The Participation Payoff: Challenges and Opportunities for Multimedia Access in Networked Communities

Naeem Ramzan
Multimedia & Vision Group
School of Electronic Eng. and Computer Science
Queen Mary, University of London, United Kingdom
naeem.ramzan@elec.qmul.ac.uk

Martha Larson
Multimedia Information Retrieval Lab
Elec. Eng., Mathematics and Computer Science
Delft University of Technology, Netherlands
m.a.larson@tudelft.nl

Frédéric Dufaux
Multimedia Signal Processing Group
Institute of Electrical Engineering
Ecole Polytechnique Fédérale de Lausanne, Switzerland
frederic.dufaux@epfl.ch

Kai Clüver
Communication Systems Group
Technische Universität Berlin, Germany
cluever@nue.tu-berlin.de

ABSTRACT

Increasingly, multimedia collections are associated with networked communities consisting of interconnected groups of users who create, annotate, browse, search, share, view, critique and remix collection content. Information arises within networked communities via connections among users and in the course of interactions between users and content. Community-derived information can be exploited to improve user access to multimedia. This paper provides a survey of techniques that make use of a combination of three information sources: community-contributed information (e.g., tags and ratings), network structure and techniques for multimedia content analysis. This *triple synergy* offers a wide range of opportunities for improving access to multimedia in networked communities. We focus our survey on three areas important for multimedia access: annotation, distribution and retrieval. The picture that emerges is promising: information derived from the social community is remarkably effective in improving access to multimedia content, and participation in networked communities has a high payoff.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**] H.3.1: Content Analysis and Indexing – *indexing methods*. H.3.3 Information Search and Retrieval H.3.4 Systems and Software – *distributed systems.*

General Terms

Algorithms, Design, Performance

Keywords

Social Media, Tagging, Content distribution, Multimedia retrieval

1. INTRODUCTION

The rapid growth in Web technologies has changed the role of individuals from passive content consumers to active content creator/producers. Prominent examples of social multimedia communities include Flickr (http://www.flickr.com/) and You-Tube (http://www.youtube.com/). The unprecedented popularity of these applications lend credence to the claim that in a few years *everyone* will be a multimedia content mediator and a multimedia content producer in addition to being a multimedia content consumer.

The on-going evolutionary trends of the Web, often associated with the term 'Web 2.0', emphasize community, collaboration and the rise of distribution platforms that deliver content directly [1]. As collection sizes grow larger, the need for new multimedia access technology becomes evermore urgent. This need translates into a pressing demand for exploitation of the full range of possibilities offered by recent developments in social multimedia. Users require access to ever-growing collections of multimedia to be fast, if not instantaneous, and they expect video retrieval to be as easy and effective as text retrieval.

This paper presents a survey of techniques that can be used to realize effective exploitation of the potential of networked communities for the purposes of multimedia access. We focus on three major areas of research related to multimedia access: annotation, distribution and retrieval. The context of this paper is PetaMedia (http://www.petamedia.eu/), a Network of Excellence project funded by the European Commission. The concepts presented in this paper have been formulated and refined by exploration, investigation and research effort within the PetaMedia consortium. Specifically, the goal of PetaMedia is to promote a new paradigm of multimedia research that is based on a *triple synergy* constituted by three information sources:

- Annotations contributed by users (e.g., tags, ratings, comments)
- Network structures (social networks & P2P networks)
- Multimedia content analysis

New techniques for multimedia access and retrieval that simultaneously derive benefit from all three of these information sources are currently being developed at research sites not only within the PetaMedia network, but also across the globe.

The paper is structured into three sections examining three different areas important for multimedia access: annotation, distribution and retrieval. For each of these areas, combination of conventional techniques with information derived from networked communities presents new potential for advancing the current state of technology. Opportunities and challenges are discussed for each area and examples of work that is currently ongoing are also provided.

2. ANNOTATION

Networked communities engage in tagging activities. Tagging involves assigning a word or a group of words to a multimedia item that captures characteristics of that item. Users are free to choose the tags they want to use and to invent new tags when the necessity arises. Tagging behavior gives rise to a non-hierarchical structure, often referred to as a folksonomy, that supports efficient search and retrieval by bridging the semantic gap. Users' motivations for tagging are not always entirely transparent, but it is often assumed that tags are assigned for the purposes of re-finding items later. In a social community, it is possible that tagging is motivated by more altruistic considerations. For a recent overview of collective tagging, also called collaborative tagging, see [2]. Here, the assumption is made that tags have a function for access within the social community, and we describe methods for generating new tags. Often cited disadvantages of using tags over using a closed list of keywords is the lack of control over synonyms, homonyms, and normalized forms (e.g., singular versus plural). Although some techniques have been proposed to overcome these limitations, it is commonly recognized that the most useful tags will emerge if enough users tag the same content. Automatic methods offer solutions to increase the number and improve the quality of tags characterizing objects in a collection. Theses methods are discussed in the remainder of this section.

2.1 Tags derived from audiovisual content

Tags can be considered to be a form of multimedia annotation, and automatic metadata generation methods can be used to create them automatically. The simplest source of automatically generated tags is the context information (e.g., time and date stamp) stored by the device that created the content. The value of contextual features and surface video features is discussed in detail by [3]. Even when more sophisticated techniques are applied, such features should not be discarded or overlooked.

The prohibitively high cost of detailed manual annotation has motivated the development of powerful multimedia content analysis and classification techniques. These techniques provide a basic method for attributing semantic values to audiovisual data. A broad array of technologies is capable of automatic metadata generation: automatic content segmentation and clustering, audio analysis, speech recognition, object detection, face recognition, gesture recognition and music analysis. The output of each of these techniques can be encoded in the form of a tag and used to annotate multimedia content. Techniques for automatic tag generation all face the challenge of the semantic gap [4] – the mismatch between the information carried directly in the multimedia signal and human notions of meaning that give rise to user search criteria.

In the case of multimedia with an audio or speech track, semantic information can automatically be extracted by using audio analysis and speech recognition techniques. Speech transcripts are a valuable source of tags, even when the speech conditions are less than ideal and word error rates are higher than optimal levels [5]. Rich transcription [6] techniques make it possible to mark speaker turn boundaries and audio events such as laughter and applause. Text classification and retrieval techniques can be applied to speech transcripts in order to automatically generate subject labels that can be used to tag videos, cf. [7].

Techniques for detecting concepts in video for the purpose of multimedia retrieval are becoming increasingly more advanced [8][9]. In order to build a system to detect – and thereby annotate – semantic concepts in multimedia data, the basic approach is to manually label training data and optimize a classifier that makes use of low-level features extracted from these training data. Extracting the same features from unlabeled data, the concept detector will then perform an automatic annotation. The usefulness of this – necessarily imperfect – classification depends on the application, and on whether the semantic concepts chosen for training are appropriate. Improvements in annotation accuracy may be obtained, for example, with use of more sophisticated features and/or classifiers. Introducing local or temporal context in the analysis, i.e., also analyzing neighboring objects or temporal progress before making the final decision, forms a further enhancement of the annotation results [10][11]. In addition to concepts, visual classifiers can be trained to identify characteristics of video such as topic and genre [12].

The drawbacks of the basic approach are the amount of labeled training data necessary for a reasonable accuracy and the inflexibility of the classifier in view of huge quantities of multimedia data and varying sets of concepts expected by the users. Features that can be more directly extracted from multimedia – for example via clustering – are useful for structuring, navigation and browsing, but are difficult to relate to user queries, which typically encode abstract and context-dependent content needs. These drawbacks may be mitigated using efficient modeling and clustering methods and employing a cumulative training algorithm [13]. Leveraging key effects within the multimedia signal is another promising approach [14]. The benefit of combining visual and audio information is well established in the area of video retrieval. Here, multimodal approaches range from exploiting face recognition and speech transcripts, a combination which has been used since the early days of news retrieval [15], to more recent approaches that fuse modalities using reranking [16].

Finally, we would like to mention the growing awareness in the multimedia content analysis community of the importance of annotations that encode affective information. Multimedia content analysis techniques exploiting relatively low-level features have been shown to be effective for this purpose, cf. e.g., [17]. Also related, is work on video classification that makes use of connoted visual codes [18]. The output of the approach is a class label that can be used as tag for annotating the video. The connoted-visual-code approach directly targets the capture of emotional meaning, e.g., romance or longing, that is associated with what is visually depicted in the content, e.g., a sunset.

Other methods for automatically generating tags exploit information that is generated by the user community and we devote the remainder of this section to discussion of these techniques.

2.2 Using tags to create new tags

A relatively straightforward method of generating new tags is to base them on previously existing tags, e.g., tags derived from previously performed content-based tagging or assigned by users in the community. Although these annotations may be sparse and noisy, they are invaluable for searching and browsing and they can also considerably help to improve the accuracy and efficiency of multimedia content analysis. Combining the annotations with the results of a feature extraction from the content itself yields the starting point for tag propagation, i.e., attributing identical annotations to similar multimedia data [19][20][21].

As mentioned above, shortcomings associated with user-assigned tags include inconsistencies caused by the use of synonyms or hypernyms, or other ambiguities. To address these shortcomings, hierarchical word databases can be employed for a semantic analysis of the text tags. The choice of a suitable similarity measure may be critical, and depends on the task given [22]. In order to additionally be able to process named entities used in the textual annotations, large encyclopedic corpora may be searched for hypernyms [23]. In [24], information about user tagging behavior and about tags assigned within the collection of items is used in order to detect tag synonyms.

Exploiting existing tags to generate additional tags involves a loop that includes both manual and automatic processes. A particularly promising way to realize that loop involves applying relevance feedback with the aim of further refining or enhancing existing annotations. In such a loop, automatic tag recommendations [25][26] are presented to users, who then have the option of either accepting or declining them.

2.3 User incentives for tagging

A direct method in order to increase the number of tags in a multimedia community is to offer users incentives in order to entice them to tag content. In the context of a networked community, incentives can be based on the users' own drive towards social interaction. The range of motivations that prompt users to tag is wide. For example, users may want to express value judgments and share them with others, to identify themselves by tagging items in they appear with their names, or to attract other users' attention to their content (cf. also [2][27]).

A popular approach to motivate users to tags makes use of play- and competition-driven methods. This approach encourages users to perform a task that is challenging for computers, e.g., multimedia tagging, by formulating it in the form of a game. For instance, random users are paired together to find agreement on semantic image labels in [28]. Other examples are based on the same principle. Peekaboom is a game for object recognition [29] and Phetch is for image retrieval [30]. In [31], the incentive to use a region-based labeled database of images promotes participation of the researchers in the labeling process itself. PodCastle [32] offers functionality that lets users correct speech recognition transcripts. Although this functionality is less conventionally game-like than other applications, users do make corrections. The corrected transcripts are used as metadata annotations to index podcasts, but are also used in order to retrain the acoustic models. As a result, the speech recognition system that produces the transcripts is continuously improved.

Explicit tagging by users may potentially lead to high quality semantic metadata. However, it requires a sufficient large number of participating users in order to be viable, which in turn demands strong incentives. An alternative to motivating users to generate more tags is to use user behavior indirectly as a method for tag generation. We now turn to discussion of examples of such techniques.

2.4 Implicit tagging

Users generate implicit information when they interact with each other and when they interact with content in a multimedia collection. Implicit tagging is an important tool to create, propagate and evaluate annotations. The method offers an alternative to traditional tagging, which requires an explicit action by subjects and is helpful to tackle the challenge of creating tags for huge volumes of un-annotated data.

One method of implicit tagging involves the behavior of users while they watch multimedia content. For instance, clues about interest, emotional state, agreement, focus of attention, can be obtained by analysis of facial expression, gaze direction, head and body posture, as well as acoustic signals. More specifically, the combination of these low level features, in conjunction with machine learning techniques, may reveal high level information about the affective state of the user, including interest, frustration and boredom. In [33], acoustic and visual sensors detect laughter when a user is watching multimedia content, enabling to label the data with tags.

Similarly, it is also possible to measure physiological signals from the user. In [34], features such as respiration, Galvanic skin resistance, skin temperature, eye blinking rate, electromyogram and blood flow are measured to obtain emotional tags. An electroencephalogram-based brain-computer interface system is proposed for implicit emotional tagging of multimedia content in [35]. More precisely, the system analyzes the P300 evoked potential recorded from user's brain to assign an emotional tag to a given video clip. EEG signals are also valuable for the task of checking the appropriateness of existing tags. In [36], subjects are presented with both a video and a tag and the EEG modality is used for the purpose of tag validation, i.e., to determine if the tag is congruent with the video.

2.5 Tagging and the triple synergy

The *triple synergy* of user-contributed information, user interactions within the social network and multimedia content analysis throws open new and promising opportunities to improve the quantity and quality of multimedia annotations in the form of tags. Going forward, we anticipate multiple challenges that must be faced. First, although much work has been devoted to multimodal combination of information from different sources, more research is necessary to develop methods that are generally applicable or can be easily optimized in particular use scenarios. A prime example of creative combination is the work presented in [37], which uses information on how people position themselves with respect to each other when pictures are taken in social settings to support automatic understanding of image semantics. Second, additional research effort must be devoted to the challenge of translating affective reactions of individual users at a given moment into tags that are stable enough across time and across users to be useful. Issues relating to the convenience of sensor equipment and also of user acceptance of systems that record their emotional reactions will continue to be important in the area of implicit tagging. Third, much information can be derived from user comments and reviews. Currently, these sources are underexploited for tag generation. Finally, it should be kept in mind that tag systems have developed on the Web are governed by a set of conventions that emerge as a product of interactions between humans and media. Automatically generated tags must fit within the system in order to be useful to users, cf. e.g., [38].

3. DISTRIBUTION

Multimedia content distribution systems must necessarily cope with the complexity of handling the different types of networks, service platforms and the different types of end user terminals that are available on the market now and in the near future. This section describes techniques that make it possible to split the functionality of content distribution over the different parties involved in the distribution process. Peer-to-peer (P2P) content distribution is the most renowned application area of P2P systems and it contains file sharing systems, e.g., Gnutella, distributed storage applications, and content delivery networks. These applications offer the possibility to publish, store and exchange files and other digital/multimedia content. Recently, increasing amounts of attention have been paid to distribution in P2P systems, cf. e.g., [39] [40].

Multimedia content is created as a result of intellectual activities in industry as well as in the scientific, literary and artistic fields. The distribution of this content is usually protected by copyright. Copyrights are intended to safeguard, human creativity by providing incentives to creators with assurance that their work can be disseminated without the fear of illegal copying. Nowadays, not only industries, but also consumers, are increasingly using the Internet as a medium for the distribution of content [40]. The increasing demand for digital content threatens to overwhelm the infrastructure of online content providers. An attractive approach for commercial online content distribution is the use of P2P protocols. This approach does not require a content provider to overprovision its bandwidth to handle peak demands, nor does it require the provider to purchase service from a third party. Rather, a P2P protocol such as BitTorrent (http://www.bittorrent.com/) harnesses its clients' bandwidth for file distribution, and saves the bandwidth and computing resources of a content provider. For this reason, copyright issues are significant for content distribution over P2P networks.

Application communities perform selection, targeted collection, development and exploitation of content using existing solutions that allow:

- Collaborative distributed environment for sharing/editing and enhancing media content
- Access and acquisition to new information
- Efficient management of resources, security and intellectual property rights handling
- Facilities to leverage human and cultural knowledge
- Support for the creation of user communities that can access a wide range of information independently of their location or technical constraints

Thus, overall evaluation of content distribution in distributed settings includes technical, end user and socio-economic evaluation, with a focus on both the social and also the economic payoffs.

Various existing and proposed solutions for content distribution in P2P networks have been investigated. The most important challenges of content distribution systems are discussed in more detail in the remainder of this section.

3.1 Content Adaptation and Scalability

The ability to adapt content appropriately for the context is a key challenge of content distribution in P2P and distributed communities.

Digital content must be constantly and transparently adapted to user preferences and terminal characteristics. Adaptation can be achieved via selecting the appropriate media parameters and modalities based on information concerning the context. A conventional video coding system encodes a video sequence in a desired, fixed bit-stream that is adequate for a given application. For this reason, serving different clients requires transcoding of a given video sequence, which is not computationally efficient and may reduce video quality.

Another solution lies in encoding of the content into many versions, each aimed for delivery to a certain group of receivers having similar decoding and display capabilities, which are then stored at the video server. As those two main traditional approaches do not provide at the same time a low-cost adaptation and rational storage requirements, a need for a new technology is evident. Scalable Video Coding [41]-[44] provides a straightforward solution for a universal system for video coding that can serve a broad range of applications.

In a Scalable Video Coding system, the adaptation of the video bit-stream is done in a low-complexity fashion, by simple bit-stream parsing. Since the adaptation is performed by means of the scaling of compressed video parameters, the scalable bit-stream has to be encoded in such way that the bit-stream parts are hierarchically encoded according to these parameters. Basic types of scalability or adaptation parameters are spatial (resolution), quality (often referred to as signal-to-noise ratio scalability, or SNR scalability), temporal (frame rate) and combined scalability.

3.2 Performance and Persistence

Another important challenge facing P2P content distribution systems is ensuring balance between upload and download data for each peer. One of the metrics used to measure the performance of a P2P system is the time required to perform a search and retrieve the requested multimedia content. Additionally, balance can be maintained by implementing a fairness policy that enforces users to offer and consume resources in a fair and balanced manner [45]. Various P2P content distribution systems aim to achieve this by giving incentives to users to share as much as possible of their upload bandwidth. The ratio of uploaded versus downloaded data chunks is a metric utilized for this purpose for instance by the BitTorrent network to prioritize the users.

3.3 Secure Distribution of Contents

The secure distribution of multimedia content is an important challenge in networked communities. The main focus on the secure distribution of contents is privacy and confidentiality by ensuring that the digital content is accessible only to authorized user. Unauthorized entities cannot change data; challengers cannot surrogate an imitation document for a requested one [46].

The secure availability and persistence of digital content and associated assets to authorized user is also an important aspect. It gives stability in the presence of failure or changing node populations.

3.4 Quality of Experience

The received video quality as perceived by the end user is the ultimate goal of the content distribution in networked communities. The received video, apart from compression artifacts, might suffer quality degradations due to network behaviors, such as packet losses, jitter and delay. In this scenario, the presence of a video quality monitoring system that is able to inform the service provider upon the quality of the delivered data could be very useful. Based on the quality feedback, the coding strategies, as well as the transmission strategies, may be optimized in order to maximize the end user experience.

The classical approach to evaluate video quality consists in resorting to subjective quality measurements performed by trained human experts who rate the overall quality of video, the obvious limitations of the subjective method being the important resource involvement and the inadequateness for systematic large scale and automatic application. An alternative is to use metrics such as Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR), which unfortunately do not in general correlate well with subjective judgments.

Objective methods were therefore augmented over the last ten years by models representing the Human Vision System (HVS) taking into account various phenomena according to the targeted video processing and application context. The incorporation of HVS models into the quality evaluation of video processing gave rise to the notion of Quality of Experience (QoE) [47][48].

With perceptually-based video quality metrics, it is possible to perform automatically an objective evaluation of video quality that reflects the human appreciation of the processed and visualized data. The QoE video quality measurement is applicable to evaluate the relative quality of a video when compared to an original signal (i.e., Full-Reference system, FR), or the absolute (or intrinsic) quality of a video, in which case the evaluation is performed without any reference (i.e., No-Reference system, NR).

3.5 Distribution & the triple synergy

In the case of content distribution, the triple synergy of user contributed inactions, user interactions within the social network and multimedia content analysis presents opportunities, but also a series of challenges. The exact methods by which these information sources can be exploited in order to determine the encoding of specific content that will promote system scalability remains to be explored in depth. If content delivery is to be adapted to user preferences and terminal characteristics, effective use of content analysis, metadata and also the structure of the distribution network is critical. Finally, research and development that directly targets the quality of the viewer's experience will ensure that new video distribution technologies are suited to optimally fulfill the needs of individual users within the networked community.

4. SOCIAL CONTENT RETRIEVAL

Social media can be characterized as 'collection goods, produced through computer-mediated collective action' [49], but social media is perhaps best understood by contrasting it with conventional content. A conventional media collection has a well-defined existence independent of those who create, curate, annotate and use it. Social content has an inherent interdependence with the people who interact with it. If YouTube suddenly gained a radical influx of new users, the composition of its content potentially could be fundamentally altered – not so with a conventional film archive, whose content and cataloguing system does not depend on its users. In a book written for the popular press [50], Shirky identifies 'the linking of symmetrical participation and amateur production' (p. 107) as the factor that distinguishes the currently ongoing emergence of social media, from other historical developments in communication tools. Social content collections are inherently dependent on the networked communities that create and curate them. This section discusses the ways in which interactions between members of networked communities can be exploited to improve the retrieval of social multimedia content.

In this context, we refer to 'social content retrieval' (or 'social content search') rather than 'social media retrieval' in order to underline that we are discussing retrieval of social multimedia content, such as is uploaded on YouTube and Flickr, and not discussing other forms of search in social networks or community collections. Among the forms of search we exclude are the following: people search in Facebook, product/review search on Epinions.com, event search on eventful.com or fact search on Wikipedia. Further, it is important to differentiate social content retrieval from social search, in which the search activity itself and not necessarily the target collection is social in nature. This sort of social search, which we do not consider further in this context, involves multiple searchers who communicate with each other during the search process, cf. e.g., [51].

Social content search is a young and rapidly growing field. Social media is attracting growing attention, as witnessed by a number of recent workshops at multimedia and information retrieval conferences [52][53][54]. In addition to its roots in information retrieval, and in particular, multimedia information retrieval, social content search draws on a number of other disciplines. An important contribution is made by social network analysis [55], which concerns itself with analyzing the relationships between actors. Social network analysis places emphasis on connections and exchanges between pairs of people, rather than on groups or hierarchies and was initially developed to study human interaction in the real world. A relatively younger discipline that contributes to social content retrieval is the study of complex networks [56]. Network theory analyzes sets of objects related by pair-wise connections. Complex networks are networks characterized by interesting, non-trivial patterns of connections, a key example being the connectivity patterns that arise in social networks. Social networks can be modeled as graphs in which users are the nodes and interactions between users are the connections between nodes. Communities are defined as sets of nodes that interact less strongly with external nodes than they do with other members of the community. Network structure analysis of social communities reveals patterns of interaction, cf. [57], which can be used to inform the development of social content search techniques.

Within a social community, users play a variety of roles. A recent social-network marketing book [58] lays out a 'Social Technographics' ladder that encodes groups of social-network participants in the order of increasing involvement. The categories are: *inactives, spectators, joiners, collectors, critics and creators*. In the remainder of this section, we examine these categories and discuss the opportunities and the challenges created by the information products and byproducts that members of these categories produce.

4.1 Users as spectators and joiners

Although many users are relatively inactive within networked communities, it is still possible to make use of their contributions. Such users participate as *spectators* (in the sense that they search, browse and view collection content) and as *joiners* (in the sense that they sign up for membership), but their activity does not extend beyond this basic level. Despite their relative inactivity, the search patterns of such users are an important source of information that has been shown to be invaluable for improving web search. Implicit user feedback can be collected in the form of which items in a results list a user chooses to click on. Although a user choice does not always indicate relevance to the query, implicit feedback has been shown to be useful in improving retrieval [59][60][61].

These approaches can be extended to multimedia content, where information can be collected about user search behavior and viewing patterns. Applied to photo search, transaction log analysis and observation of behavior patterns when a user is explicitly refining a query have been shown to be helpful [62][63]. In [64], transaction logs from a large audiovisual archive are used in order to analyze the search behavior of professional searchers. Mining of information on the search trails left through a video collection is investigated by [65]. Users interacting with content contribute 'behavioral semantics,' for example, while viewing a video a viewer might pause, fast-forward, play or stop viewing entirely. This interaction behavior provides information about the relevance or appeal of specific portions of the video [66]. The information that users contribute when quoting a scene from a video, commenting on a scene or reaction to others' comments can be exploited to generate annotations [67]. It should be kept in mind, that even basic information about viewing behavior can be extremely helpful for retrieval. It can be as simple as recording the number of times a given video is viewed cf. 'views' on YouTube. This popularity information can be used to refine result ranking. More complex behavior information must be shown to have a higher utility that popularity in order to justify the effort of collecting it and integrating it into the system.

Additional information helpful for social content retrieval is provided by *joiners* who become part of the network and, in the ideal case, set up explicit connections to other members of the group. Basic information about connections between the members of a social community is valuable for improving social content retrieval. Exploiting social information has been the subject of intense investigation; and early example is [68]. Users connect to each other and via these connections also connect to content. Systems in which users create explicit connections with each other involve defining 'friendship' or 'trust' relationships [69]. Some approaches induce these relationships, when they are not explicit in the data set [70]. Relationships between users are taken to reflect common interests or preferences and are valuable because they can be used to propagate semantics. As suggested by [66], members of specific social groups can be presented with results list generated using the same ranking mechanism. The ranking mechanism takes into account the dominant behavior of the group. *Community profiling* [71] has attracted much interest within the area of multimedia modeling and has a great potential for improving multimedia search.

4.2 Users as collectors, critics and creators

More active users participate in the community as *collectors*, *critics* and *creators*. Typical *collectors* group items into categories and draw attention to or disseminate items of interest, by way of social bookmarking, or, in the case of Twitter users, tweeting or re-tweeting. *Critics* rate items, comment on items, write reviews and, in the case of Epinions.com, rate reviews. *Creators* contribute content to the community, which may include original content, or also co-created content or remixes.

Collaborative recommendation [72][73] makes use of sets of items contained in the item sets of individual users. The items may be items that the user has stored, viewed or purchased (i.e., as in the case of *collectors*) or items that the user has rated (i.e., in the case of *critics*). Recommendation can be considered a form of retrieval in which the user's item set (i.e., user profile) acts as an implicit query. Item-based recommendation, well known for its use on Amazon (http://www.amazon.com/) [74] involves exploitation of item co-occurrences in user profiles. User-based collaborative filtering involves exploited groups of like-minded users, where similarity is determined by user profile similarity. Collaborative filtering approaches can be combined with content analysis, for example in the area of music recommendation [75]. Appropriate combination of collaborative information with user-assigned tags has been demonstrated to be beneficial for the personalization of tagging, browsing and search [76].

Social networks can be represented as graphs or users, but many successful approaches combine users, content and other information into a single structure. In [77], a model is presented that combines users and documents and can be used to guide navigation. A graph consisting of users, items and tags is used by [78] to improve the results for queries consisting of only a single word (i.e., a click on a tag).

Community members create content and contribute it to the social media collection. The challenge of this content is that it is often variable in quality and not adequately annotated. So-called SPUG content (Semi-Professional User Generated content) may be relatively clean and well structured, but nonetheless stands to benefit from improved annotations. Canonical processes of media production [79] can still be expected to apply to the creation of user-generated content. The opportunities that arise to exploit the creation process are many. For example, dependencies between items are generated as part of the content creation process. Users create content sometimes in the form of series or shows and often within the conventions of established genres, such as videos reviewing books and movies or do-it-yourself videos. Their imitation of mainstream media examples creates regularities in form and style throughout the collection. Users also imitate each other's videos and reply to one another in video form. Such interaction leads to the creation of rich conglomerates of related material. Another effect is the development of emergent genres or themes, referred to as *Internet memes*. A classical example is the Diet Coke + Mentos eruption videos on YouTube. Because creators are also viewers, a type of loop emerges within a networked community that channels content into dynamically developing, yet readily identifiable categories. Such categories can be exploited to improve retrieval.

When users create content, that content is often associated with a particular location, collected by the camera or added by the user. In [80], it is shown that combination of tag-based, location-based and content-based analysis makes possible an improved understanding of a social content collection, specifically, Flickr. Other work on Flickr [81] demonstrates that when location information is not explicitly supplied, appropriate techniques can leverage tags to automatically place photos on a map.

Not all users make equal contributions to the network and some even fail to participate in the network in good faith. Here, we cite examples of research from outside multimedia that aims to handle these cases. In [82], the authority status of users is calculated by analyzing the network structure. A method for handling vote spam is presented in [83]. In [84], an approach that simultaneously calculates user reputation and content quality is proposed.

4.3 Social content search & the triple synergy

If multimedia information retrieval is indeed the 'next major frontier of search' [85], then social multimedia retrieval could be considered the leading edge of that frontier. In [86], one of the main technical challenges facing multimedia information retrieval is identified as, *How to Best Combine Human and Machine Intelligence?* This challenge is particularly relevant for social media retrieval, since contributions reflecting human intelligence form the fabric of the social media collection. The *triple synergy* (i.e., user-contributed information, user interactions within the social network and multimedia content analysis) is clearly an opportunity to improve multimedia access in social media communities. In fact, the *triple synergy* can be considered to completely characterize social media communities, and, as such, it is impossible to overemphasize its importance. The challenges in appropriately exploiting the *triple synergy* for retrieval are daunting, but worth the effort facing. We feel that one of the most important challenges is realizing the potential of the networked community to contribute to unlocking meaning in multimedia content. As Ramesh Jain reminds us, '…semantics does not come only from the data or from the user, but it emerges as an interaction of the data and the user.'[87]. Understanding motivations for participation in the networked community is also key. Ultimately, the goal of social multimedia research goes beyond retrieval and encompasses making sense of complete collections, the explicit aim of work such as [80]. Not every user online becomes a part of a networked community. The study reported in [88] finds that most do not. Those that do use available tools form many social connections. Ever-improving exploitation of information derived from networked communities could perhaps encourage a trend towards greater participation.

5. SUMMARY

Networked communities have experienced enormous growth and serve as loci of multimedia collections. Information that derives from these communities, ranging from user-contributed tags to information about user relations through analysis of network structure can be exploited to improve user access to these collections. Three particular areas have been examined in which such information can be applied: annotation, distribution and retrieval. In each of these areas, information derived from networked communities can be used to enhance and improve existing approaches and techniques. In sum, network communities hold within themselves a high potential for improving multimedia access for their users. Participation in a networked community potentially has a high payoff.

ACKNOWLEDGMENTS

This research was supported by the European Commission under contract FP7-216444 PetaMedia. We acknowledge and thank all our colleagues in the PetaMedia Network of Excellence, whose collective efforts have made possible the formulation of the concepts of multimedia retrieval in networked communities presented here.

6. REFERENCES

[1] Boll, S. 2007. MultiTube – Where Web 2.0 and Multimedia Could Meet. IEEE MultiMedia 14, 1 (Jan. 2007), 9-13.

[2] Trant, J. 2009. Studying social tagging and folksonomy: a review and framework. Journal of Digital Information 10,1.

[3] de Jong, F. M. G., Westerveld, T., and de Vries, A.P. 2007. Multimedia search without visual analysis: the value of linguistic and contextual information. IEEE Transactions on Circuits and Systems for Video Technology 17, 3, 365-371.

[4] Smeulders, A. W., Worring, M., Santini, S., Gupta, A., and Jain, R. 2000. Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intell. 22, 12 (Dec. 2000), 1349-1380.

[5] Tsagkias, M., Larson, M., and de Rijke, M. 2008. Term clouds as surrogates for user generated speech. In Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Singapore, Singapore, July 20 - 24, 2008). SIGIR '08. ACM, New York, NY, 773-774.

[6] Zweig, G., Makhoul, J., Stolke, A., August 2006. Introduction to the special section on rich transcription. IEEE Trans. On Audio, Speech, and Lang. Proc. 14, 5, 1490-1491.

[7] Larson, M., Newman, E., Jones, G. 2009. Overview of VideoCLEF 2008: Automatic generation of topic-based feeds for dual language audio-visual content. In Peters, C. et al., (eds.) Evaluating Systems for Multilingual and Multimodal Information Access, CLEF 2008 Revised Selected Papers, LNCS 5706, 906-917.

[8] Hauptmann, G., Christel, M. G., and Yan, R. 2008. Video retrieval based on semantic concepts. Proceedings of the IEEE 96, 4, (April 2008), 602-622.

[9] Snoek, C.G.M. and Worring, M. 2009. Concept-based video retrieval. Foundations and Trends in Information Retrieval 2, 4, (2009), 215-322.

[10] Mylonas, P., Spyrou, E., Avrithis, Y, and Kollias, S. 2009. Using visual context and region semantics for high-level concept detection. IEEE Trans. on Multimedia 1, 2 (Feb. 2009) 229-243.

[11] Lin, Y.-R., Sundaram, H., De Choudhury, M., and Kelliher, A. 2009. Temporal patterns in social media streams: theme discovery and evolution using joint analysis of content and context. In Proceedings of the Int. Conf. on Multimedia and Expo (New York, NY, Jun./Jul. 2009) ICME 2009.

[12] R. Glasberg, S. Schmiedeke, P. Kelm, T. Sikora. 2008. An automatic system for real-time video-genres detection using high-level descriptions and a set of classifiers. Proceedings of the 12th IEEE Int. Symp. on Consumer Electronics (Algarve, Portugal, April 2008) ISCE 2008.

[13] Li, J. and Wang, J.Z. 2008. Real-time computerized annotation of pictures. IEEE Trans. on Pattern Analysis and Machine Intelligence 30, 6 (June 2008), 985-1002.

[14] Cai, R., Lu, L., Hanjalic, A., Zhang, H.-J., and Cai, L.-H. 2006. A flexible framework for key audio effects detection and auditory context inference. IEEE Trans. on Audio, Speech, and Language Processing 14, 3, 1026-1039.

[15] Wactlar, H. D., Hauptmann, A. G., Christel, M. G., Houghton, R. A., and Olligschlaeger, A.M. 2000. Complementary video and audio analysis for broadcast news archives. Comm. of the ACM 43, 2 (Feb. 2000), 42-47.

[16] Hsu, W. H., Kennedy, L. S., and Chang, S. 2006. Video search reranking via information bottleneck principle. In Proceedings of the 14th Annual ACM International Conf. on Multimedia (Santa Barbara, CA, USA, October 23 - 27, 2006). MULTIMEDIA '06. ACM, New York, NY, 35-44.

[17] Hanjalic, A. April 2006. Extracting moods from pictures and sounds: towards truly personalized TV. Signal Processing Magazine, IEEE 23, 2, 90-100.

[18] Cavet, R., Volmer, S., Leopold, E., Kindermann, J., and Paass, G. 2004. Revealing the connoted visual code: a new approach to video classification. Computers & Graphics 28, 3, 361-369.

[19] Sharma, A., Hua, G., Liu, Z. and Zhang, Z. 2008. Meta-tag propagation by co-training an ensemble classifier for improving image search relevance, Proceedings of the IEEE CVPR Workshop on Internet Vision (Anchorage AK, USA, June 2008). CVPR 2008.

[20] Siersdorfer, S., San Pedro, J., and Sanderson, M. 2009. Automatic video tagging using content redundancy. In Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (Boston, MA, USA, July 19 - 23, 2009). SIGIR '09. ACM, New York, NY, 395-402.

[21] Li, X., Snoek, C.G.M. and Worring, M. 2009. Learning social tag relevance by neighbor voting, IEEE Trans. On Multimedia 11, 7 (Nov. 2009), 1310-1322.

[22] Cattuto, C., Benz, D., Hotho, A., and Stumme, G. 2008. Semantic analysis of tag similarity measures in collaborative tagging systems. In Proc. of the 3rd Workshop on Ontology Learning and Population. (Patras, Greece, July 2008) OLP3.

[23] Kliegr, T., Chandramouli, K., Nemrava, J., Svatek, V. and Izquierdo, E. 2008. Combining image captions and visual analysis for image concept classification. In Proceedings of the 9th Int. Workshop on Multimedia Data Mining. (Las Vegas, Aug 2008). MDM/KDD 2008.

[24] Clements, M., de Vries, A. P., and Reinders, M. J. 2008. Detecting synonyms in social tagging systems to improve content retrieval. In Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Singapore, Singapore, July 20 - 24, 2008). SIGIR '08. ACM, New York, NY, 739-740.

[25] Garg, N. and Weber, I. 2008. Personalized, interactive tag recommendation for Flickr. In Proceedings of the 2008 ACM Conference on Recommender Systems (Lausanne, Switzerland, October 23 - 25, 2008). RecSys '08. ACM, New York, NY, 67-74.

[26] Sigurbjörnsson, B. and van Zwol, R. 2008. Flickr tag recommendation based on collective knowledge. In Proceeding of the 17th International Conference on World Wide Web (Beijing, China, April 21 - 25, 2008). WWW '08. ACM, New York, NY, 327-336.

[27] Marlow, C., Naaman, M., Boyd, D., and Davis, M. 2006. Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead. In Proc. of the Collab. Web Tagging Workshop at WWW'06.

[28] Von Ahn, L. and Dabbish, L. 2004. Labeling images with a computer game. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA). CHI '04. ACM Press, 319–326.

[29] Von Ahn, L., Liu, R., and Blum, M. 2006. Peekaboom: a game for locating objects in images. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (New York, NY, USA). CHI '06. ACM, 55–64.

[30] Von Ahn, Ginosar, S., Kedia, M., Liu, R. and Blum, M. 2006. Improving accessibility of the web with a computer game. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (New York, NY, USA, 2006). CHI '06. ACM, 79–82.

[31] Russell, B.C., Torralba, A., Murphy, K.P., and Freeman, W.T. 2008. LabelMe: A database and web-based tool for image annotation. Int. J. Comput. Vision 77, 1-3, 157–173.

[32] Ogata, J. and Goto, M. 2009. PodCastle: Collaborative training of acoustic models on the basis of wisdom of crowds for podcast transcription. In Proceedings of Interspeech 2009 — 10th Annual Conference of the International Speech Communication Association, 1491-1494.

[33] Petridis, S. and Pantic M. 2008. Audiovisual laughter detection based on temporal features. In Proceedings of the Int. Conf. on Multimodal Interfaces, (Chania, Crete, Greece, October 2008), 37-44.

[34] Soleymani, M., Chanel, G., Kierkels J., and Pun, T. 2008. Affective characterization of movie scenes based on multimedia content analysis and user's physiological emotional responses. In Proceedings of the Int. Symposium on Multimedia (Berkeley, CA, Dec. 15-17, 2008).

[35] Yazdani, A., Lee, J.-S., and Ebrahimi, T. 2009. Implicit emotional tagging of multimedia using EEG signals and brain computer interface. In Proceedings of the ACM Multimedia Workshop on Social Media. (Beijing, China).

[36] Koelstra, S., Muehl, C., and Patras, I. 2009. EEG analysis for implicit tagging of video data. In Proc. of the Workshop on Affective Brain-Computer Interfaces, Proc. ACII, 27-32.

[37] Gallagher, A. and Chen, T. 2009. Understanding Images of Groups of People, IEEE Conference on Computer Vision and Pattern Recognition 2009.

[38] Melenhorst, M., Grootveld, M., van Setten, M., and Veenstra, M. 2008. Tag-based information retrieval of video content. In *Proceeding of the 1st international Conference on Designing interactive User Experiences For TV and Video* (Silicon Valley, CA, Oct. 22-24, 2008). UXTV '08, vol. 291.

[39] Nguyen, T., Kolazhi, K. Kamath, R., Cheung, S.S. and Tran, D.A. 2008. Efficient multimedia distribution in source constraint networks. IEEE Trans. on Multimedia 10, 3 (Apr. 2008), 523-537.

[40] Androutsellis-Theotokis, S. and Spinellis, D. 2004. A survey of peer-to-peer content distribution technologies. ACM Computing Surveys 36, 4 (2004) 335-371.

[41] Ramzan, N., Zgaljic, T., and Izquierdo, E. 2009. An efficient optimisation scheme for scalable surveillance centric video communications. Signal Processing: Image Communication 24, 6 (July 2009), 510-523.

[42] Ramzan N. and Izquierdo, E. 2009. Scalable video coding as basis for networked media internet. In Proceedings of the European Commission NEM Summit Towards Future Media Internet. (Saint-Malo, France, Sept. 2009).

[43] Ramzan, N. and Izquierdo, E. 2009. An optimised error protection scheme for wavelet based scalable video. In Proc. of Latin American Conference on Networked Electronic Media (Bogota, Columbia, Aug. 2009). LACNEM.

[44] Baccichet, P., Schierl, T., Wiegand T., and Girod, B. 2007. Low-delay peer-to-peer streaming using scalable video coding. In Proceedings of the 16th International Packet Video Workshop (Lausanne, Switzerland, Nov. 2007).

[45] Li, H. and Chen, G. 2007. Data persistence in structured P2P networks with redundancy schemes. In Proceedings of the Sixth International Conference of Grid and Cooperative Computing (Los Alamitos, CA, 2007). GCC.

[46] Palomar, E., Estevez-Tapiador, J. M., Hernandez-Castro, J. C., and Ribagorda, A. 2006. A protocol for secure content distribution in pure P2P networks. In Proceedings of the 17th International Conference on Database and Expert Systems Applications. DEXA'06. 712-716.

[47] De Simone, F., Naccari, M., Tagliasacchi, M., Dufaux, F., Tubaro, S., and Ebrahimi, T. 2009. Subjective assessment of H.264/AVC video sequences transmitted over a noisy channel. In Proceedings of the First International Workshop on Quality of Multimedia Experience (2009). QoMEX..

[48] Mathieu, B. 2008. A network and QoE aware P2P video streaming solution. In Proc. of the 8th Int. Conference on New Technologies in Distributed Systems (Lyon, France, June 23 - 27, 2008). NOTERE '08. ACM, NY, NY, 1-9.

[49] Smith, M., Barash, V., Getoor, L., and Lauw, H. W. 2008. Leveraging social context for searching social media. In Proceedings of the 2008 ACM Workshop on Search in Social Media (Napa Valley, California, USA, October 26- 30, 2008). SSM '08. ACM, New York, NY, 91-94.

[50] Shirky, C., 2008. Here Comes Everybody: The Power of Organizing Without Organizations. The Penguin Press HC.

[51] Evans, B. M. and Chi, E. H. 2008. Towards a model of understanding social search. In Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work (San Diego, CA, USA, November 08 - 12, 2008). CSCW '08. ACM, New York, NY, 485-494.

[52] Agichtein, E., Hearst, M. A., and Soboro, I. 2009. The Search and Social Media Workshop at SIGIR 2009. SIGIR Forum 43, 2 (Dec. 2009), 53-56.

[53] Marchand-Maillet, S., Vries, A. P., and Naaman, M. 2009. Workshop on Information Retrieval over Social Networks. In Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval (Toulouse, France, April 06 - 09, 2009) 803.

[54] Boll, S., Hoi, S. C., Luo, J., Jin, R., King, I., and Xu, D. 2009. First ACM SIGMM International Workshop on Social Media. WSM'09. In Proceedings of the 17th ACM International Conference on Multimedia (Beijing, China, October 19 - 24, 2009), MM '09. ACM, New York, NY, 1161-1162.

[55] Haythornthwaite, C. 1996. Social network analysis: An approach and technique for the study of information exchange. Library & Information Science Research. 18, 4 (Autumn 1996), 323-342.

[56] Newman, M. E. J. 2003. The structure and function of complex networks, SIAM Review 45, 167-256.

[57] Leskovec, J., Lang, K. J., Dasgupta, A., and Mahoney, M. W. 2008. Statistical properties of community structure in large social and information networks. In Proceedings of the 17th International Conference on World Wide Web (Beijing, China, April 21 - 25, 2008). WWW '08. ACM, New York, NY, 695-704.

[58] Li, C. and Bernoff, J. 2008. Groundswell: Winning in a World Transformed by Social Technologies. 1st Edition. Harvard Business School Press.

[59] Agichtein, E., Brill, E., and Dumais, S. 2006. Improving web search ranking by incorporating user behavior information. In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Seattle, Washington, USA, August 06 - 11, 2006). SIGIR '06. ACM, New York, NY, 19-26.

[60] Xue, G., Zeng, H., Chen, Z., Yu, Y., Ma, W., Xi, W., and Fan, W. 2004. Optimizing web search using web click-through data. In Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management (Washington, D.C., USA, November 08 - 13, 2004). CIKM '04. ACM, New York, NY, 118-126.

[61] Craswell, N. and Szummer, M. 2007. Random walks on the click graph. In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Amsterdam, Netherlands, July 23 - 27, 2007). SIGIR '07. ACM, New York, NY, 239-246.

[62] Müller, H., Müller, W., Squire, D., Marchand-Maillet, S., and Pun, T. 2000. Long-term learning from user behavior in content-based image retrieval. Tech. report. University of Geneva 2000.

[63] Marchand-Maillet, S. and Bruno, E. 2003. Exploiting user interaction for semantic content-based image retrieval. Tech. report. Computer Vision and Multimedia Laboratory, Computing Centre, University of Geneva.

[64] Huurnink, B., Hollink, L., van den Heuvel, W., and de Rijke. M. 2010. The Search Behavior of Media Professionals at an Audiovisual Archive: A Transaction Log Analysis. Journal of the American Society for Information Science and Technology. 2010.

[65] Hopfgartner, F., Vallet, D., Halvey, M., and Jose, J. 2008. Search trails using user feedback to improve video search. In Proceedings of the 16th ACM International Conference on Multimedia (Vancouver, British Columbia, Canada, October 26 - 31, 2008). MM '08. ACM, New York, NY, 339-348.

[66] Davis, S. J., Ritz, C. H., and Burnett, I. S. 2009. Using social networking and collections to enable video semantics acquisition. IEEE MultiMedia 16, 4, 52-61.

[67] Yamamoto, D., Masuda, T., Ohira, S., and Nagao, K. 2008. Video scene annotation based on web social activities. IEEE MultiMedia 15, 3 (Jul. 2008), 22-32.

[68] Shardanand U. and Maes P. 1995. Social information filtering: algorithms for automating 'word of mouth'. In Proceedings of SIGCHI, (Denver, Colorado, 1995), 210–217.

[69] Kazai, G. and Milic-Frayling, N. 2008. Trust, authority and popularity in social information retrieval. In Proceedings of the 17th ACM Conference on Information and Knowledge Management (Napa Valley, California, USA, October 26 - 30, 2008). CIKM '08. ACM, New York, NY, 1503-1504.

[70] Golbeck, J. and Hendler, J. 2006. Inferring binary trust relationships in web-based social networks. ACM Trans. Internet Technol. 6, 4 (Nov. 2006), 497-529.

[71] Daylamani Zad, D. and Agius, H. 2009. Multimedia Metadata 2.0: Challenges of collaborative content modeling. Advances in Semantic Media Adaptation and Personalization 2, M.C. Angelides, M. Wallace and P. Mylonas, Eds. CRC Press, Boca Raton, FL, USA, 1-20.

[72] Resnick P., Iacovou N., Suchak M., Bergstorm P., and Riedl J. Grouplens: An open architecture for collaborative filtering of netnews. In the Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work, (Chapel Hill, North, Carolina, 1994) ACM, 175–186.

[73] Adomavicius, G. and Tuzhilin, A. 2005. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Transactions on Knowledge and Data Engineering (Jun. 2005), 734-749.

[74] Linden, G., Smith, B., and York, J. 2003. Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Computing 7, 1 (Jan./Feb. 2003), 76-80.

[75] Yoshii, K., Goto, M., Komatani, K., Ogata, T., and Okuno, H. G. 2008. An efficient hybrid music recommender system using an incrementally trainable probabilistic generative model. Audio, Speech, and Language Processing, IEEE Transactions on. 16, 2, 435-447.

[76] Wang J., Clements M., Yang, J., de Vries, A. P. and Reinders, M. J. T. 2010. Personalization of tagging systems. Information Processing and Management 46, 58-70.

[77] Marchand-Maillet, S., Szekely, E., and Bruno, E. 2009. Optimizing strategies for the exploration of social-networks and associated data collections. In Proceedings of the International Workshop on Image Analysis for Multimedia Interactive Services. (London, UK). WIAMIS'09.

[78] Clements, M., de Vries, A. P., and Reinders, M. J. T. 2008. Optimizing single term queries using a personalized Markov random walk over the social graph. Proceedings of the Workshop on Exploiting Semantic Annotations in Information Retrieval (March 2008). ESAIR.

[79] Hardman, L., Obrenović, Ž., Nack, F., Kerhervé, B., Piersol, K. 2008. Canonical processes of semantically annotated media production. Multimedia Systems,14, 6, (December 2008) 327-340.

[80] Kennedy, L., Naaman, M., Ahern, S., Nair, R., and Rattenbury, T. 2007. How Flickr helps us make sense of the world: context and content in community-contributed media collections. In Proceedings of the 15th International Conference on Multimedia (Augsburg, Germany, September 25 - 29, 2007). MULTIMEDIA '07. ACM, New York, NY, 631-640.

[81] Serdyukov, P., Murdock, V., and van Zwol, R. 2009. Placing Flickr photos on a map. In Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (Boston, MA, USA, July 19 - 23, 2009). SIGIR '09. ACM, New York, NY, 484-491.

[82] Kirsch, S. M., Gnasa, M., Won, M., and Cremers, A. B. 2007. From page rank to social rank: Authority-based retrieval in social information spaces. In Social information retrieval systems: Emerging technologies and applications for searching the web effectively. D. H.Gohand S. Foo (eds.), Idea Group Inc., 2007.

[83] Bian, J., Liu, Y., Agichtein, E., and Zha, H. 2008. A few bad votes too many?: Towards robust ranking in social media. In Proceedings of the 4th International Workshop on Adversarial Information Retrieval on the Web (Beijing, China, April 22 - 22, 2008). C. Castillo, K. Chellapilla, and D. Fetterly, (eds.) AIRWeb '08, 295. ACM, NY, NY, 53-60.

[84] Bian, J., Liu, Y., Zhou, D., Agichtein, E., and Zha, H. 2009. Learning to recognize reliable users and content in social media with coupled mutual reinforcement. In Proceedings of the 18th International Conference on World Wide Web (Madrid, Spain, April 20 - 24, 2009). WWW '09. ACM, New York, NY, 51-60.

[85] Lew, M. S., Sebe, N., Djeraba, C., and Jain, R. 2006. Content-based multimedia information retrieval: state of the art and challenges. ACM Trans. Multimedia Comput. Commun. Appl. 2, 1, 1-19.

[86] Kankanhalli, M. S., and Rui, Y. March 2008. Application potential of multimedia information retrieval. Proceedings of the IEEE 96, 4, 712-720.

[87] Jain, R. 2008. Multimedia information retrieval: watershed events. In Proceeding of the 1st ACM International Conference on Multimedia Information Retrieval (MIR '08) ACM, New York, NY, USA, 229-236.

[88] Halvey, M. J. and Keane, M. T. 2007. Exploring social dynamics in online media sharing. In Proceedings of the 16th International Conference on World Wide Web (Banff, Alberta, Canada, May 08 - 12, 2007). WWW '07. ACM, New York, NY, 1273-1274.