



Personalized Video

Krishnan Ramanathan, Yogesh Sankarasubramaniam, Vidhya Govindaraju

HP Laboratories
HPL-2011-176

Keyword(s):

Personalization; Online Video; Multimedia retrieval

Abstract:

Despite the tremendous growth in online video, the consumption interfaces are still modeled on search-and-browse. We propose a solution to create a much richer online video experience modeled on the television channel metaphor. This is especially useful in the emerging market context where people are more likely to consume the video web compared to the text web. Our solution allows users to consume online videos from multiple sources in a personalized manner. We have implemented a number of interactions to facilitate easy consumption such as personalized recommendations, query assistance, similar videos, documents related to videos and video annotations. We believe our solution enables a differentiating online video experience.

External Posting Date: October 6, 2011 [Fulltext] Approved for External Publication
Internal Posting Date: October 6, 2011 [Fulltext]

Personalized Video

Krishnan Ramanathan, Vidhya Govindaraju, Yogesh Sankarasubramaniam

HP Labs India

ABSTRACT

Despite the tremendous growth in online video, the consumption interfaces are still modeled on search-and-browse. We propose a solution to create a much richer online video experience modeled on the television channel metaphor. This is especially useful in the emerging market context where people are more likely to consume the video web compared to the text web. Our solution allows users to consume online videos from multiple sources in a personalized manner. We have implemented a number of interactions to facilitate easy consumption such as personalized recommendations, query assistance, similar videos, documents related to videos and video annotations. We believe our solution enables a differentiating online video experience.

1. Introduction

Video content on the web is exploding rapidly (users upload more than 48 hours of video every minute on Youtube). Consumers are shifting video media time from television to online video. Online video sites allow consumers to upload user generated content and share it with other consumers. Video is also the second most popular application on the Internet (after Facebook) in terms of Internet minutes spent. Widespread adoption of broadband and wireless 3G services allow high video streaming throughput. Newer device form factors such as tablets (e.g. Apple ipad) are seeing increased video consumption due to the affordances (such as watching videos when lying down).

Although the content on online video sites is burgeoning, the online video navigation and consumption experience is similar to web browsing and very unlike the channel metaphor users are comfortable with for television viewing. Users have to search for online videos on multiple sites even if it is about the same topic. The interactions offered vary from site to site and the user has to learn multiple interfaces. Because the online video download happens in real time over a thin pipe, online video sites reduce the resolution of the videos for acceptable download speeds. In spite of this, video download is often slow and jitter-ful leading to a bad experience. Thus the online video experience today is very dissimilar to the jitter-free video experience on television where content is organized as channels and the user has an easy interface to skip to other content.

Online video has a special relevance to emerging markets like India. Despite many efforts, the number of Internet users in India is experiencing slow growth. One of the main reasons is availability of text content; creating text content in local Indian languages is quite challenging because of the diversity in scripts and different keyboards required thereof. The levels of literacy required to query, navigate and consume the text web is also quite high. This lack of text content and complexity of navigation has limited the value proposition of the Internet (and hence computers) for consumers in these markets. Video doesn't suffer from these problems. Creation of video content doesn't require specialized hardware such as local language keyboards. The growth of mobile phones with video cameras has allowed users to create video content on the move. The high penetration of televisions in India is indicative of the low literacy requirements for consuming video; people with minimal screen literacy can operate a television [1].

We believe there is an opportunity to substantially improve the online video consumption experience both in the enterprise as well as in the consumer space. There is a wide range of devices (desktops, laptops, handhelds and television) on which online video is consumed.

2. System overview

We have solved the problem of making online video easier by creating a personalized client side interface for consuming online video. Figure 1 shows the architecture of the solution. The main innovations are: Channel

metaphor for video consumption, a personalized recommendation engine that allows videos to be sourced from multiple online sites, a user interface with touch and gesture support and an interaction engine that provides a rich experience, viewing similar videos, recommendation of text content related to video, video prioritization based on user feedback and video annotation and playback.

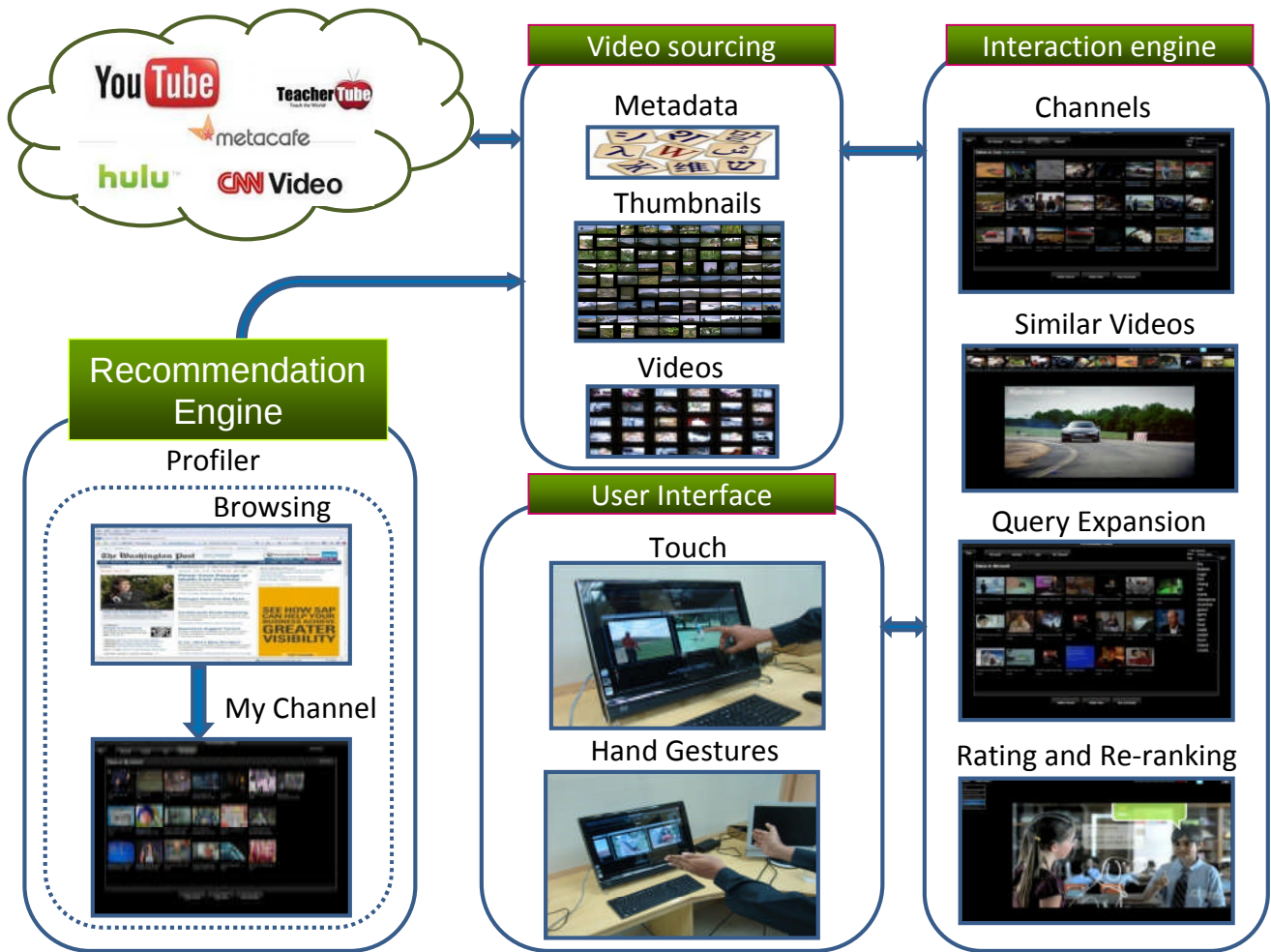


Figure 1 High level architecture of Personalized Video

Video meta-data is sourced from multiple sources. This meta-data includes the title of the video, the description, tags, user comments etc. The user interface could be a keyboard/mouse, touch or gestures depending on the capabilities of the device. The interaction engine incorporates various interactions that a user can perform while watching the video. The recommendation engine builds a user profile based on the videos viewed by the user across different channels and attempts to recommend related and adjacent videos with the goal of expanding the users interests. The videos are currently streamed from the source (e.g .Youtube) but could be cached to provide offline access. We have implemented the application using the Windows Presentation Framework (WPF). The application can be downloaded from [9].

We now describe the various features of the Personalized Video application.

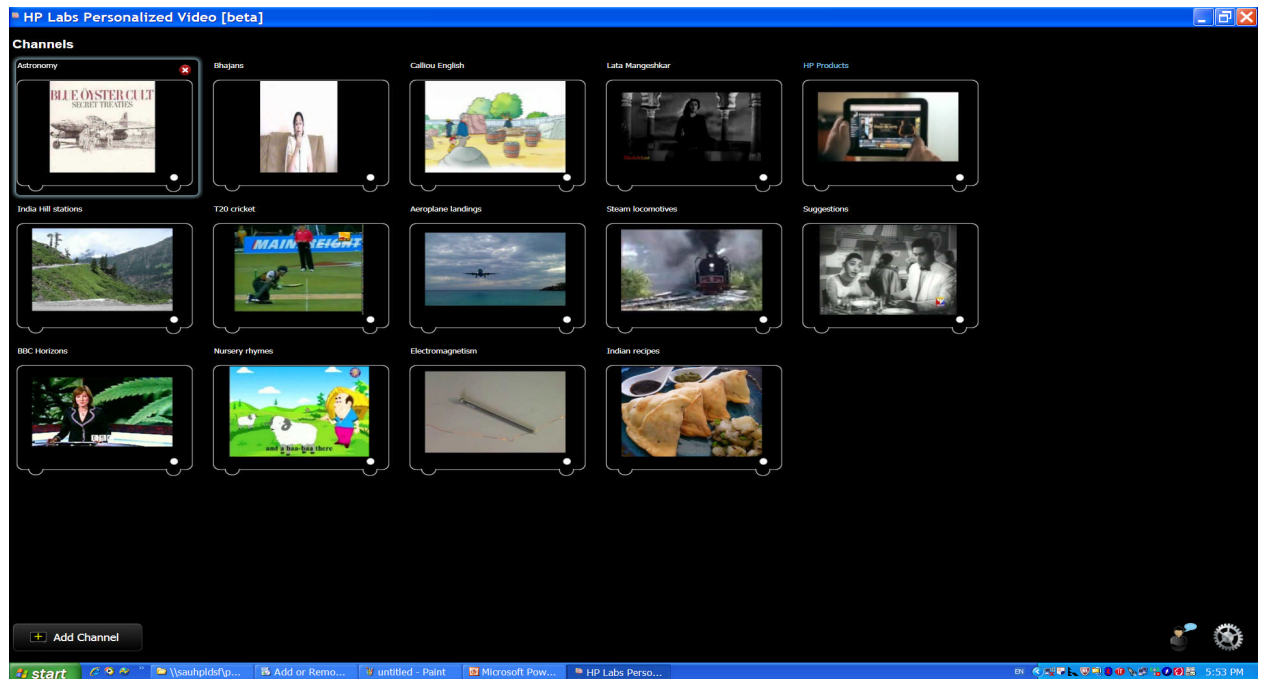


Figure 2 View of the channels within the application

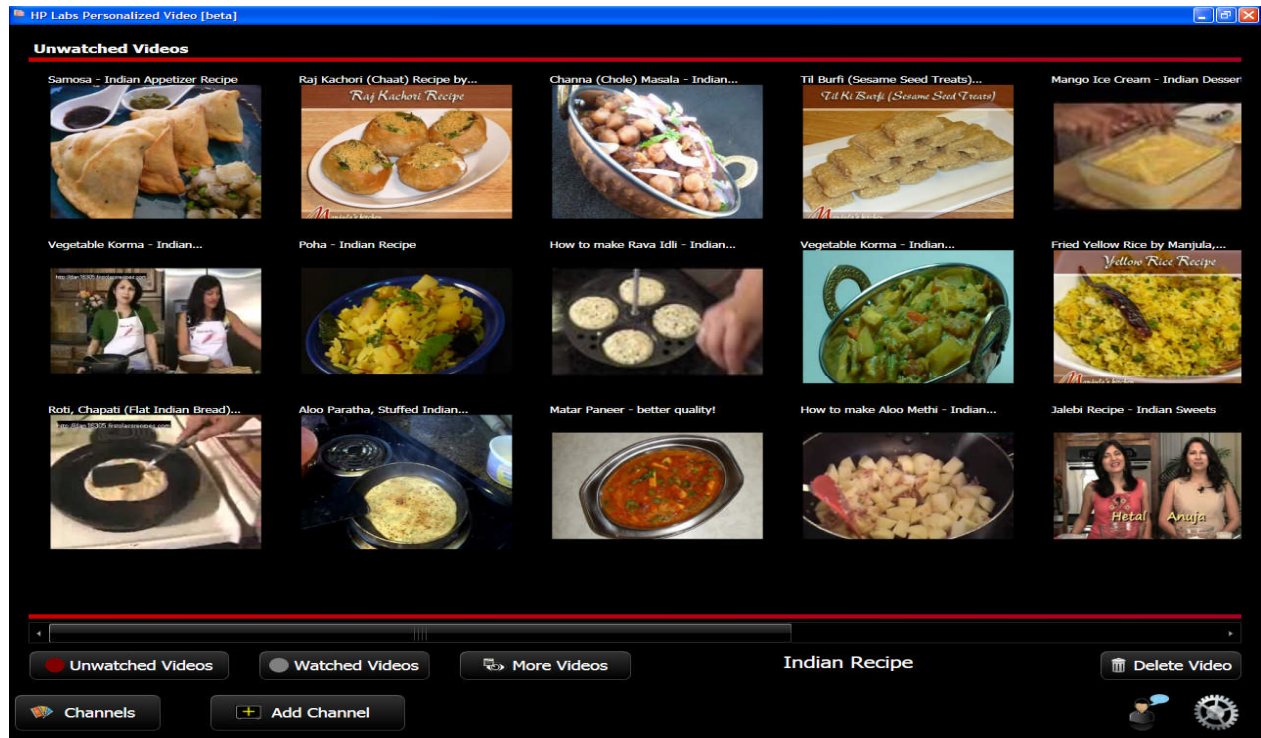


Figure 3 Videos within a channel

Leanback online video consumption using the channel metaphor - Our solution aggregates videos from multiple online video providers using a topic based “channel” metaphor. A channel is any longstanding interest the user has, videos are pushed into the channel without requiring users to make the same query again. Videos are cached on the device to avoid network jitter and offline consumption. Figure 2 shows the application with channels such as “Indian recipes”, “Steam Locomotives, and “Nursery rhymes”. Clicking a channel pulls up the list of videos associated with the channel (figure 3).

Personalization – This is accomplished by video recommendation based on user profiles and using user feedback to prioritize videos within a channel. The video recommender constructs the user profile by extracting keyphrases from the metadata of the videos (title, tags, description), weighted based on user rating, number of views, and other parameters. The recommender sources videos using this profile, and makes it available in the “Recommended” channel. Within a channel, users are allowed to rate videos as “like” or “dislike”. We then use the Naïve Bayes algorithm to predict the rating on the unseen videos in the channel. The videos in the channel are then re-ordered based on this predicted rating.

Group consumption of videos – We have integrated a face recognition system and group recommender to recognize multiple users and provide an ambient viewing experience where the content is interesting to all the viewers (people identified with face recognition). Thus, relevant channels and videos are automatically populated and adapted according to the current user group. This is particularly useful in an Indian household setting where the entire family has access to a single shared device.

Input methods – We have integrated the application with the Gesture keyboard (developed at HP Labs India) that allows input in Indian language (e.g. Hindi), see figure 4.

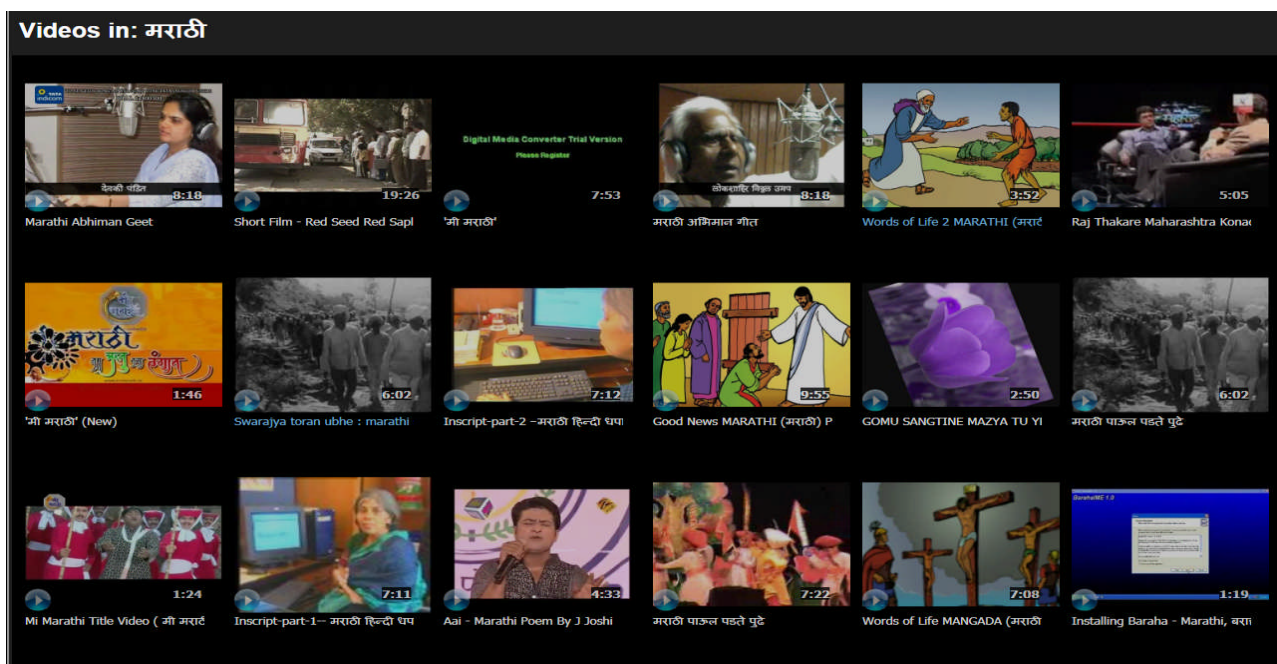


Figure 4 A channel in Marathi language

We have also implemented an application on Symbian that allows a Symbian phone to be used for input at a distance (e.g. entering channel names). For video navigation, the application allows hand gesture navigation (in addition to touch). We currently support six gestures for playing the videos, navigating to next/previous videos and for zoomin/zoomout videos in a channel. Users can also provide feedback on the videos (like or dislike) with thumbs up/down gesture. We have also tested the user experience with the Logitech air mouse.

Integration with Facebook and Twitter – Users can share a video with friends on Facebook and Twitter with customized messages.

Query assistance – Users are suggested additional querying words after they enter a channel. This is performed using pseudo-relevance feedback. The video source (e.g. Youtube) is queried with the channel name. The high frequency words and key phrases from the results are extracted and presented to the user (figure 5).

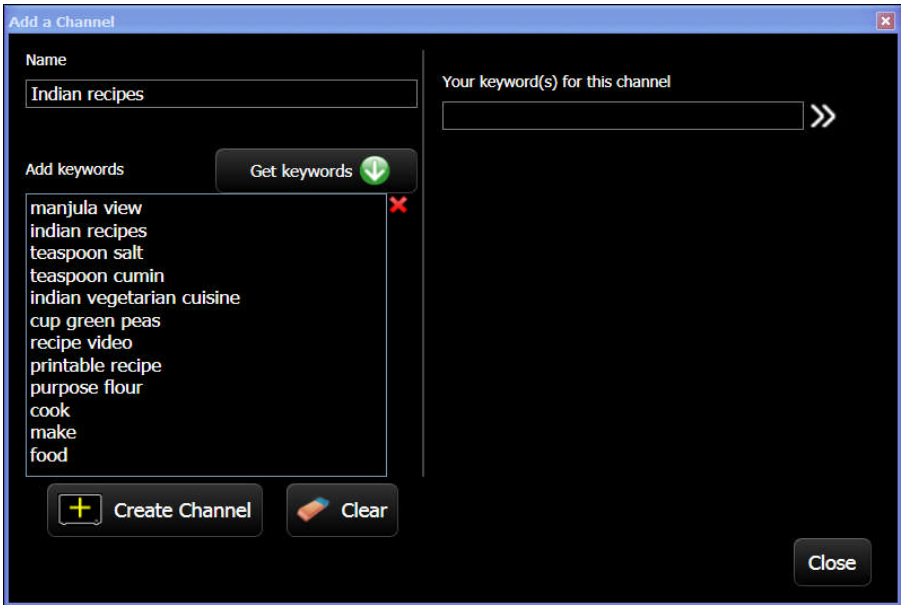


Figure 5 The query assistance interface

Similar videos search – Users can request to watch videos related to the current video using the “Similar videos” feature (see Fig. 6). This works by first ranking the keyphrases (in a similar manner as was done for the personalization feature described earlier), querying video engines with the keyphrases, and then assessing the similarity of sourced content. The same method can be used to also source related textual content. For e.g. a culinary video can point to the corresponding recipe, which the user can download or print.



Figure 6 The “Similar Videos” interface

Video annotation – The system allows the user to annotate videos. The user can mark points in the video with text annotations and visit those at a later time. This can be used for revisiting parts of a lecture or a procedure in a recipe (figure 7).



Figure 7 Interface for Video annotation

3. Technologies

The Personalized Video application incorporates a number of technologies including personalization technologies, similar content retrieval and Video prioritization using user feedback. In this section we outline some of the research and technologies that went into the application.

3.1 User Profiling

User profiling is used in the Personalized Video application for recommending videos based on user tastes. We started out using the profiler based on web pages [4] but eventually decided to create a user profile based on the videos the user had seen in the user created channels.

We elicit the metadata information available for the videos to map the features of a video to the interests of the user. Key phrases from video metadata represent the interests of the user. The key phrases are weighted based on multiple factors like the corresponding video ratings, days since the video related to the interest was last seen, the ratio of number of videos watched to the number of videos recommended based on the key phrase. The key phrases are then used in a recommender system to fetch video recommendations. The user profile is constantly updated based on the user behavior.

3.2 Group recommendation

We have built a group recommender to automatically source and display video content that is of interest to all the viewers. First, users are identified by a face recognizer. Then, a joint user profile is constructed as follows: individual interests are first expanded into representative topics (key phrases from video metadata), and then a reciprocal rule is used to combine/discard topics, which has the effect of jointly satisfying all the users. Videos are then sourced and presented based on the joint profile. Our approach allows a strong discovery of joint interests, rather than merely identifying matches in the individual profiles. Thus, even when the individual profiles do not show direct overlap, our solution can identify joint interests among users.

A study on subgroups of 30 users indicated that the retrieval precision is around 70%; users are also very satisfied with the quality of recommendations. More details on the algorithm and implementation, along with examples, can be found in [3]. Our approach is also of interest in other group scenarios such as shared printing, group advertizing, online group reading etc.

3.3 Similar videos and related content search

In the Personalized Video application, the users can request to watch videos similar in context to the current video. The algorithm extracts key terms from the video metadata and queries a video source to fetch similar videos. The same method is used to source related text content. The key terms extracted from the video are used to query a text source to fetch similar text content. More details of this algorithm are provided in [7].

Sometimes the related videos can slightly differ in topic from the original video. While this may be unacceptable in a traditional search system, it offers a way of diversifying and discovering new content in a video consumption application like Personalized Video. Slaney [7] points out that the main difference is that in multimedia search, people are looking for entertainment and not information. Hence, users are willing to even click on results with low precision if they are entertaining enough.

3.4 Video Prioritization using user feedback

In the Personalized video channel, users can rate videos based on their interest as “like” or “dislike”. The ratings are then used to prioritize videos so that the videos of interest to the user are ranked higher. We devised two algorithms for video prioritization. In the first interest-video rank algorithm, the metadata information of the videos is used to map the interest of the users. Key terms (referred to as concepts) from the video which includes both key phrases and keywords are used in a ranking function to get a ranked list of videos based on user ratings. A bipartite graph is created between the videos and the concepts. The weight on the videos (rating of the videos) is transferred to the concepts based on a weighing function and the weight from the concepts is transferred back to the videos. The algorithm outputs a ranked list of videos.

The second algorithm is based on Naïve Bayes Classifier. The Naïve Bayes Classifier used in our re-ranking algorithm trains n -bins (n corresponding to the number of possible ratings for a video, in our case it is 2 – like and dislike) based on user interactions. The features extracted from an video rated i are used to train the i^{th} bin. The classifier then predicts a rating for unrated videos based on the features extracted from the video. All the words from the title, tags and description of a video are taken and are used to train the classifier. In the Personalized video application, we currently use this algorithm to prioritize the videos because of its faster execution speed.

4. The User Interface

Figure 8 shows the elements of the user interface for personalized video. The interface allows users to switch between new and watched videos. It allows users to create and delete channels as well as video metadata. It also allows interactions to get similar videos, provide feedback based ranking, get web pages related to the video and share videos with Twitter and Facebook friends. We have integrated the gestures with the application using the Microsoft Kinect API and Kinect camera. This allows users to navigate videos from a distance. We have also integrated the speech interface that comes with Kinect to allow users to navigate videos using speech commands.



Figure 8 Personalized Video User Interface

Personalized video was designed to be a client-cloud solution. We have also implemented a mechanism to create a channel on the cloud that is pushed to all clients of personalized video. This can be used to broadcast content (and perhaps advertisements) to client devices.

5. Related work

There are a number of attempts to make the online video experience richer. CoolIris is a Firefox plugin that displays result thumbnails on a 3-D wall. Timetube allows users to navigate Youtube videos based on the video upload date. Apple TV allows users to stream photos and videos to a TV, supports the h.264 and Mpeg-4 codecs and has a remote control interface. Boxee is an open source software media player that allows streaming web video to TV. Boxee also has social networking features and allows people to see what their friends are watching. Miro is an open source player that allows subscribing to RSS videos, searching video sites and downloading videos.

Acknowledgements

We would like to acknowledge the efforts of the talented engineers at HP Labs India who are responsible for creating a ship-worthy application. Specifically, we would like to thank Gali Prasad, Jnana Prakash Rana and Swetha Reddy.

REFERENCES

1. Marti Hearst, The decline of text, http://edge.org/q2009/q09_9.html#hearst
2. <http://www.bothsidesofthetable.com/2010/10/19/the-future-of-television-the-digital-living-room/>
3. Sankarasubramaniam, Y.; Ramanathan, K.; Lahiri, A.; Govindaraju, V. (2010). Group recommendations using Joint Profiling. HPL-2010-126. HP Labs Technical Report
4. Krishnan Ramanathan and Komal Kapoor, Creating user profiles using Wikipedia, ER 2009 conference, LNCS Volume 5829/2009, 415-427, DOI: 10.1007/978-3-642-04840-1_31, Springer Verlag.
5. C.G.M. Snoek and M. Worring, Concept based video retrieval, Foundations and trends in information retrieval, vol. 2, no. 4 (2008), pp. 215-322.
6. Loeb, S.: Architecting Personal Delivery of Multimedia Information. Communications of the ACM 35(12) (1992), pp. 39-48.
7. Vidhya Govindaraju and Krishnan Ramanathan, Similar Document Search and Recommendation, HPL-2011-150, HP Labs Technical Report.
8. M.Slaney, Precision-Recall is wrong for multimedia, IEEE Multimedia, July-September 2011.
9. Personalized Video download, <http://www.personalized-video.com>