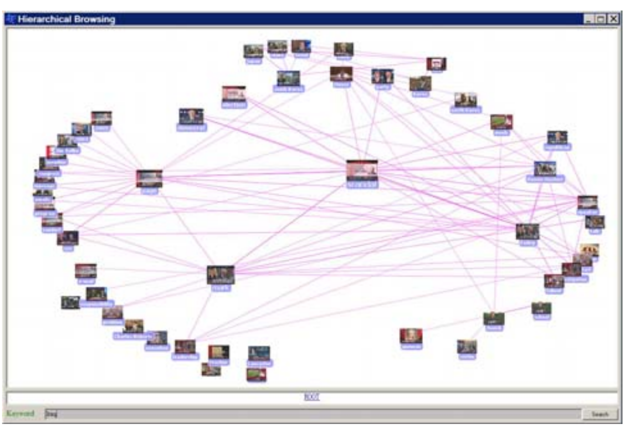

Fig. 3 One portion of our topic network for organizing large-scale news videos.

can precisely characterize the interestingness of the relevant news topics, and thus it can be used to assist news seekers on making better search decisions.

To integrate the topic network for supporting user-adaptive topic recommendation, it is very attractive to achieve graphical representation and visualization of the topic network, so that news seekers can obtain a good global overview of large-scale news videos at the first glance and make better search decisions in the process of interactive topic network exploration and navigation. Unfortunately, visualizing large-scale topic network in a 2D system interface with a limited screen size is not a trivial task. To achieve more effective visualization of large-scale topic network, we have developed multiple innovative techniques: (a) highlighting the news topics according to their interestingness scores for allowing news seekers to obtain the most important insights at the first glance; (b) integrating hyperbolic geometry to create more space for large-scale topic network visualization and exploration.

3.3 Interestingness Scores of News Topics

We have integrated both the popularity of the news topics and the importance of the news topics to determine their interestingness scores. The popularity of a given news topic is related to the number of TV channels or news programs which have discussed or reported the given news topic. If one news topic is discussed or reported by more TV channels or news programs, it tends to be more interesting. The importance of a given news topic is also related to its linkage structure with other news topics on the topic network. If one news topic is related to more news topics on the topic network, it tends to be more interesting [17]. For example, the news topic for “roadside bond in Iraq" may relate to the news topics of “gap price increase" and “stock decrease". Thus the interestingness score $\rho(C_i)$ for a given news topic C_i is defined as:

$$\rho(C) = \lambda \cdot \log(m(C_i) + \sqrt{m^2(C_i) + 1}) + \gamma \cdot \log(k(C_i) + \sqrt{k^2(C_i) + 1}), \lambda + \gamma = 1 \quad (2)$$

where $m(c_i)$ is the number of TV channels or news programs which have discussed or reported the given news topic C_i, $k(c_i)$ is the number of news topics linked with the given news topic C_i on the topic network. Thus the interestingness score for a given news topic increases adaptively with both the number of the relevant TV channels or news programs and the number of the linked news topics. Such interestingness scores can be used to highlight the most interesting news topics and eliminate the less interesting news topics for reducing the visual complexity for large-scale topic network visualization and exploration.

3.4 Hyperbolic Topic Network Visualization

Supporting graphical representation and visualization of the topic network can provide an effective solution for exploring large-scale news videos at the topic level and recommend the *news topics of interest* interactively for assisting news seekers to make

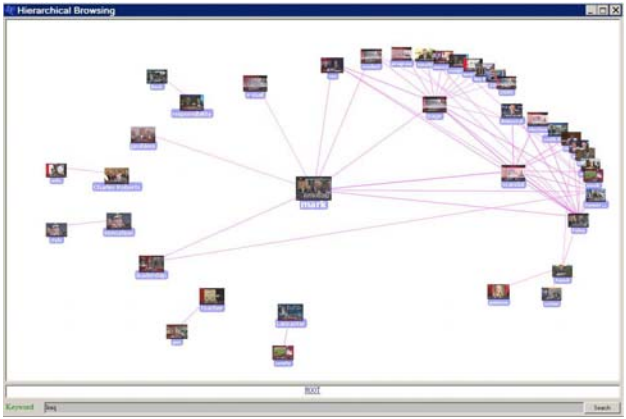

Fig. 4 One view of hyperbolic visualization of our topic network.

better search decisions. However, visualizing large-scale topic network in a 2D system interface with a limited screen size is a challenging task. We have investigated multiple solutions to tackle this challenge task: (a) A string-based approach is incorporated to visualize the topic network with a nested view, where each news topic node is displayed closely with the most relevant news topic nodes according to the values of their associations. The underlying inter-topic contextual relationships are represented as the linkage strings. (b) The geometric closeness of the news topic nodes is related to the strength of their inter-topic contextual relationships, so that such graphical representation of the topic network can reveal a great deal about how these news topics are connected. (c) Both geometric zooming and semantic zooming are integrated to adjust the levels of visible details automatically according to the discerning constraint on the number of news topic nodes that can be displayed per view.

Our approach for topic network visualization exploits hyperbolic geometry [14-16]. The hyperbolic geometry is particularly well suited for achieving graph-based layout of the topic network, and it has “more space" than Euclidean geometry. The essence of our approach is to project the topic network onto a hyperbolic plane according to the inter-topic contextual relationships, and layout the topic network by mapping the relevant news topic nodes onto a circular display region. Thus our topic network visualization scheme takes the following steps: (a) The news topic nodes on the topic network are projected onto a hyperbolic plane according to their inter-topic contextual relationships, and such projection can usually preserve the original contextual relationships between the news topic nodes. (b) After such context-preserving projection of the news topic nodes is obtained, Poincaré disk model [14-16] is used to map the news topic nodes on the hyperbolic plane to a 2D display coordinate. Poincaré disk model maps the entire hyperbolic space onto an open unit circle, and produces a non-uniform mapping of the news topic nodes to the 2D display coordinate.

Our approach for topic network visualization relies on the representation of the hyperbolic plane, rigid transformations of the hyperbolic plane and mappings of

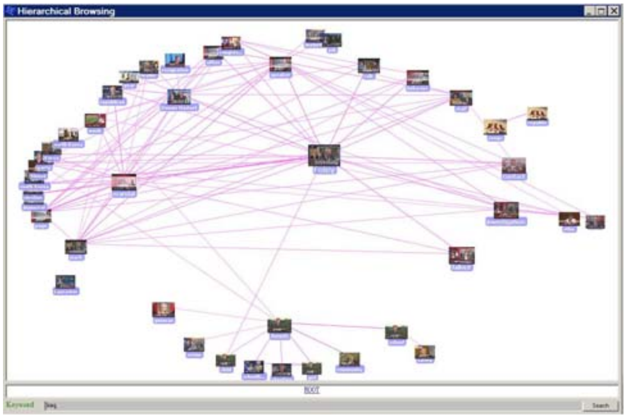

Fig. 5 Another view of hyperbolic visualization of our topic network.

the news topic nodes from the hyperbolic plane to the unit disk. Internally, each news topic node on the graph is assigned a location $z = (x, y)$ within the unit disk, which represents its Poincaré coordinates. By treating the location of the news topic node as a complex number, we can define such a mapping as the linear fractional transformation [14-16]:

$$z_t = \frac{\theta z + P}{1 + \bar{P}\theta z} \tag{4}$$

where P and θ are the complex numbers, $|P| < 1$ and $|\theta| = 1$, and $\bar{P}$ is the complex conjugate of P. This transformation indicates a rotation by θ around the origin following by moving the origin to P (and $-P$ to the origin).

To incorporate such transformation for topic network visualization, the layout routine is structured as a recursion that takes a news topic node and a wedge in which to lay out the news topic node and its relevant news topic nodes. It places the news topic node at the vertex of the wedge, computes a wedge for each relevant news topic node and recursively calls itself on each relevant news topic node. The relevant news topic nodes are placed in the middle of their subwedges at a distance computed by the formula:

$$d = \sqrt{\left(\frac{(1-s^2)sin(a)}{2s}\right)^2 + 1} - \frac{(1-s^2)sin(a)}{2s} \tag{5}$$

where a is the angle between the midline and the edge of the subwedge and s is the desired distance between a relevant news topic node and the edge of its subwedge. In our current implementation, we set $s = 0.18$. The result, d, is the necessary distance from current news topic node to its relevant news topic node. If the value of d is less than that of s, we set d to s for maintaining a minimum space between the current news topic node and the relevant news topic node. Both s and d are represented as the hyperbolic tangent of the distance in the hyperbolic plane.

3.5 Personalized Topic Network Generation

After the hyperbolic visualization of the topic network is available, it can be used to enable interactive exploration and navigation of large-scale news videos at the topic level via *change of focus*. The *change of focus* is implemented by changing the mapping of the news topic nodes from the hyperbolic plane to the unit disk for display, and the positions of the news topic nodes in the hyperbolic plane need not to be altered during the focus manipulation. As shown in Fig. 4 and Fig. 5, news seekers can change their focuses of the news topics by clicking on any visible news topic node to bring it into the focus at the screen center, or by dragging any visible news topic node interactively to any other screen location without losing the contextual relationships between the news topics, where the rest of the layout of the topic network transforms appropriately. In such interactive topic network navigation and exploration process, news seekers can obtain the *news topics of interest* interactively, build up their mental search models easily and make better search decision effectively by selecting the visible news topics directly. Because our hyperbolic visualization framework can assign more spaces for the news topic node in current focus and ignore the details for the residual news topic nodes on the topic network, it can theoretically avoid the overlapping problem by supporting change of focus and thus it can spporting large-scale topic network visualization and navigation.

On the other hand, such interactive topic network exploration process has also provided a novel approach for capturing the search interests of news seekers automatically. We have developed a new algorithm for generating personalized topic network automatically from the current search actions of news seeker while the new seeker navigates the topic network. Thus the *personalized interestingness score* for a given news topic C_i on the topic network is defined as:

$$\rho(C_i)=\rho^{org}(C_i)+\rho^{org}(C_i)\left\{\alpha\frac{e^{v(C_i)}-e^{-v(C_i)}}{e^{v(C_i)}+e^{-v(C_i)}}+\beta\frac{e^{s(C_i)}-e^{-s(C_i)}}{e^{s(C_i)}+e^{-s(C_i)}}+\delta\frac{e^{d(C_i)}-e^{-d(C_i)}}{e^{d(C_i)}+e^{-d(C_i)}}\right\} \tag{6}$$

where $\alpha+\beta+\delta=1$, $\rho^{org}(C_i)$ is the original interestingness score for the given news topic C_i as defined in Eq. (2), $v(C_i)$ is the visiting times of the given news topic C_i from the particular news seeker, $s(C_i)$ is the staying seconds on the given news topic C_i from the particular news seeker, $d(C_i)$ is the interaction depth for the particular user to interact with the news topic C_i and the relevant news videos which are relevant to the given news topic C_i, α, β and δ are the weighting factors. The visiting times $v(C_i)$, the staying seconds $s(C_i)$, and the interaction depth $d(C_i)$ can be captured automatically in the user-system interaction procedure. Thus the personalized interestingness scores of the news topics are determined immediately when such user-system interaction happens, and they will increase adaptively with the visiting times $v(C_i)$, the staying seconds $s(C_i)$, and the interaction depth $d(C_i)$.

After the personalized interestingness scores for all these news topics are learned from the current search actions of news seeker, they can further be used to weight the news topics for generating a *personalized topic network* to represent the user profiles (i.e., search preferences of news seeker) precisely. Thus the news topics

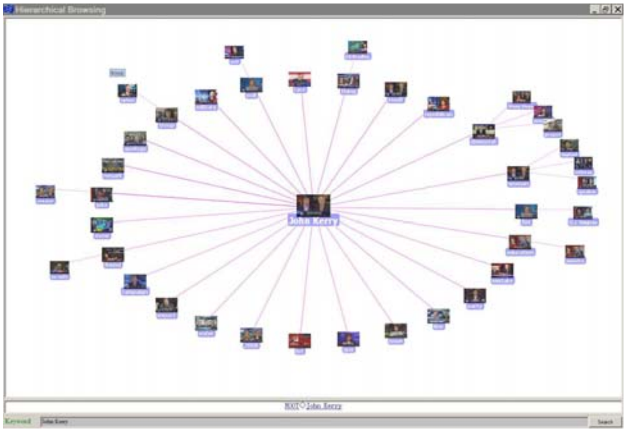

Fig. 6 The most relevant news topics for interestingness propagation.

with smaller values of the personalized interestingness scores can be eliminated automatically from the topic network, so that each news seeker can be informed by the most interesting news topics according to his/her personal preferences.

The search interests of news seeker may be changed according to his/her timely observations of news topics, and one major problem for integrating the user's profiles for topic recommendation is that the user's profiles may over-specify the search interests of news seeker and thus they may hinder news seeker to search other interesting news topics on the topic network. Based on this observation, we have developed a novel algorithm for propagating the search preferences of news seeker over other relevant news topics on the topic network, i.e., the news topics which have stronger correlations with the news topics which have been accessed by the particular news seeker. Thus the personalized interestingness score $v(C_j)$ for the news topic C_j to be propagated is determined as:

$$\chi(C_j) = \rho(C_j)\bar{\phi}(C_j), \quad \bar{\phi}(C_j) = \sum_{l \in \Omega} \phi(C_l, C_j) \tag{7}$$

where Ω is the set of the accessed news topics linked with the news topic C_j to be propagated as shown in Fig. 6 and Fig. 7, $\phi(C_l, C_j)$ is the inter-topic association between the news topic C_j and the news topic C_l which is linked with C_j and has been accessed by the particular news seeker, and $\rho(C_j)$ is the interestingness score for the news topic C_j to be propagated. Thus the news topics, which have larger values of the personalized interestingness scores $\chi(\cdot)$ (strongly linked with some accessed news topics on the topic network), can be propagated adaptively.

By integrating the inter-topic correlations for automatic propagation of the preferences of news seeker, our proposed framework can precisely predict his/her hidden preferences (i.e., search intentions) from his/her current actions. Thus the user's profiles can be represented precisely by using the personalized topic network, where the interesting news topics can be highlighted according to their personalized

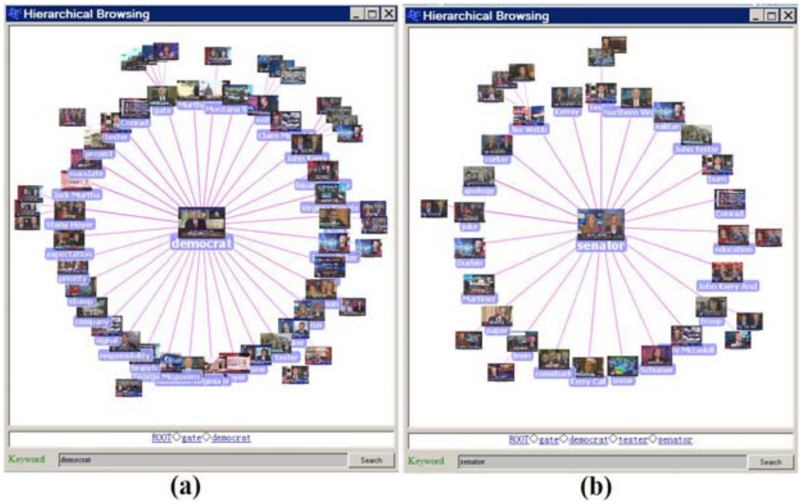

Fig. 7 The most relevant news topics for interestingness propagation.

interestingness scores as shown in Fig. 6 and Fig. 7. Such personalized topic network can further be used to recommend the news topics of interest for news seekers to make better future search decisions.

4 Personalized News Video Recommendation

Because the same news topic may be discussed many times in the same TV news program or be discussed simultaneously by multiple TV news programs, the amount of news videos under the same topic could be very large. Thus topic-based news search via keyword matching may return large amount of news videos which are relevant to the same news topic. To reduce the information overload, it is very important to develop new algorithms for ranking the news videos under the same news topic and recommending the most relevant news videos according to their importance and representativeness scores [18-19].

The news videos, which are relevant to the given news topic C_j, are ranked according to their importance and representiveness scores. For the given news topic C_j, the importance and representativeness score $\varrho(x|C_j)$ for one particular news video x is defined as:

$$\varrho(x|C_j) = \alpha e^{-\Delta t} + (1-\alpha)\left\{\beta\frac{e^{v(x|C_j)} - e^{-v(x|C_j)}}{e^{v(x|C_j)} + e^{-v(x|C_j)}} + \gamma\frac{e^{r(x|C_j)} - e^{-r(x|C_j)}}{e^{r(x|C_j)} + e^{-r(x|C_j)}} + \eta\frac{e^{q(x|C_j)} - e^{-q(x|C_j)}}{e^{q(x|C_j)} + e^{-q(x|C_j)}}\right\} \tag{8}$$

where $\beta + \lambda + \eta = 1$, Δt is the time difference between the time for the TV news programs to discuss and report the given news topic C_j and the time for the news seeker to submit their searches, $v(x|C_j)$ is the visiting times of the given news video x from all the news seekers, $r(x|C_j)$ is the rating score of the given news video x from all the news seekers, $q(x|C_j)$ is the quality of the given news video.

We separate the time factor from other factors for news video ranking because the time factor is more important than other factors for news video ranking (i.e., one topic can be treated as the news because it is new and tell people what is happening now or what is discussing now). The quality $q(x|C_j)$ is simply defined as the frame resolution and the length of the given news video x. If a news video has higher frame resolution and longer length (be discussed for longer time), it should be more important and representative for the given news topic.

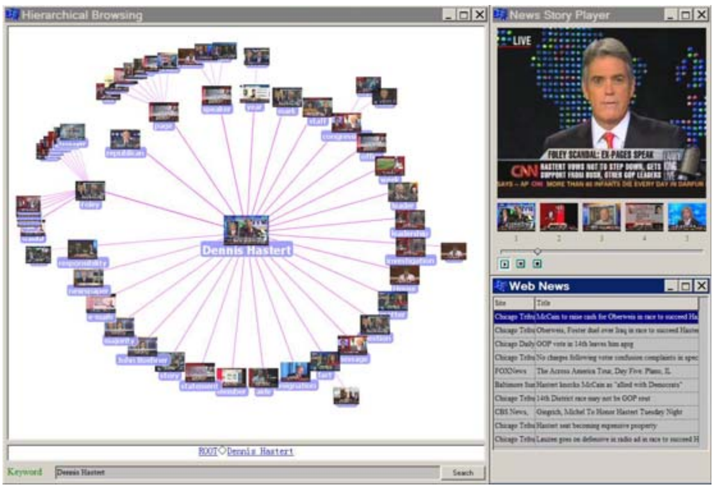

Fig. 8 Our system interface for supporting multi-modal news recommendation.

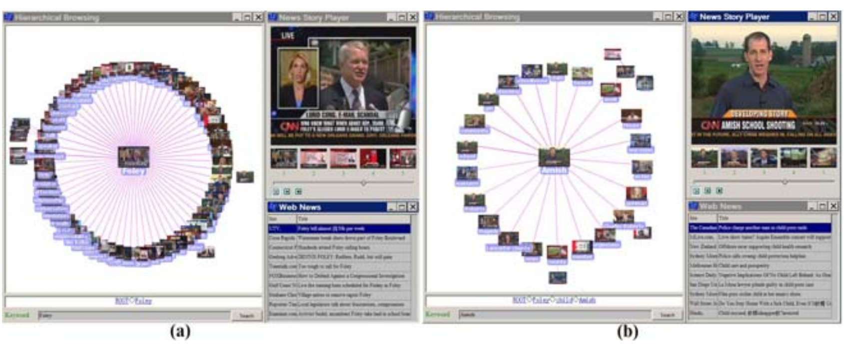

Fig. 9 Two examples for supporting multi-modal news recommendation.

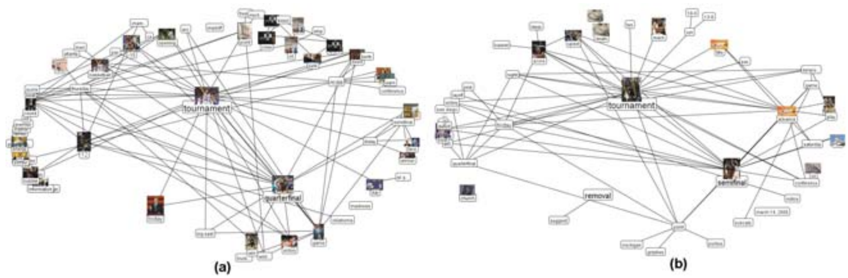

Fig. 10 Our system for supporting online news recommendation: (a) topic network for March 13; (b) topic network for March 14.

After the search goals pf news seekers (i.e., which are represented by the accessed news topics) are captured interactively, our personalized news video recommendation system can: (a) recommend top 5 news videos according to their importance and representative scores; (b) recommend the news topics of interest on the topic network which are most relevant to the accessed news topic and suggest them as the future search directions according to the current preferences of news seekers, where the accessed news topic is set as the current focus (i.e., center of the topic network); (c) recommend the most relevant online text news which are relevant with the accessed news topic, so that news seekers can also read the most relevant online web news; (d) record the search history and preferences of news seekers for generating more reliable personalized topic network to make better recommendation in the future. Some experimental results are given in Fig. 8 and Fig. 9, one can conclude that our personalized news video recommendation system can effectively support multi-modal news recommendation from large-scale collections of news videos.

We have also extended our multi-modal news analysis tools to support personalized online news recommendation. First, the news topics of interest are extracted from large-scale online news collections and the inter-topic similarity contexts are determined for topic network generation as shown in Fig. 10, one can observe that

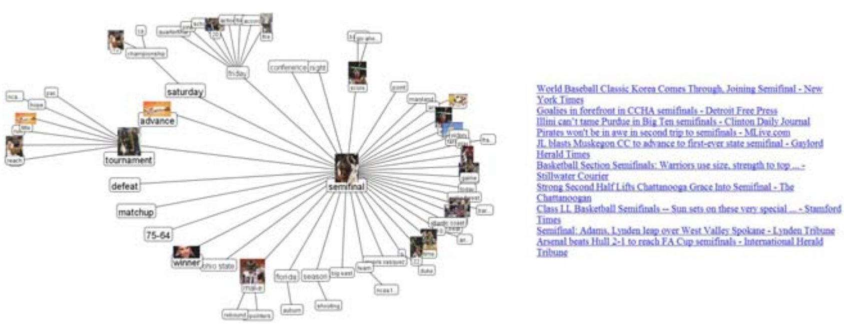

Fig. 11 Our system for supporting online news recommendation: personalized topic network and the relevant online news sources.

such the topic network can represent the news topics of interest and their inter-topic similarity contexts effectively. By incorporating the inputs of news seekers for online news recommendation, the accessed news topic is set as the current focus and the most relevant news sources are recommended as shown in Fig. 11.

5 Algorithm Evaluation

We carry out our experimental studies by using large-scale news videos. The topic network which consists of 4000 most popular news topics is learned automatically from large-scale news videos. Our work on algorithm evaluation focus on: (1) evaluating the performance of our news topic detection algorithm and assessing the advantages for integrating multiple information sources for news topic detection; (2) evaluating the response time for supporting change of focus in our system, which is critical for supporting interactive navigation and exploration of large-scale topic network to enable user-adaptive topic recommendation; (3) evaluating the performance (efficiency and accuracy) of our system for allowing news seekers to look for some particular news videos of interest (i.e., personalized news video recommendation).

Automatic news topic detection plays an important role in our personalized news video recommendation system. However, automatic topic detection is still an open problem in natural language processing community. On the other hand, automatic video understanding via semantic classification is also an open issue in computer vision community. In this chapter, we have integrated multiple information sources (audio, video and closed captions) to exploit the cross-media advantages for achieving more reliable news topic detection.

Based on this observation, our algorithm evaluation for our automatic news topic detection algorithm focuses on comparing its performance difference by combining different information sources for news topic detection. We have compared three combination scenarios for news topic detection: (a) only the closed captions are used for news topic detection; (b) the closed captions and the audio channel are integrated and synchronized for news topic detection; (c) the closed captions, the audio channel and the video channel are seamlessly integrated and synchronized for news topic detection. As shown in Fig. 12, integrating multiple information sources for news topic detection can enhance the performance of our algorithm significantly.

One critical issue for evaluating our personalized news video recommendation system is the response time for supporting change of focus to enable interactive topic network navigation and exploration, which is critical for supporting user-adaptive topic recommendation. In our system, the change of focus is used for achieving interactive exploration and navigation of large-scale topic network. The *change of focus* is implemented by changing the Poincaré mapping of the news topic nodes from the hyperbolic plane to the display unit disk, and the positions of the news topic nodes in the hyerbolic plane need not to be altered during the focus manipulation. Thus the response time for supporting change of focus depends on two components: (a) The computational time T_1 for re-calculating the new Poincaré mapping of large-scale topic network from a hyperbolic plane to a 2D display unit

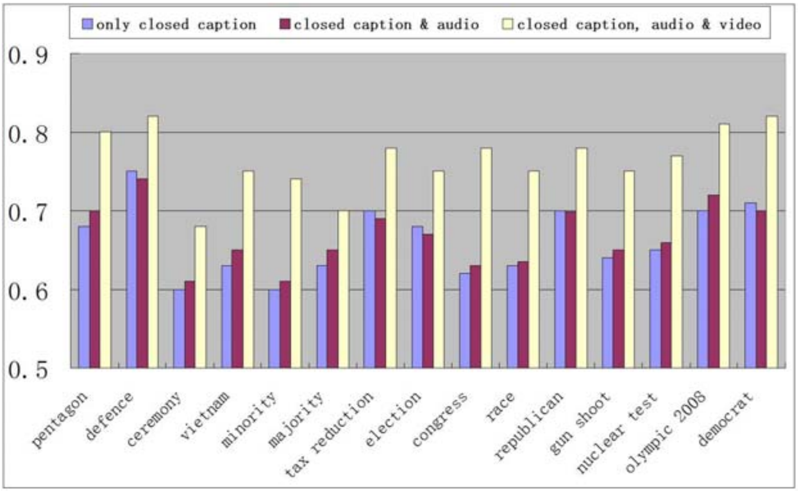

Fig. 12 The comparision results of our automatic news topic detection algorithm by integrating different sources.

disk, i.e., re-calculating the Poincaré position for each news topic node; (b) The visualization time T_2 for re-layouting and re-visualizing the new Poincaré mapping of large-scale topic network on the display disk unit. As shown in Fig. 13, one can find that the computational time T_1 is not sensitive to the number of news topics, and thus re-calculating the Poincaré mapping for large-scale topic network can almost be achieved in real time. We have also evaluated the empirical relationship between the visualization time T_2 and the number of news topic nodes. In our experiments, we have found that re-visualization of large-scale topic network is not sensitive to the number of news topics, and thus our system can support re-visualization of large-scale topic network in real time.

When the news topics of interest are recommended, our system can further allow news seekers to look for the most relevant news videos according to their importance and representative scores. For evaluating the effeciency and the accuracy of our personalized news video recommendation system, the *benchmark metric* includes *precision* P and *recall* R. The precision P is used to characterize the accuracy of

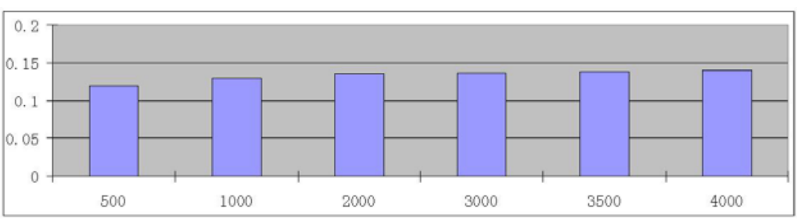

Fig. 13 The empirical relationship between the computational time T_1 (seconds) and the number of news topic nodes.

Table 1 The precision and recall for supporting personalized news video recommendation.

news topics	policy	pentagon	change	insult
precision/recall	95.6% /97.3%	98.5% /98.9%	100% /99.2%	92.8% /93.6%
news topics	implant	wedding	haggard	bob
precision/recall	90.2% /93.5%	96.3% /94.5%	96.5% /92.8%	90.3% /97.4%
news topics	gate	steny hoyer	democrat	urtha
precision/recall	95.9% /96.8%	96.5% /96.2%	96.3% /97.1%	93.6% /94.3%
news topics	majority	leader	confirmation	defence
precision/recall	99.2% /98.6%	93.8% /99.3%	94.5% /93.8%	100% /99.6%
news topics	secretary	veterm	ceremony	service
precision/recall	100% /98.8%	99.8% /99.2%	99.3% /96.6%	91.2% /93.2%
news topics	honor	vietnam	lesson	submit
precision/recall	91.2% /93.5%	98.8% /96.7%	90.3% /91.6%	91.2% /91.5%
news topics	minority	indonesia	president	trent lott
precision/recall	100% /99.6%	96.8% /97.7%	100% /96.8%	92.5% /92.3%
news topics	o.j. sinpson	trial	money	book
precision/recall	95.6% /99.4%	90.5% /90.3%	100% /90.6%	96.8% /93.6%
news topics	john kerry	military	race	mandate
precision/recall	100% /96.5%	100% /93.2%	100% /97.8%	92.6% /92.5%
news topics	election	leadship	school gun shoot	execution
precision/recall	100% /95.5%	92.8% /90.3%	100% /96.7%	90.6% /91.3%
news topics	responsibility	sex	message	congress
precision/recall	92.1% /91.5%	97.5% /98.2%	88.3% /87.6%	100% /96.3%
news topics	north korea	japan	china	white house
precision/recall	100% /99.3%	98.5% /95.6%	97.3% /95.2%	100% /94.8%
news topics	nuclear test	republican	amish	gun shoot
precision/recall	100% /97.6%	91.6% /92.8%	99.5% /91.6%	100% /99.8%
news topics	teacher	conduct	program	olmypic 2008
precision/recall	93.8% /94.5%	87.92% /88.3%	83.5% /90.2%	100% /99.3%
news topics	beijing	child	tax reduction	shooting
precision/recall	99.2% /97.3%	91.3% /91.5%	98.5% /96.9%	99.6% /98.4%
news topics	safety	investigation	ethic	committee
precision/recall	94.5% /94.8%	93.3% /96.5%	93.3% /95.6%	91.8% /95.2%
news topics	scandal	dennis hastert	preseident candidates	matter
precision/recall	96.6% /97.3%	95.3% /88.3%	98.5% /97.3%	85.2% /85.3%

our system for finding the particular news videos of interest, and the recall R is used to characterize the efficiency of our system for finding the particular news videos of interest. They are defined as:

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + TN} \tag{9}$$

where TP is the set of true positive news videos that are relevant to the need of news seeker and are recommended correctly, FP is the set of fause positive news videos that are relevant to the need of news seeker and are not recommended, and TN is the set of true negative news videos that are relevant but are recommended incorrectly. Table 1 gives the precision and recall of our personalized news video recommendation system. From these experimental results, one can observe that our system can support personalized news video recommendation effectively, thus news seekers are allowed to search for some particular news videos of interest effectively.

6 Conclusions

In this chapter, we have developed an interactive framework to support personalized news video recommendation and allow news seekers to access large-scale news videos more effectively. To allow news seekers to obtain a good global overview of large-scale news videos at the topic level, topic network and hyperbolic visualization are seamlessly integrated to achieve user-adaptive topic recommendation. Thus news seekers can obtain the *news topics of interest* interactively, build up their mental search models easily and make better search decisions by selecting the visible news topics directly. Our system can also capture the search intentions of news seekers implicitly and further recommend the most relevant news videos according to their importance and representativeness scores. Our experiments on large-scale news videos have provided very positive results.

References

1. Marchionini, G.: Information seeking in electronic environments. Cambridge University Press, Cambridge (1997)
2. Fan, J., Keim, D., Gao, Y., Luo, H., Li, Z.: JustClick: Personalized Image Recommendation via Exploratory Search from Large-Scale Flickr Images. IEEE Trans. on Circuits and Systems for Video Technology 19(2), 273–288 (2009)
3. Yang, B., Mei, T., Hua, X.-S., Yang, L., Yang, S.-Q., Li, M.: Online video recommendation based on multimodal fusion and relevance feedback. In: ACM Conf. on Image and Video Retrieval (CIVR 2007), pp. 73–80 (2007)
4. Yang, H., Chaisorn, L., Zhao, Y., Neo, S.-Y., Chua, T.-S.: VideoQA: question answering on news video. ACM Multimedia, 632–641 (2003)
5. Borthwick, A., Sterling, J., Agichtein, E., Grishman, R.: NYU: Description of the MENE named entity system as used in MUC-7. In: Proc. of the Seventh Message Understanding Conf., MUC-7 (1998)
6. McDonald, D., Chen, H.: Summary in context: Searching versus browsing. ACM Trans. Information Systems 24(1), 111–141 (2006)
7. Schmid, H.: Probabilistic part-of-speech tagging using decision trees. In: Intl. Conf. on New Methods in Language Processing (1994)
8. A.I. Inc., "Lingpipe", `http://www.alias-i.com/lingpipe/`
9. Swan, R.C., Allan, J.: TimeMine: visualizing automatically constructed timelines. In: ACM SIGIR (2000)
10. Havre, S., Hetzler, B., Whitney, P., Nowell, L.: ThemeRiver: Visualizing thematic changes in large document collections. IEEE Trans. on Visualization and Computer Graphics 8(1), 9–20 (2002)

11. Luo, H., Fan, J., Yang, J., Ribarsky, W., Satoh, S.: Large-scale new video classification and hyperbolic visualization. In: IEEE Symposium on Visual Analytics Science and Technology (VAST 2007), pp. 107–114 (2007)
12. Luo, H., Fan, J., Yang, J., Ribarsky, W., Satoh, S.: Exploring large-scale video news via interactive visualization. In: IEEE Symposium on Visual Analytics Science and Technology (VAST 2006), pp. 75–82 (2006)
13. van Wijk, J.: Bridging the gaps. IEEE Computer Graphics and Applications 26(6), 6–9 (2006)
14. Walter, J.A., Ritter, H.: On interactive visualization of high-dimensional data using the hyperbolic plane. In: ACM SIGKDD (2002)
15. Lamping, J., Rao, R.: The hyperbolic browser: A focus+content technique for visualizing large hierarchies. Journal of Visual Languages and Computing 7, 33–55 (1996)
16. Furnas, G.W.: Generalized fisheye views. In: ACM CHI, pp. 16–23 (1986)
17. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. In: WWW (1998)
18. Wang, J., Chen, Z., Tao, L., Ma, W.-Y., Liu, W.: Ranking user's relevance to a topic through link analysis on web logs. In: WIDM, pp. 49–54 (2002)
19. Lai, W., Hua, X.-S., Ma, W.-Y.: Towards content-based relevance ranking for video search. ACM Multimedia, 627–630 (2006)
20. Teevan, J., Dumais, S., Horvitz, E.: Personalized search via automated analysis of interests and activities. In: ACM SIGIR (2005)
21. Goldberg, K., Roeder, T., Gupta, D., Perkins, C.: Eigentaste: A constant time collaborative filtering algorithm. Information Retrieval 4(2), 133–151 (2001)
22. Mooney, R., Roy, L.: Content-based book recommending using learning for text categorization. In: ACM Conf. on Digital Libraries, pp. 195–204 (2000)
23. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Boston (1998)
24. Naphade, M., Smith, J.R., Tesic, J., Chang, S.-F., Hsu, W., Kennedy, L., Hauptmann, A., Curtis, J.: Large-scale concept ontology for multimedia. IEEE Multimedia (2006)
25. Fan, J., Gao, Y., Luo, H.: Integrating concept ontology and multi-task learning to achieve more effective classifier training for multi-level image annotation. IEEE Trans. on Image Processing 17(3) (2008)
26. Fan, J., Gao, Y., Luo, H., Jain, R.: Mining multi-level image semantics via hierarchical classification. IEEE Trans. on Multimedia 10(1), 167–187 (2008)
27. Fan, J., Luo, H., Gao, Y., Jain, R.: Incorporating concept ontology to boost hierarchical classifier training for automatic multi-level video annotation. IEEE Trans. on Multimedia 9(5), 939–957 (2007)
28. Fan, J., Yau, D.K.Y., Elmagarmid, A.K., Aref, W.G.: Automatic image segmentation by integrating color edge detection and seeded region growing. IEEE Trans. on Image Processing 10(10), 1454–1466 (2001)
29. Wactlar, H., Hauptmann, A., Gong, Y., Christel, M.: Lessons learned from the creation and deployment of a terabyte digital video library. IEEE Computer 32(2), 66–73 (1999)
30. Adams, W.H., Iyengar, G., Lin, C.-Y., Naphade, M.R., Neti, C., Nock, H.J., Smith, J.R.: Semantic indexing of multimedia content using visual, audio and text cues. EURASIP JASP 2, 170–185 (2003)
31. Naphade, M.R., Huang, T.S.: A probabilistic framework for semantic video indexing, filtering, and retrival. IEEE Trans. on Multimedia 3, 141–151 (2001)